

Olasılık Teorisi

Açık Matematik

Açık Matematik — açık kaynaklı, reklamsız Türkçe matematik notları
acik-matematik.com

Bu çalışma [Creative Commons BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/) lisansı ile paylaşılmıştır: kaynak göstererek ve aynı lisansla, ticari olmayan amaçlarla kopyalayabilir, dağıtabilir ve uyarlayabilirsiniz.

Bu sürüm: 20 Eylül 2026

İçindekiler

Giriş	9
1 Rastgele Deney, Örnek Uzay ve Olaylar	11
1.1 Rastgele Deney ve Örnek Uzay	11
1.2 Olaylar Üzerinde İşlemler	13
1.3 Sınıflar ve Kuvvet Kümesi	16
1.4 Alıştırmalar	18
2 Cebirler, σ-Cebirler ve Borel Cebri	21
2.1 Cebirler	21
2.2 σ -Cebirler	21
2.3 σ -Cebirlerin Temel Özellikleri	24
2.4 σ -Cebirlerin Arakesiti ve Üretilen σ -Cebir	25
2.5 Borel Cebri	28
2.6 Alıştırmalar	31
3 Olasılık Ölçüsü ve Olasılık Uzayı	35
3.1 Kolmogorov Aksiyomları	35
3.2 Olasılık Ölçüsünün Temel Özellikleri	36
3.3 Toplama Kuralı ve İçerme–Dışarma Formülü	37
3.4 Boole Eşitsizliği	41
3.5 Olasılığın Sürekliliği	42
3.6 Alıştırmalar	45
4 Olasılık Uzayı Örnekleri ve Geometrik Olasılık	49
4.1 Sonlu Örnek Uzaylar	49
4.2 Sayılabilir Örnek Uzaylar	52
4.3 Gerçel Sayılar Üzerinde Olasılık Ölçüleri	53
4.4 Ölçü ve Geometrik Olasılık	59
4.5 Alıştırmalar	64
5 Sayma Teknikleri	67
5.1 Toplama ve Çarpma Kuralları	67
5.2 Permütasyon	69
5.3 Kombinasyon	70
5.4 Tekrarlı Permütasyon ve Çok Terimli Katsayı	71
5.5 Binom Teoremi	73
5.6 Tekrarlı Kombinasyon: Yıldız–Çubuk Yöntemi	74

5.7	Karma Sayma Problemleri	77
5.8	Laplace Uzayında Olasılık Hesapları	78
5.9	Alıştırmalar	79
6	Koşullu Olasılık, Toplam Olasılık ve Bayes Teoremi	81
6.1	Koşullu Olasılık	81
6.2	Çarpma Teoremi	86
6.3	Sonlu Parçalanma ve Toplam Olasılık Formülü	87
6.4	Bayes Teoremi	89
6.5	Alıştırmalar	93
7	Bağımsız Olaylar	95
7.1	Bağımsızlığın Tanımı	95
7.2	Bağımsız ve Ayrık Olaylar	96
7.3	Tümleyenlerin Bağımsızlığı	98
7.4	Koşullu Olasılıkların Eşitliği ve Bağımsızlık	99
7.5	Çoklu Bağımsızlık	100
7.6	Bağımsızlık Hesapları	104
7.7	Alıştırmalar	105
8	Rastgele Değişken	109
8.1	Rastgele Değişkenin Tanımı	109
8.2	İlk Örnekler	111
8.3	Hangi Fonksiyonlar Rastgele Değişkendir?	114
8.4	Borel Ölçütü	115
8.5	Rastgele Değişkenin Olasılık Dağılımı	117
8.6	Alıştırmalar	119
9	Dağılım Fonksiyonu	123
9.1	Dağılım Fonksiyonunun Tanımı	123
9.2	Dağılım Fonksiyonu Örnekleri	123
9.3	Dağılım Fonksiyonunun Temel Özellikleri	132
9.4	Dağılım Fonksiyonundan Olasılık Hesabı	133
9.5	Dağılım Fonksiyonu Olasılık Dağılımını Belirler	136
9.6	Alıştırmalar	137
10	Kesikli Rastgele Değişkenler ve Olasılık Fonksiyonu	141
10.1	Kesikli Rastgele Değişken	141
10.2	Olasılık Fonksiyonu	141
10.3	Olasılık Fonksiyonu ile Dağılım Fonksiyonu	143
10.4	Örnekler	147
10.5	Alıştırmalar	153

11	Sürekli Rastgele Değişkenler ve Yoğunluk Fonksiyonu	157
11.1	Sürekli Rastgele Değişken ve Yoğunluk Fonksiyonu	157
11.2	Dağılım Fonksiyonunun Sürekliliği ve Nokta Olasılıkları	158
11.3	Hangi Fonksiyonlar Yoğunluk Olabilir?	162
11.4	Yoğunluğun Dağılım Fonksiyonundan Elde Edilmesi	163
11.5	Örnekler	165
11.6	Alıştırmalar	170
12	Rastgele Değişkenin Fonksiyonunun Dağılımı	173
12.1	Rastgele Değişkenin Fonksiyonu	173
12.2	Kesikli Rastgele Değişkenin Dönüşümü	174
12.3	Dağılım Fonksiyonu Tekniği	176
12.4	Dönüşüm Kesikli mi, Sürekli mi?	180
12.5	Dönüşüm Teoremi	181
12.6	Bir Yoğunluğun Baştan Sona İncelenmesi	185
12.7	Alıştırmalar	187
13	Rastgele Vektörler ve Ortak Dağılımlar	191
13.1	Rastgele Vektör ve Olasılık Dağılımı	191
13.2	Ortak Dağılım Fonksiyonu	193
13.3	Kesikli Rastgele Vektörler	195
13.4	Sürekli Rastgele Vektörler	199
13.5	Alıştırmalar	204
14	Bağımsız Rastgele Değişkenler ve Koşullu Dağılımlar	207
14.1	Bağımsız Rastgele Değişkenler	207
14.2	Koşullu Dağılımlar	210
14.3	Bağımsızlık ile Koşullu Dağılım Arasındaki İlişki	213
14.4	Kesikli Örnekler	214
14.5	Sürekli Örnekler	219
14.6	Alıştırmalar	223
15	Rastgele Vektörlerin Dönüşümleri	225
15.1	Kesikli Rastgele Vektörlerin Dönüşümleri	225
15.2	Sürekli Rastgele Vektörler: Jacobian Yöntemi	228
15.3	Tek Bir Dönüşüm: Yapay Değişken ve Dağılım Fonksiyonu Tekniği	231
15.4	Bağımsız Değişkenlerin Toplamı: Konvolüsyon	234
15.5	İki Kapsamlı Örnek	235
15.6	Alıştırmalar	239
16	Beklenen Değer, Momentler ve Varyans	243
16.1	Bir Rastgele Değişkenin Fonksiyonunun Beklenen Değeri	243
16.2	Momentler, Varyans ve Standart Sapma	246
16.3	Beklenen Değerin Lineerliği	247

16.4	Varyansın Hesaplanması	249
16.5	Örnekler	250
16.6	Markov ve Chebyshev Eşitsizlikleri	253
16.7	Alıştırmalar	256
17	Karakteristik ve Moment Üreten Fonksiyonlar	259
17.1	Karakteristik Fonksiyon	259
17.2	Moment Üreten Fonksiyon	265
17.3	Çarpımsal Moment Üreten Fonksiyon	269
17.4	Bir Dönüşümün Karakteristik Fonksiyonu	271
17.5	Alıştırmalar	273
18	Rastgele Vektörlerde Beklenen Değer, Kovaryans ve Korelasyon	275
18.1	Rastgele Vektörün Fonksiyonunun Beklenen Değeri	275
18.2	Kovaryans ve Korelasyon Katsayısı	279
18.3	Lineer Birleşimin Varyansı	286
18.4	Çözümlü Örnekler	287
18.5	Alıştırmalar	292
19	Koşullu Beklenen Değer	295
19.1	Koşullu Beklenen Değerin Tanımı	295
19.2	Koşullu Beklenen Değerin Özellikleri	299
19.3	Rastgele Değişken Olarak Koşullu Beklenen Değer	301
19.4	Koşullu Varyans ve Toplam Varyans Kuralı	303
19.5	Rastgele Sayıda Terimin Toplamı	305
19.6	En İyi Tahmin Olarak Koşullu Beklenen Değer	307
19.7	Alıştırmalar	308
20	Bernoulli ve Binom Dağılımları	311
20.1	Bernoulli Deneyi ve Bernoulli Dağılımı	311
20.2	Binom Dağılımı	312
20.3	Beklenen Değer ve Varyans	315
20.4	Moment Üreten Fonksiyon ve Toplamlar	317
20.5	Uygulama: Kalite Kontrol	320
20.6	En Olası Değer	321
20.7	Alıştırmalar	323
21	Poisson Dağılımı	327
21.1	Tanım ve Olasılık Fonksiyonu	327
21.2	Beklenen Değer ve Varyans	330
21.3	Moment Üreten ve Karakteristik Fonksiyon	331
21.4	Bağımsız Poisson Değişkenlerinin Toplamı	332

21.5	Binom Dağılımına Poisson Yaklaşımı	334
21.6	Alıştırmalar	338
22	Geometrik, Negatif Binom ve Hipergeometrik Dağılımlar . . .	341
22.1	Kesikli Düzgün Dağılım	341
22.2	Geometrik Dağılım	342
22.3	Negatif Binom Dağılımı	347
22.4	Hipergeometrik Dağılım	350
22.5	Kesikli Dağılımlar Arasındaki İlişkiler	354
22.6	Alıştırmalar	355
23	Düzgün, Üstel ve Normal Dağılımlar	359
23.1	Sürekli Düzgün Dağılım	359
23.2	Üstel Dağılım	362
23.3	Normal Dağılım	366
23.4	Alıştırmalar	373
23.5	Olasılık Teorisi'nin Sonu	374

Giriş

Olasılık teorisi, belirsizliği ölçülebilir bir matematiksel nesneye dönüştürür. Bir deneyin olası sonuçlarını bir küme (Ω) olarak, sorulabilecek soruları bu kümenin altkümeleri (olaylar) olarak, “ne kadar olası” sorusunun yanıtını ise bu altkümelere sayı atayan bir ölçü olarak yazar. Kolmogorov’un üç aksiyomuyla kurulan bu yapı, rastgele değişkenler, dağılım fonksiyonları, beklenen değer ve üretici fonksiyonlara uzanır; matematiksel istatistiğin ve stokastik süreçlerin ortak temelidir.

Bu notlar bir dönemlik Olasılık Teorisi dersini beş kısımda izler. Her kavram tanım kutusuyla verilmiş, her teorem ispatlanmış, her örnek sonuna kadar çözülmüştür; kesikli ve sürekli hâller her zaman yan yana ele alınır.

i Not

Olasılık Teorisi notlarının tamamı hazırdır ve aşağıdaki bağlantılardan ulaşılabilir.

1. Rastgele Deney, Örnek Uzay ve Olaylar

Bir zar attığımızda hangi yüzün üste geleceğini önceden bilemeyiz. Yine de bu işlem hakkında söyleyebileceğimiz kesin şeyler vardır: olabilir sonuçlar 1, 2, 3, 4, 5, 6'dır; zar düzgünse hiçbir yüzün ötekine üstünlüğü yoktur; zarı çok kez atarsak altı gelme sıklığının $\frac{1}{6}$ dolayında toplandığını gözleriz. **Olasılık teorisi**, sonucu önceden bilinmeyen olguları bu türden kesin ifadelerle betimleyen ve bu ifadelerden ispatlanabilir sonuçlar çıkaran matematik dalıdır. Bunun için üç nesne kurulur: olabilir sonuçların kümesi olan **örnek uzay**, bu kümenin **olay** adı verilen alt kümeleri ve her olaya, olayın **olasılığı** denen, 0 ile 1 arasında bir sayı karşılık getiren bir fonksiyon. Bu üçlünün uyması gereken kurallar 1933'te Andrey Kolmogorov tarafından birkaç aksiyomla belirlenmiştir; teoremin bütün sonuçları bu aksiyomlardan türetilir. Teori bir kez kurulduktan sonra "zarı çok kez atınca sıklık $\frac{1}{6}$ 'ya yaklaşır" türünden gözlemler birer sezgi olmaktan çıkar, ispatlanan teoremlere dönüşür.

Kitap beş kısımdan oluşur. **Birinci kısımda** olasılık uzayı kurulur: örnek uzay ve olaylar, olayları bir arada tutan σ -cebirlere, Kolmogorov aksiyomları, sonlu ve geometrik örnek uzaylar, sayma teknikleri, koşullu olasılık, Bayes teoremi ve bağımsızlık. **İkinci kısımda** deneyin sonucuna bir sayı karşılık getiren rastgele değişkenler ile onların dağılım fonksiyonları, kesikli ve sürekli halleriyle incelenir. **Üçüncü kısım** birden çok rastgele değişkeni birlikte ele alan rastgele vektörlere, ortak ve koşullu dağılımlara ve bunların dönüşümlerine ayrılmıştır. **Dördüncü kısımda** beklenen değer, varyans, kovaryans ve korelasyon ile moment üreten ve karakteristik fonksiyonlar gelir. **Beşinci kısım** uygulamada en sık karşılaşılan dağılımları, Bernoulli, binom, Poisson, geometrik, negatif binom, hipergeometrik, düzgün, üstel ve normal dağılımları, toplu olarak inceler.

Bu ilk bölümde yapının en alt katını döşüyoruz. Rastgele deney, örnek uzay ve olay kavramlarını tanımlayacak; sonlu, sayılabilir sonsuz ve sayılamaz örnek uzaylara birer örnek verecek; olaylar arasında "ve", "veya", "değil" bağlaçlarıyla kurulan işlemleri küme işlemlerine çevirecek ve bu işlemlerin uyduğu kuralları ispatlayacağız. Bölümün sonunda olayların toplandığı **sınıf** ve **kuvvet kümesi** kavramları gelir; bunlar, bir sonraki bölümde tanımlanacak olan σ -cebirin hazırlığıdır.

1.1 Rastgele Deney ve Örnek Uzay

Günlük dilde "deney" denince laboratuvar aklı gelir; olasılık teorisinde sözcük çok daha geniş anlamda kullanılır. Zar atmak, bir madeni parayı fırlatmak, yarın öğlen sıcaklığı ölçmek, bir ampulün ömrünü gözlemek, üretim bandından rastgele seçilen bir parçanın kusurlu olup olmadığına bakmak: bunların hepsi bir deneydir. Ortak yanları şudur: **ne olabileceğini** önceden biliriz, **ne olacağını** bilmeyiz.

Tanım 1.1 (Rastgele Deney). Olabilir sonuçlarının kümesi önceden belli olan, ancak hangi sonucun ortaya çıkacağı önceden söylenemeyen bir işleme **rastgele deney** (rastgele sonuçlu deney) ya da kısaca **deney** denir.

Tanımda iki koşul vardır: olabilir sonuçların kümesi belli olmalı ve hangi sonucun çıkacağı önceden söylenememelidir. Bir zarın atılması ikisini de sağlar. Deniz düzeyinde suyun kaç derecede kaynadığını ölçmek ikinci koşulu sağlamaz: sonuç önceden bilinir. "Bir kişinin aklına gelen ilk şeyi söylemesi" ise birinci koşulu sağlamaz: olabilir sonuçların kümesi belli değildir. Olasılık teorisinin dışında kalanlar, birinci koşulu sağlamayan işlemlerdir. Sonucu önceden bilinen işlemler rastgele deney değildir; yine de tek sonuçlu bir örnek uzayla teoremin diline çevrilebilir ve yozlaşmış birer özel hâl olarak hiçbir güçlük çıkarmazlar.

Tanım 1.2 (Örnek Uzay). Bir rastgele deneyin bütün olabilir sonuçlarının kümesine deneyin **örnek uzayı** (örneklem uzayı) denir ve Ω ile gösterilir. Ω 'nın her ω elemanına deneyin bir **sonucu** ya da bir **örnek noktası** denir.

Deney yapıldığında Ω 'nın tam olarak bir elemanı ortaya çıkar: sonuçlar birbirini dışlar ve Ω olabilir sonuçların hepsini kapsar. Örnek uzay sonlu, sayılabilir sonsuz ya da sayılamaz olabilir; aşağıdaki üç örnek sırasıyla bu üç durumu gösterir.

Tanım 1.3 (Olay). Ω örnek uzayının her alt kümesine bir **olay** denir. Deneyin sonucu ω olduğunda, $\omega \in A$ ise **A olayı gerçekleşti**, $\omega \notin A$ ise **A olayı gerçekleşmedi** denir. Özel olarak Ω 'ya **kesin olay**, $\emptyset$ 'e **imkânsız olay**, tek elemanlı $\{\omega\}$ kümelerine **basit olay** denir.

Kesin olay her deneyde gerçekleşir, çünkü sonuç ne olursa olsun Ω 'nın elemanıdır; imkânsız olay hiçbir zaman gerçekleşmez, çünkü hiçbir sonuç $\emptyset$ 'in elemanı değildir. Bir **sonuç** ile bir **basit olay** arasındaki farka dikkat edelim: ω örnek uzayın bir elemanıdır, $\{\omega\}$ ise örnek uzayın bir alt kümesidir. Olasılık ölçüsü olaylar, yani kümeler üzerinde tanımlanacağı için “ ω sonucunun olasılığı” denince kastedilen her zaman $\{\omega\}$ basit olayının olasılığıdır.

Tanımda “her alt kümesi” dedik. Sonlu örnek uzaylarda bu kararın hiçbir sakıncası yoktur. Sayılamaz örnek uzaylarda ise, ileride görüleceği gibi, her alt kümeye aralık uzunluğuyla uyumlu tutarlı bir olasılık atamak mümkün olmaz; bu yüzden hangi alt kümelerin olay sayılacağı ayrıca seçilir. Bu seçimin uyması gereken kuralları σ -cebiri kavramıyla (Tanım 2.2) verilir. Bu bölümde bütün alt kümeleri olay saymaya devam edeceğiz.

Örnek 1.1 (Zar Atma Deneyi ve Olayları). Düzgün bir tavla zarının atılması ve üste gelen yüzün gözlenmesi bir rastgele deneydir. Olabilir sonuçlar altı yüzdür; örnek uzay

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

kümesidir. Bu deneyle ilgili olaylardan bazıları şunlardır:

$A = \{1\}$: bir gelmesi olayı,

$B = \{1, 3, 5\}$: tek sayı gelmesi olayı,

$C = \{1, 2, 3\}$: dörtten küçük bir sayı gelmesi olayı,

$D = \{6\}$: altı gelmesi olayı,

$G = D^c = \{1, 2, 3, 4, 5\}$: altı gelmemesi olayı,

$H = A \cup D = \{1, 6\}$: bir **veya** altı gelmesi olayı,

$I = B \cap C = \{1, 3\}$: tek **ve** dörtten küçük bir sayı gelmesi olayı.

Zar 3 gelirse B , C , G ve I olayları gerçekleşir; A , D ve H gerçekleşmez. Zar 6 gelirse yalnız D ve H gerçekleşir. A ve D basit olaylardır; H , bu iki basit olayın birleşimidir. Ω 'nın $2^6 = 64$ alt kümesi vardır (Önerme 1.2) ve bunların her biri bu deneyin bir olayıdır; yukarıdakiler bunlardan yalnızca yedisidir.

i Örnek uzay bir modelleme seçimidir

Aynı fiziksel işlem için birden çok örnek uzay yazılabilir; seçim, neyi gözlediğimize bağlıdır. Zar örneğinde yalnızca gelen sayının tek mi çift mi olduğuyla ilgileniyorsak $\Omega' = \{\text{tek}, \text{çift}\}$ de bir örnek uzaydır. Ama bu uzayda “bir gelmesi” bir olay değildir; sorulabilecek sorular kabalaşmıştır. Kural şudur: örnek uzay, yanıtını merak ettiğimiz her sorunun bir olaya karşılık geleceği kadar ince seçilmelidir. Aşağıdaki sıcaklık örneğinde -10 ile 35 sınırları da bir seçimdir; $\Omega = \mathbb{R}$ almak da mümkündür.

Örnek 1.2 (Tura Gelinceye Kadar Para Atma Deneyi). Bir madeni paranın tura gelinceye kadar atılması deneyini ele alalım; yazı Y , tura T ile gösterilsin. Deney ilk turada biter ve sonuç, o ana kadar gözlenen dizidir:

$$\Omega = \{T, YT, YYT, YYYYT, YYYYYT, \dots\}.$$

n -inci sonucu ω_n ile gösterelim: ω_n , $n - 1$ tane yazının ardından bir turadan oluşan dizidir; $\omega_1 = T$, $\omega_2 = YT$, $\omega_3 = YYT$ ve böyle sürer. Böylece $\Omega = \{\omega_1, \omega_2, \omega_3, \dots\}$ **sayılabilir sonsuz** bir kümedir. Bu deneyle ilgili olaylardan bazıları şunlardır:

$A = \{T\} = \{\omega_1\}$: ilk atışta tura gelmesi olayı,

$B = \{T, YT, YYT\} = \{\omega_1, \omega_2, \omega_3\}$: dördüncü atıştan önce tura gelmesi olayı,

$C = \{\omega_2, \omega_4, \omega_6, \dots\}$: turanın çift sayılı bir atışta gelmesi olayı,

$B^c = \{\omega_4, \omega_5, \omega_6, \dots\}$: en az dört atış gerekmesi olayı.

Ω sonsuz olduğundan bu deneyle ilgili sonsuz tane olay tanımlanabilir. Örneğin her $n \geq 2$ için “tura n -inci atışta gelir” basit olayı $\{\omega_n\}$ ile “tura en geç n -inci atışta gelir” olayı $\{\omega_1, \dots, \omega_n\}$ birbirinden farklı olaylardır ve n değiştikçe her biri değişir. ($n = 1$ için bu iki olay çıkarır: ikisi de $\{\omega_1\}$ 'dir.)

Örnek uzayı yazarken paranın hiç tura gelmemesi, yani sonsuz yazı dizisi, dışarıda bırakıldı. Bu bir seçimdir; istenirse bu dizi de Ω 'ya bir sonuç olarak eklenebilir. Olasılık ölçüsü kurulduğunda bu sonucun olasılığının sıfır olduğu görülecektir; bu yüzden onu dışarıda bırakmak hiçbir hesabı değiştirmez.

Örnek 1.3 (Sıcaklık Gözleme Deneyi). Yarın saat 12'de hava sıcaklığının santigrat derece olarak gözlenmesi bir rastgele deneydir. Sıcaklığın -10 ile 35 derece arasında kalacağını varsayarsak örnek uzay

$$\Omega = \{t : -10 < t < 35\} = (-10, 35)$$

aralığıdır. Bu küme **sayılamazdır**. Bu deneyle ilgili olaylardan bazıları şunlardır:

$A = \{t : 15 < t < 25\} = (15, 25)$: sıcaklığın 15 ile 25 derece arasında olması olayı,

$B = \{t : t = 10\} = \{10\}$: sıcaklığın tam 10 derece olması olayı,

$C = \{t : t \leq 0\} = (-10, 0]$: sıcaklığın sıfırın altında ya da sıfır olması olayı,

$A^c = (-10, 15] \cup [25, 35)$: sıcaklığın 15 ile 25 derece arasında olmaması olayı.

Ω 'nın her alt aralığı, her tek elemanlı alt kümesi ve aralıkların her birleşimi bir olaydır; dolayısıyla bu deneyle ilgili de sonsuz tane olay vardır. $A \cap B = \emptyset$ olduğuna dikkat edelim: sıcaklık aynı anda hem 10 derece hem de 15 ile 25 arasında olamaz.



Şekil 1.1: Üç tür örnek uzay. Solda iki zar atıldığında sonuç bir sayı çifti, örnek uzay ise $6 \times 6 = 36$ noktalık sonlu bir kümedir. Ortada para tura gelinceye kadar atılır; sonuçlar $\omega_1 = T, \omega_2 = YT, \omega_3 = YYT, \dots$ diye sayılarak sonsuza dek sürer, örnek uzay sayılabilir sonsuzdur. Sağda yarınki sıcaklık $(-10, 35)$ aralığındaki herhangi bir gerçel sayı olabilir; bu örnek uzay sayılamazdır ve olaylar aralıklar ile onların birleşimleridir.

1.2 Olaylar Üzerinde İşlemler

Zar örneğinde “bir veya altı gelmesi” ile “tek ve dörtten küçük bir sayı gelmesi” olaylarını yazarken bağlaçları küme işlemlerine çevirdik: “veya” birleşim, “ve” kesişim oldu; “altı gelmemesi” ise tümleyendi. Bu çeviri, olaylar hakkında konuşurken kullandığımız dilin tamamını küme kuramına taşır. Şimdi bunu genel olarak yapalım.

Tanım 1.4 (Olaylar Üzerinde İşlemler). Ω örnek uzayında A ve B iki olay olsun.

(i) **Birleşim.** $A \cup B = \{\omega \in \Omega : \omega \in A \text{ veya } \omega \in B\}$ olayına “**A veya B** olayı” denir. $A \cup B$ 'nin gerçekleşmesi, A ile B 'den **en az birinin** gerçekleşmesi demektir.

(ii) **Kesişim.** $A \cap B = \{\omega \in \Omega : \omega \in A \text{ ve } \omega \in B\}$ olayına “**A ve B** olayı” denir. $A \cap B$ 'nin gerçekleşmesi, A ile B 'nin **ikisinin birden** gerçekleşmesi demektir.

(iii) **Tümleyen.** $A^c = \Omega \setminus A = \{\omega \in \Omega : \omega \notin A\}$ olayına “**A değil** olayı” ya da A 'nın **tümleyeni** denir. A^c 'nin gerçekleşmesi, A 'nın **gerçekleşmemesi** demektir.

(iv) **Fark.** $A \setminus B = \{\omega \in \Omega : \omega \in A \text{ ve } \omega \notin B\}$ olayına A ile B 'nin **farkı** denir. $A \setminus B$ 'nin gerçekleşmesi, A 'nın gerçekleşip B 'nin gerçekleşmemesi demektir.

(v) **Ayrık olaylar.** $A \cap B = \emptyset$ ise A ile B 'ye **ayrık** (bağdaşmaz) olaylar denir: ikisi birden gerçekleşemez.

(vi) **Gerektirme.** $A \subseteq B$ ise “ A olayı B olayını **gerektirir**” denir: A gerçekleştiğinde B de gerçekleşir.

Sayılabılır bir olay ailesi $A_1, A_2, A_3, \dots$ için

$$\bigcup_{n=1}^{\infty} A_n = \{\omega \in \Omega : \text{en az bir } n \text{ için } \omega \in A_n\}, \quad \bigcap_{n=1}^{\infty} A_n = \{\omega \in \Omega : \text{her } n \text{ için } \omega \in A_n\}$$

olayları sırasıyla “ A_n ’lerden **en az biri**” ve “ A_n ’lerin **hepsi**” olaylarıdır.

Sözlük biçiminde toplarsak:

Dilde	Küme olarak	Gerçekleşme koşulu
A veya B	$A \cup B$	en az biri gerçekleşir
A ve B	$A \cap B$	ikisi birden gerçekleşir
A değil	A^c	A gerçekleşmez
A ama B değil	$A \setminus B$	A gerçekleşir, B gerçekleşmez
A ile B birlikte olamaz	$A \cap B = \emptyset$	ayrık olaylar
A ise B	$A \subseteq B$	A gerçekleşince B de gerçekleşir
A_n ’lerden en az biri	$\bigcup_n A_n$	bir n için A_n gerçekleşir
A_n ’lerin hepsi	$\bigcap_n A_n$	her n için A_n gerçekleşir

Bu sözlüğün değeri şuradadır: olasılık ölçüsünün aksiyomları yalnız birleşim ve tümleyen üzerinden yazılacak, ama “en az biri”, “hepsi”, “hiçbiri”, “tam olarak biri” gibi bütün günlük ifadeler bu tabloya bakılarak kümelere çevrilebilecektir.

⚠ Ayrık olaylar ile tümleyen aynı şey değildir

Zar deneyinde $A = \{1\}$ ve $D = \{6\}$ ayrık olaylardır: zar aynı anda hem bir hem altı gelemmez. Ama A, D ’nin tümleyeni değildir; $D^c = \{1, 2, 3, 4, 5\}$ kümesi A ’dan çok daha büyüktür. Tümleyen olmak, ayrık olmanın ötesinde birleşimin bütün örnek uzayı doldurmasını ister: $A \cap B = \emptyset$ **ve** $A \cup B = \Omega$ ise $B = A^c$ ’dir. Ayrık olaylar için yalnız birinci koşul vardır. Zar 3 geldiğinde ne A ne D gerçekleşir; oysa bir olayla tümleyeninden her deneyde tam olarak biri gerçekleşir.

Olaylar birer küme olduğundan küme işlemlerinin bütün kuralları olaylar için de geçerlidir. Bu kuralları burada bir araya toplayıp ispatlıyoruz, çünkü sonraki bölümlerde σ -cebiri özellikleri (Teorem 2.1) ve olasılık ölçüsünün özellikleri (Teorem 3.1) ispatlanırken hepsine sürekli başvurulacaktır; özellikle De Morgan kuralları, “birleşim altında kapalı” olmakla “kesişim altında kapalı” olmak arasında köprü kuracaktır. İspatlar tek bir yöntemle dayanır: iki kümenin eşit olduğunu göstermek için keyfi bir ω alınır ve ω ’nın birinci kümede olmasının ikinci kümede olmasıyla eşdeğer olduğu gösterilir.

Önerme 1.1 (Olay Cebirinin Özellikleri). Ω örnek uzayında A, B, C olaylar ve $A_1, A_2, A_3, \dots$ bir olay dizisi olsun.

(a) **Değişme:** $A \cup B = B \cup A$ ve $A \cap B = B \cap A$.

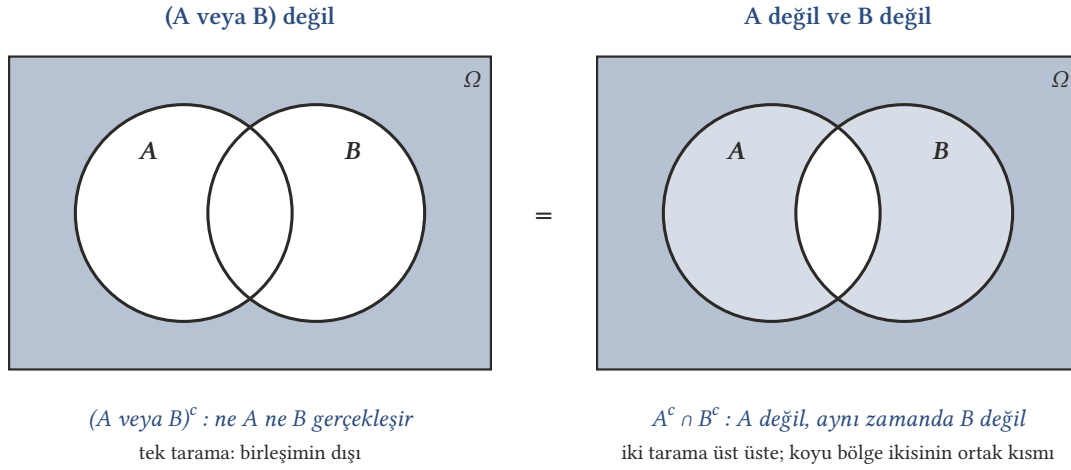
(b) **Birleşme:** $(A \cup B) \cup C = A \cup (B \cup C)$ ve $(A \cap B) \cap C = A \cap (B \cap C)$.

(c) **Dağılma:** $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ ve $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

(d) **Tümleyen:** $(A^c)^c = A$, $A \cup A^c = \Omega$, $A \cap A^c = \emptyset$, $\Omega^c = \emptyset$ ve $\emptyset^c = \Omega$.

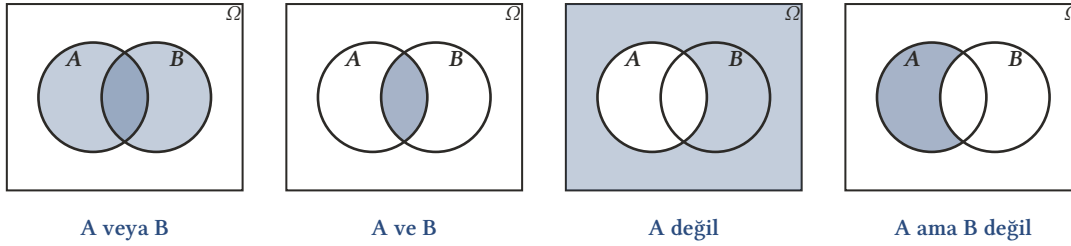
(e) **De Morgan kuralları:** $(A \cup B)^c = A^c \cap B^c$ ve $(A \cap B)^c = A^c \cup B^c$; daha genel olarak

$$\left(\bigcup_{n=1}^{\infty} A_n\right)^c = \bigcap_{n=1}^{\infty} A_n^c, \quad \left(\bigcap_{n=1}^{\infty} A_n\right)^c = \bigcup_{n=1}^{\infty} A_n^c.$$



Şekil 1.2: De Morgan kuralı $(A \cup B)^c = A^c \cap B^c$. Solda birleşimin tümleyeni bir kerede taranmıştır: ne A ne B gerçekleşir. Sağda önce A^c, sonra B^c ayrı ayrı taranmış, iki taramanın üst üste geldiği koyu bölge ortaya çıkmıştır; bu bölge soldakiyle aynıdır. Kural, "veya"nın değilini "ve"nin değillerine çevirir; $(A \cap B)^c = A^c \cup B^c$ biçimi de aynı yolla okunur.

(f) **Ayrıştırma:** $A \setminus B = A \cap B^c$; $A = (A \cap B) \cup (A \setminus B)$ ve $A \cup B = A \cup (B \setminus A)$ eşitliklerinde sağ taraftaki iki olay ayrıktır.



Şekil 1.3: Olay işlemlerinin Venn şemaları. Dikdörtgen örnek uzay Ω 'yı, daireler A ve B olaylarını gösterir. Soldan sağa: birleşim $A \cup B$, kesişim $A \cap B$, tümleyen A^c ve fark $A \setminus B = A \cap B^c$ (A'nın gerçekleşip B'nin gerçekleşmediği sonuçlar).

İspat.

Bütün ispatlarda $\omega \in \Omega$ keyfî bir sonuçtur.

(a) " $\omega \in A$ veya $\omega \in B$ " önermesi ile " $\omega \in B$ veya $\omega \in A$ " önermesi aynı şeyi söyler; dolayısıyla $\omega \in A \cup B \Leftrightarrow \omega \in B \cup A$ olur. "Ve" bağlacı için aynı akıl yürütme kesşimi verir.

(b) $\omega \in (A \cup B) \cup C$ olması, ω 'nın $A \cup B$ kümesinde ya da C'de olması, yani A, B, C kümelerinden en az birinde bulunması demektir. $\omega \in A \cup (B \cup C)$ olması da tam olarak bunu söyler. Kesişim için: $\omega \in (A \cap B) \cap C$ olması, ω 'nın üç kümenin her birinde bulunması demektir; $\omega \in A \cap (B \cap C)$ de aynı koşuldur.

(c) Birinci eşitlik. $\omega \in A \cap (B \cup C)$ olsun. O hâlde $\omega \in A$ ve $\omega \in B \cup C$ 'dir; ikincisi $\omega \in B$ ya da $\omega \in C$ demektir. $\omega \in B$ ise $\omega \in A \cap B$, $\omega \in C$ ise $\omega \in A \cap C$ olur; her iki hâlde de $\omega \in (A \cap B) \cup (A \cap C)$ 'dir. Tersine $\omega \in (A \cap B) \cup (A \cap C)$ olsun. $\omega \in A \cap B$ ise $\omega \in A$ ve $\omega \in B \subseteq B \cup C$ 'dir; $\omega \in A \cap C$ ise $\omega \in A$ ve $\omega \in C \subseteq B \cup C$ 'dir. Her iki hâlde de $\omega \in A \cap (B \cup C)$ olur.

İkinci eşitlik. $\omega \in A \cup (B \cap C)$ olsun. $\omega \in A$ ise ω hem $A \cup B$ hem $A \cup C$ kümesindedir. $\omega \in B \cap C$ ise $\omega \in B \subseteq A \cup B$ ve $\omega \in C \subseteq A \cup C$ 'dir. Her iki hâlde de $\omega \in (A \cup B) \cap (A \cup C)$ olur. Tersine $\omega \in (A \cup B) \cap (A \cup C)$ olsun. $\omega \in A$ ise iş bitmiştir. $\omega \notin A$ ise $\omega \in A \cup B$ olması $\omega \in B$ 'yi, $\omega \in A \cup C$ olması $\omega \in C$ 'yi gerektirir; dolayısıyla $\omega \in B \cap C \subseteq A \cup (B \cap C)$ olur.

(d) Tanım gereği $\omega \in (A^c)^c \Leftrightarrow \omega \notin A^c \Leftrightarrow \omega \in A$; yani $(A^c)^c = A$. Her $\omega \in \Omega$ için “ $\omega \in A$ ” ile “ $\omega \notin A$ ” önermelerinden tam olarak biri doğrudur. En az biri doğru olduğundan her ω , $A \cup A^c$ kümesindedir: $A \cup A^c = \Omega$. İkisi birden doğru olamadığından hiçbir ω , $A \cap A^c$ kümesinde değildir: $A \cap A^c = \emptyset$. Son olarak $\Omega^c = \{\omega \in \Omega : \omega \notin \Omega\} = \emptyset$ ve $\emptyset^c = \{\omega \in \Omega : \omega \notin \emptyset\} = \Omega$ ’dır.

(e) Önce genel biçimin birincisini gösterelim:

$$\omega \in \left(\bigcup_{n=1}^{\infty} A_n \right)^c \Leftrightarrow \omega \notin \bigcup_{n=1}^{\infty} A_n \Leftrightarrow \text{hiçbir } n \text{ için } \omega \in A_n \text{ değil} \Leftrightarrow \text{her } n \text{ için } \omega \in A_n^c \Leftrightarrow \omega \in \bigcap_{n=1}^{\infty} A_n^c.$$

İkincisi birincisinden çıkar: birinci kuralı A_n^c dizisine uygulayıp (d)’deki $(A_n^c)^c = A_n$ eşitliğini kullanırsak

$$\left(\bigcup_{n=1}^{\infty} A_n^c \right)^c = \bigcap_{n=1}^{\infty} (A_n^c)^c = \bigcap_{n=1}^{\infty} A_n$$

olur; iki yanın tümleyenini alıp yine (d)’yi kullanınca $\bigcup_{n=1}^{\infty} A_n^c = \left(\bigcap_{n=1}^{\infty} A_n \right)^c$ elde edilir. İki olay için verilen kurallar bunun özel hâlidir: dizide $A_1 = A$ ve her $n \geq 2$ için $A_n = B$ alınırsa $\bigcup_n A_n = A \cup B$ ve $\bigcap_n A_n = A \cap B$ olur; genel kurallar da $(A \cup B)^c = A^c \cap B^c$ ve $(A \cap B)^c = A^c \cup B^c$ biçimini alır. (f) $\omega \in A \setminus B$ olması, $\omega \in A$ ve $\omega \notin B$, yani $\omega \in A$ ve $\omega \in B^c$ olması demektir; bu da $\omega \in A \cap B^c$ ’dir. $A = (A \cap B) \cup (A \setminus B)$ için: $\omega \in A$ olsun. $\omega \in B$ ise $\omega \in A \cap B$, $\omega \notin B$ ise $\omega \in A \setminus B$ olur; her hâlde ω sağ taraftadır. Tersine, sağ taraftaki iki küme de A ’nın alt kümesidir. İki parça ayrıktır, çünkü bir ω aynı anda hem B ’de hem B ’nin dışında olamaz.

$A \cup B = A \cup (B \setminus A)$ için: $B \setminus A \subseteq B$ olduğundan sağ taraf sol tarafın alt kümesidir. Tersine $\omega \in A \cup B$ olsun. $\omega \in A$ ise sağ taraftadır. $\omega \notin A$ ise $\omega \in B$ olmak zorundadır ve $\omega \notin A$ olduğundan $\omega \in B \setminus A$ ’dır. A ile $B \setminus A$ ayrıktır, çünkü $B \setminus A \subseteq A^c$ ’dir.

■

Önermenin (f) şıkkındaki iki ayrıştırma ileriye sık kullanılacaktır: bir olayı, başka bir olayla kesişen ve kesişmeyen parçalarına ayırmak, olasılık hesaplarında toplama kuralının (Teorem 3.2) ve toplam olasılık formülünün (Teorem 6.4) çıkış noktasıdır.

Kuralları somut bir örnekte doğrulayalım.

Örnek 1.4 (Zar Deneyinde De Morgan Kurallarının Doğrulanması). Örnek 1.1’ndeki $B = \{1, 3, 5\}$ (tek sayı) ve $C = \{1, 2, 3\}$ (dörtten küçük sayı) olayları için $(B \cup C)^c = B^c \cap C^c$ ve $(B \cap C)^c = B^c \cup C^c$ eşitliklerini kümeleri açıkça yazarak doğrulayınız; $B \setminus C$ ve $C \setminus B$ olaylarını bulup sözle ifade ediniz.

Çözüm.

Önce tümleyenleri yazalım: $B^c = \{2, 4, 6\}$ (çift sayı gelmesi) ve $C^c = \{4, 5, 6\}$ (dörtten küçük olmayan sayı gelmesi).

Birinci kural. $B \cup C = \{1, 2, 3, 5\}$ olduğundan $(B \cup C)^c = \{4, 6\}$ ’dır. Öte yandan $B^c \cap C^c = \{2, 4, 6\} \cap \{4, 5, 6\} = \{4, 6\}$ ’dir. İki küme eşittir. Sözle: “tek ya da dörtten küçük olmaması”, “çift ve dörtten küçük olmaması” ile aynı olaydır; her ikisi de zarın 4 ya da 6 gelmesi demektir.

İkinci kural. $B \cap C = \{1, 3\}$ olduğundan $(B \cap C)^c = \{2, 4, 5, 6\}$ ’dır. Öte yandan $B^c \cup C^c = \{2, 4, 6\} \cup \{4, 5, 6\} = \{2, 4, 5, 6\}$ ’dır. İki küme yine eşittir. Sözle: “hem tek hem dörtten küçük olmaması”, “çift ya da dörtten küçük olmaması” ile aynı olaydır.

Farklar. $B \setminus C = B \cap C^c = \{1, 3, 5\} \cap \{4, 5, 6\} = \{5\}$: tek olup dörtten küçük olmayan sayı gelmesi, yani beş gelmesi olayı. $C \setminus B = C \cap B^c = \{1, 2, 3\} \cap \{2, 4, 6\} = \{2\}$: dörtten küçük olup tek olmayan sayı gelmesi, yani iki gelmesi olayı. Ayrıştırma kuralına göre $B = (B \cap C) \cup (B \setminus C) = \{1, 3\} \cup \{5\}$ olmalıdır; gerçekten de bu birleşim $\{1, 3, 5\} = B$ ’dir ve parçalar ayrıktır.

■

1.3 Sınıflar ve Kuvvet Kümesi

Bir deneyle ilgilendiğimiz olaylar genellikle tek tek değil, topluca ele alınır: olasılık ölçüsü, olayların hepsi üzerinde birden tanımlanan bir fonksiyon olacaktır. Bu yüzden olayları bir arada tutan nesneye bir ad vermek gerekir.

Tanım 1.5 (Sınıf). Ω örnek uzayının bazı alt kümelerinden oluşan boş olmayan bir aileye Ω üzerinde bir **sınıf** (olaylar sınıfı, olaylar ailesi) denir ve $\mathcal{U}$ ile gösterilir. Bir sınıfın elemanları Ω 'nın alt kümeleri, yani olaylardır.

Zar deneyinde $\mathcal{U}_1 = \{B, B^c\} = \{\{1, 3, 5\}, \{2, 4, 6\}\}$ yalnız tek-çift olaylarından oluşan bir sınıftır; $\mathcal{U}_2 = \{\emptyset, D, D^c, \Omega\}$ altı gelip gelmemesiyle ilgili olayların sınıfıdır; tek elemanlı $\mathcal{U}_3 = \{A\}$ da bir sınıftır. Sınıf tanımında hiçbir kapalılık koşulu yoktur: $\mathcal{U}_1$, elemanlarının birleşimi olan Ω 'yı içermez ve yine de bir sınıftır. Kapalılık koşulları σ -cebiri tanımıyla (Tanım 2.2) gelecektir.

💡 Üç düzey: sonuç, olay, sınıf

Bu bölümde iç içe üç düzey vardır ve karıştırılmaları en sık yapılan hatadır. En altta **sonuçlar** vardır: $\omega \in \Omega$. Bir üstte **olaylar** vardır: $A \subseteq \Omega$, yani $\omega \in A$. En üstte **sınıflar** vardır: $A \in \mathcal{U}$. Zar deneyinde 3 bir sonuçtur, $\{1, 3, 5\}$ bir olaydır, $\{\{1, 3, 5\}, \{2, 4, 6\}\}$ bir sınıftır. Bir sonuç bir olayın **elemanıdır**; bir olay bir sınıfın **elemanı**, örnek uzayın ise **alt kümesidir**. “ $3 \in \mathcal{U}_1$ ” ya da “ $\{1, 3, 5\} \subseteq \mathcal{U}_1$ ” yazmak yanlıştır: bu iki önerme iyi tanımlıdır, ama ikisi de sağlanmaz; doğru yazım $\{1, 3, 5\} \in \mathcal{U}_1$ 'dir.

Tanım 1.6 (Kuvvet Kümesi). Ω kümesinin bütün alt kümelerinin oluşturduğu sınıfa Ω 'nın **kuvvet kümesi** denir ve

$$\mathcal{P}(\Omega) = \{A : A \subseteq \Omega\}$$

ile gösterilir.

Kuvvet kümesi Ω üzerindeki en büyük sınıftır: her sınıf için $\mathcal{U} \subseteq \mathcal{P}(\Omega)$ olur. $\emptyset \subseteq \Omega$ ve $\Omega \subseteq \Omega$ olduğundan $\emptyset \in \mathcal{P}(\Omega)$ ve $\Omega \in \mathcal{P}(\Omega)$ 'dır. Örneğin $\Omega = \{1, 2, 3\}$ için

$$\mathcal{P}(\Omega) = \{ \emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\} \}$$

sekiz elemanlı bir sınıftır. Burada $\emptyset$ ile $\{\emptyset\}$ arasındaki fark önemlidir: $\{\emptyset\}$, tek elemanı imkânsız olay olan bir sınıftır; boş küme değildir.

Sonlu bir örnek uzayın kaç olayı ve kaç sınıfı olduğunu kesin olarak sayabiliriz. Bu sayı, sonlu örnek uzaylarda “bütün alt kümeleri olay saymanın” ne kadar cömert bir karar olduğunu gösterir ve iki elemanlı bir örnek uzayın üzerinde bile on beş farklı sınıf kurulabildiğini görmemizi sağlar.

Önerme 1.2 (Kuvvet Kümesinin Eleman Sayısı ve Sınıf Sayısı). Ω , k elemanlı sonlu bir küme olsun; yani $|\Omega| = k$.

(a) Ω 'nın 2^k tane alt kümesi vardır: $|\mathcal{P}(\Omega)| = 2^k$.

(b) Ω üzerinde $2^{2^k} - 1$ tane farklı sınıf tanımlanabilir.

İspat.

(a) k üzerinden tümevarım yapalım. $k = 0$ ise $\Omega = \emptyset$ 'dir ve tek alt kümesi $\emptyset$ 'dir; $2^0 = 1$ ile uyuşur. Şimdi k elemanlı her kümenin 2^k alt kümesi olduğunu varsayalım ve Ω , $k + 1$ elemanlı olsun. Ω 'dan bir x elemanı seçip $\Omega' = \Omega \setminus \{x\}$ diyelim; Ω' kümesi k elemanlıdır, dolayısıyla varsayım gereği 2^k alt kümesi vardır. Ω 'nın her A alt kümesi iki türden biridir:

$x \notin A$ ise $A \subseteq \Omega'$ olur; bu türden alt kümeler tam olarak Ω' kümesinin alt kümeleridir ve sayıları 2^k 'dir. $x \in A$ ise $A' = A \setminus \{x\}$ kümesi Ω' kümesinin bir alt kümesidir ve $A = A' \cup \{x\}$ olarak geri kazanılır. Tersine, her $A' \subseteq \Omega'$ için $A' \cup \{x\}$ kümesi x 'i içeren bir alt kümedir. Öyleyse $A \mapsto A \setminus \{x\}$ eşlemesi, x 'i içeren alt kümeler ile Ω' kümesinin alt kümeleri arasında birebir ve örten bir eşlemedir; bu türden alt kümelerin sayısı da 2^k 'dir.

Toplam $2^k + 2^k = 2^{k+1}$ alt küme vardır; tümevarım tamamlanır. Sonucun sezgisel nedeni şudur: bir alt küme kurmak, Ω 'nın her elemanı için “alınсын” ya da “alınmasın” diye karar vermek demektir; k eleman için birbirinden bağımsız k ikili seçim, 2^k farklı alt küme üretir.

(b) Bir sınıf, tanımı gereği, Ω 'nın alt kümelerinden oluşan boş olmayan bir ailedir; başka bir deyişle $\mathcal{P}(\Omega)$ 'nın boş olmayan bir alt kümesidir. (a) gereği $\mathcal{P}(\Omega)$ kümesi 2^k elemanlıdır; (a)'yı bu kez $\mathcal{P}(\Omega)$ kümesine uygularsak $\mathcal{P}(\Omega)$ 'nın 2^{2^k} tane alt kümesi olduğu bulunur. Bunlardan tam olarak biri, hiçbir alt küme içermeyen boş ailedir ve boş aile bir sınıf değildir. Geriye $2^{2^k} - 1$ sınıf kalır.

■

Formüldeki -1 'in nedeni ispatta görüldü: $\mathcal{P}(\Omega)$ 'nın alt kümeleri arasında boş aile de vardır, ama sınıf tanımı boş olmayan aileleri kapsar. Bu ayrım yalnız bir sayma inceliği değildir. Boş aile hiçbir olay içermez; üzerinde olasılık ölçüsü tanımlanacak tek bir olay bile sunmaz. Öte yandan $\{\emptyset\}$ sınıfı boş **değildir**: tek elemanı imkânsız olaydır. Sayılan $2^{2^k} - 1$ sınıfın içinde $\{\emptyset\}$ vardır, boş aile yoktur.

Sınıf sayısı örnek uzayın büyüklüğüyle çok hızlı büyür:

$ \Omega = k$	$ \mathcal{P}(\Omega) = 2^k$	sınıf sayısı $2^{2^k} - 1$
1	2	3
2	4	15
3	8	255
4	16	65 535
6	64	$2^{64} - 1 \approx 1,8 \cdot 10^{19}$

Altı elemanlı zar uzayı üzerinde bile sınıf sayısı yirmi basamaklıdır. Bir sonraki bölümde bu sınıfların arasından, olasılık ölçüsü tanımlamaya elverişli olanları, σ -cebirleri seçeceğiz; onların sayısı çok daha azdır. Şimdi en küçük ilginç durumu, iki elemanlı örnek uzayı, tam olarak inceleyelim.

Örnek 1.5 (İki Elemanlı Örnek Uzay Üzerindeki Bütün Sınıflar). $\Omega = \{a, b\}$ olsun. Ω üzerinde tanımlanabilecek bütün sınıfları belirleyiniz.

Çözüm.

Önce kuvvet kümesini yazalım. Ω 'nın $2^2 = 4$ alt kümesi vardır:

$$\mathcal{P}(\Omega) = \{ \emptyset, \{a\}, \{b\}, \Omega \}.$$

Bir sınıf, bu dört olaydan oluşan boş olmayan bir ailedir; **Önerme 1.2** gereği $2^4 - 1 = 15$ tane sınıf vardır. Bunları eleman sayısına göre sıralayarak listeleyelim.

Tek olaylı sınıflar (dört tane):

$$\mathcal{U}_1 = \{\emptyset\}, \quad \mathcal{U}_2 = \{\{a\}\}, \quad \mathcal{U}_3 = \{\{b\}\}, \quad \mathcal{U}_4 = \{\Omega\}.$$

İki olaylı sınıflar (altı tane; dört olaydan ikisini seçmenin altı yolu vardır):

$$\mathcal{U}_5 = \{\emptyset, \{a\}\}, \quad \mathcal{U}_6 = \{\emptyset, \{b\}\}, \quad \mathcal{U}_7 = \{\emptyset, \Omega\},$$

$$\mathcal{U}_8 = \{\{a\}, \{b\}\}, \quad \mathcal{U}_9 = \{\{a\}, \Omega\}, \quad \mathcal{U}_{10} = \{\{b\}, \Omega\}.$$

Üç olaylı sınıflar (dört tane; her biri, dört olaydan birinin dışarıda bırakılmasıyla elde edilir):

$$\mathcal{U}_{11} = \{\emptyset, \{a\}, \{b\}\}, \quad \mathcal{U}_{12} = \{\emptyset, \{a\}, \Omega\},$$

$$\mathcal{U}_{13} = \{\emptyset, \{b\}, \Omega\}, \quad \mathcal{U}_{14} = \{\{a\}, \{b\}, \Omega\}.$$

Dört olaylı sınıf (bir tane):

$$\mathcal{U}_{15} = \{\emptyset, \{a\}, \{b\}, \Omega\} = \mathcal{P}(\Omega).$$

Toplam $4 + 6 + 4 + 1 = 15$ sınıf bulundu; on altıncı aile olan boş aile sınıf sayılmadığından liste tamdır.

Bu on beş sınıfın tam olarak sekizi Ω 'yı içerir: $\mathcal{U}_4, \mathcal{U}_7, \mathcal{U}_9, \mathcal{U}_{10}, \mathcal{U}_{12}, \mathcal{U}_{13}, \mathcal{U}_{14}$ ve $\mathcal{U}_{15}$. Bunun nedeni, Ω 'yı içeren bir ailenin geri kalan üç olaydan istenen herhangi bir alt kümesiyle tamamlanabilmesidir: $2^3 = 8$ seçenek.

Bu on beş sınıftan yalnız ikisi, $\mathcal{U}_7 = \{\emptyset, \Omega\}$ ile $\mathcal{U}_{15} = \mathcal{P}(\Omega)$, bir sonraki bölümde tanımlanacak σ -cebir koşullarını sağlar. Örneğin $\mathcal{U}_{12} = \{\emptyset, \{a\}, \Omega\}$ sınıfı, $\{a\}$ olayını içerdiği hâlde onun tümleyeni olan $\{b\}$ olayını içermediği için elenecektir.

■

Sonlu bir örnek uzayda bütün alt kümeleri olay saymanın bir sakıncası yoktur; sıcaklık deneyindeki gibi sayılamaz bir örnek uzayda ise durum değişir.

1.4 Alıştırmalar

Alıştırma 1.1 (Rastgele Deney, Örnek Uzay ve Olaylar Üzerine). a) Biri 1 liralık, öteki 50 kuruşluk iki madeni para birlikte atılıyor ve her birinin üste gelen yüzü gözleniyor. Örnek uzayı yazınız. E_1 en az bir tura gelmesi, E_2 tam olarak bir tura gelmesi, E_3 hiç tura gelmemesi olayları olsun. E_1, E_2, E_3, E_1^c ve $E_1 \setminus E_2$ olaylarını küme olarak yazınız. E_2, E_1 'i gerektirir mi? Bu deneyin kaç olayı ve örnek uzayı üzerinde kaç sınıf vardır?

- b) **Örnek 1.1**'ndeki $B = \{1, 3, 5\}$, $C = \{1, 2, 3\}$ ve $D = \{6\}$ olayları için $B \cup D$, $B^c \cap C$, $(B \cup D)^c$, $C^c \setminus D$ ve $B \cap D$ olaylarını küme olarak yazıp sözle ifade ediniz. B ile D ayrık mıdır? D olayı C^c olayını gerektirir mi?
- c) **Örnek 1.2**'daki gösterimle E turanın en erken üçüncü atışta gelmesi, F turanın tek sayılı bir atışta gelmesi ve her $n \in \mathbb{N}$ için G_n turanın en geç n -inci atışta gelmesi olayı olsun. E , F ve G_n olaylarını küme olarak yazınız. $E = (\{\omega_1\} \cup \{\omega_2\})^c$, $F^c = C$, $G_n = \bigcup_{k=1}^n \{\omega_k\}$, $\bigcup_{n=1}^{\infty} G_n = \Omega$ ve $\bigcap_{n=1}^{\infty} G_n^c = \emptyset$ olduğunu gösteriniz.
- d) **Örnek 1.3**'ndeki $A = (15, 25)$ ve $B = \{10\}$ olayları ile yeni bir $K = \{t : t \geq 20\}$ olayı verilsin. $A \cap K$, $A \cup K$, $A \setminus K$, $K \setminus A$, $A^c \cap K^c$ ve $A \cap B$ olaylarını aralık olarak yazıp sözle ifade ediniz.
- e) Ω örnek uzayında A ve B olaylar olsun. Şunları ispatlayınız: (i) $A \subseteq B$, $A \cap B = A$ ve $A \cup B = B$ önermeleri birbirine denktir; (ii) $A \subseteq B$ ise $B^c \subseteq A^c$ 'dir; (iii) $A \setminus B$, $A \cap B$ ve $B \setminus A$ olayları ikiye ikiye ayrıktır ve $A \cup B = (A \setminus B) \cup (A \cap B) \cup (B \setminus A)$ 'dır.
- f) A , B , C üç olay olsun. “En az biri gerçekleşir”, “hepsi gerçekleşir”, “hiçbiri gerçekleşmez”, “yalnız A gerçekleşir”, “tam olarak biri gerçekleşir”, “en az ikisi gerçekleşir” ve “en çok biri gerçekleşir” olaylarını A , B , C cinsinden küme işlemleriyle yazınız. “Hiçbiri gerçekleşmez” olayının “en az biri gerçekleşir” olayının tümleyeni olduğunu De Morgan kuralıyla doğrulayınız.
- g) $\Omega = \{a, b, c\}$ olsun. $\mathcal{P}(\Omega)$ 'yı yazınız. Ω üzerinde kaç sınıf vardır? Bunlardan kaç Ω 'ı içerir? Kaç hem $\emptyset$ 'i hem Ω 'ı içerir? Kaç yalnız basit olaylardan oluşur?

Çözüm.

a) Sonucu, birinci harf 1 liralık paranın, ikinci harf 50 kuruşluk paranın yüzü olacak biçimde iki harfle gösterelim:

$$\Omega = \{YY, YT, TY, TT\}.$$

Paralar birbirinden ayırt edilebildiği için YT ile TY farklı sonuçlardır. $E_1 = \{YT, TY, TT\}$, $E_2 = \{YT, TY\}$ ve $E_3 = \{YY\}$ 'dir. $E_1^c = \Omega \setminus E_1 = \{YY\} = E_3$: “en az bir tura” olayının gerçekleşmemesi, hiç tura gelmemesi demektir. $E_1 \setminus E_2 = \{TT\}$: en az bir tura gelip tam olarak bir tura gelmemesi, iki tura gelmesidir. $E_2 \subseteq E_1$ olduğundan E_2, E_1 'i gerektirir: tam olarak bir tura geldiyse en az bir tura gelmiştir. Örnek uzay 4 elemanlı olduğundan **Önerme 1.2** gereği $2^4 = 16$ olay ve $2^{16} - 1 = 65535$ sınıf vardır.

b) $B \cup D = \{1, 3, 5, 6\}$: tek sayı veya altı gelmesi olayı. $B^c = \{2, 4, 6\}$ olduğundan $B^c \cap C = \{2, 4, 6\} \cap \{1, 2, 3\} = \{2\}$: çift ve dörtten küçük bir sayı gelmesi, yani iki gelmesi olayı. De Morgan kuralıyla $(B \cup D)^c = B^c \cap D^c = \{2, 4, 6\} \cap \{1, 2, 3, 4, 5\} = \{2, 4\}$: ne tek ne altı gelmesi, yani iki ya da dört gelmesi olayı; $\Omega \setminus \{1, 3, 5, 6\} = \{2, 4\}$ hesabı da aynı sonucu verir. $C^c \setminus D = C^c \cap D^c = \{4, 5, 6\} \cap \{1, 2, 3, 4, 5\} = \{4, 5\}$: dörtten küçük olmayan ve altı olmayan bir sayı, yani dört ya da beş gelmesi olayı. $B \cap D = \{1, 3, 5\} \cap \{6\} = \emptyset$: zar aynı anda hem tek hem altı gelemez; B ile D ayrıktır. $C^c = \{4, 5, 6\}$ ve $D = \{6\} \subseteq \{4, 5, 6\}$ olduğundan D, C^c olayını gerektirir: altı geldiyse dörtten küçük bir sayı gelmemiştir.

c) Turaanın en erken üçüncü atışta gelmesi, ilk iki atışın yazı olması demektir; bu, $n \geq 3$ olan ω_n sonuçlarının kümesidir: $E = \{\omega_3, \omega_4, \omega_5, \dots\}$. $F = \{\omega_1, \omega_3, \omega_5, \dots\}$ ve $G_n = \{\omega_1, \omega_2, \dots, \omega_n\}$ 'dir.

$\{\omega_1\} \cup \{\omega_2\} = \{\omega_1, \omega_2\}$ kümesinin tümleyeni, Ω 'nın bu ikisi dışındaki bütün elemanlarıdır, yani $\{\omega_3, \omega_4, \dots\} = E$ 'dir. F^c , indisi tek olmayan, yani çift olan ω_n 'lerin kümesidir: $F^c = \{\omega_2, \omega_4, \dots\} = C$. $\bigcup_{k=1}^n \{\omega_k\}$ birleşimi, $k \leq n$ olan ω_k sonuçlarının kümesidir; bu da G_n 'dir.

$\bigcup_{n=1}^{\infty} G_n \subseteq \Omega$ açıktır. Tersine herhangi bir $\omega \in \Omega$ alalım; bir $m \in \mathbb{N}$ için $\omega = \omega_m$ 'dir ve $\omega_m \in G_m$ olduğundan ω birleşimdedir. Öyleyse $\bigcup_{n=1}^{\infty} G_n = \Omega$ 'dır. Son eşitlik için De Morgan kuralını (**Önerme 1.1**) kullanalım:

$$\bigcap_{n=1}^{\infty} G_n^c = \left(\bigcup_{n=1}^{\infty} G_n \right)^c = \Omega^c = \emptyset.$$

Sözle: “her n için tura n -inci atıştan sonra gelir” olayı imkânsız olaydır; çünkü örnek uzayda her sonuç sonlu bir atışta biter.

d) $K = [20, 35]$ 'tir. $A \cap K = [20, 25)$: sıcaklığın 20 ile 25 arasında olması (20 dâhil). $A \cup K = (15, 35)$: sıcaklığın 15 dereceden yüksek olması. $A \setminus K = (15, 20)$: sıcaklığın 15 ile 20 derece arasında olması. $K \setminus A = [25, 35)$: sıcaklığın en az 25 derece olması. $A^c \cap K^c = (A \cup K)^c = (-10, 15]$: sıcaklığın 15 dereceyi geçmemesi; De Morgan kuralı burada iki tümleyenin kesişimini hesaplamayı tek bir tümleyene indirger. $A \cap B = (15, 25) \cap \{10\} = \emptyset$: A ile B ayrıktır.

e) (i) $A \subseteq B$ olsun. $A \cap B \subseteq A$ her zaman doğrudur; tersine $\omega \in A$ ise $A \subseteq B$ gereği $\omega \in B$, dolayısıyla $\omega \in A \cap B$ olur. Öyleyse $A \cap B = A$. Şimdi $A \cap B = A$ olsun. $B \subseteq A \cup B$ her zaman doğrudur; tersine $\omega \in A \cup B$ ise ya $\omega \in B$ 'dir ya da $\omega \in A = A \cap B \subseteq B$ 'dir; her hâlde $\omega \in B$. Öyleyse $A \cup B = B$. Son olarak $A \cup B = B$ olsun; $A \subseteq A \cup B = B$ olur. Böylece üç önerme birbirini dögüsel olarak gerektirir ve denktir.

(ii) $A \subseteq B$ ve $\omega \in B^c$ olsun; $\omega \notin B$ 'dir. $\omega \in A$ olsaydı $A \subseteq B$ gereği $\omega \in B$ olurdu; çelişki. Öyleyse $\omega \notin A$, yani $\omega \in A^c$. Sözle: A gerçekleşince B de gerçekleşiyorsa, B 'nin gerçekleşmemesi A 'nın gerçekleşmemesini gerektirir.

(iii) $A \setminus B \subseteq B^c$ ve $A \cap B \subseteq B$ olduğundan $(A \setminus B) \cap (A \cap B) \subseteq B^c \cap B = \emptyset$; benzer biçimde $B \setminus A \subseteq A^c$ ve $A \cap B \subseteq A$ olduğundan $(B \setminus A) \cap (A \cap B) = \emptyset$; son olarak $A \setminus B \subseteq A$ ve $B \setminus A \subseteq A^c$ olduğundan $(A \setminus B) \cap (B \setminus A) = \emptyset$. Üç olay ikişer ikişer ayrıktır. Eşitlik için **Önerme 1.1** (f)'yi iki kez uygulayalım: $A \cup B = A \cup (B \setminus A)$ ve $A = (A \cap B) \cup (A \setminus B)$; yerine koyunca $A \cup B = (A \setminus B) \cup (A \cap B) \cup (B \setminus A)$ elde edilir. Bu ayrıştırma, $A \cup B$ olayını “yalnız A ”, “hem A hem B ” ve “yalnız B ” parçalarına böler.

f) Tabloya göre çeviri yapalım. En az biri: $A \cup B \cup C$. Hepsisi: $A \cap B \cap C$. Hiçbiri: $A^c \cap B^c \cap C^c$. Yalnız A : $A \cap B^c \cap C^c$. Tam olarak biri, “yalnız A ”, “yalnız B ”, “yalnız C ” olaylarının birleşimidir:

$$(A \cap B^c \cap C^c) \cup (A^c \cap B \cap C^c) \cup (A^c \cap B^c \cap C).$$

En az ikisi, herhangi iki tanesinin birlikte gerçekleşmesidir: $(A \cap B) \cup (A \cap C) \cup (B \cap C)$. En çok biri, “en az ikisi” olayının tümleyenidir; De Morgan kuralıyla

$$((A \cap B) \cup (A \cap C) \cup (B \cap C))^c = (A^c \cup B^c) \cap (A^c \cup C^c) \cap (B^c \cup C^c).$$

Doğrulama: De Morgan kuralını üç olaya uygularsak (kural, $A_1 = A$, $A_2 = B$ ve $n \geq 3$ için $A_n = C$ alınarak genel biçimden çıkar) $(A \cup B \cup C)^c = A^c \cap B^c \cap C^c$ olur; sol taraf “en az biri gerçekleşir” olayının tümleyeni, sağ taraf “hiçbiri gerçekleşmez” olayıdır.

g) Ω üç elemanlı olduğundan $2^3 = 8$ alt kümesi vardır:

$$\mathcal{P}(\Omega) = \{ \emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \Omega \}.$$

Sınıf sayısı $2^3 - 1 = 255$ 'tir. Ω 'yı içeren bir sınıf, Ω 'ya geri kalan yedi olayın herhangi bir alt kümesinin eklenmesiyle elde edilir; bu alt küme boş da olabilir, çünkü $\{\Omega\}$ tek başına bir sınıftır. Dolayısıyla Ω 'yı içeren $2^7 = 128$ sınıf vardır. Hem $\emptyset$ 'i hem Ω 'yı içeren sınıflar, bu ikisine geri kalan altı olayın herhangi bir alt kümesinin eklenmesiyle elde edilir: $2^6 = 64$ tane. Yalnız basit olaylardan oluşan sınıflar, $\{\{a\}, \{b\}, \{c\}\}$ ailesinin boş olmayan alt kümeleridir: $2^3 - 1 = 7$ tane.

■

Bu bölümde bir deneyin sonuçlarını, olaylarını ve olaylar üzerindeki işlemleri küme kuramının diline taşıdık; olayların toplandığı sınıf kavramını ve en büyük sınıf olan kuvvet kümesini tanıdık. Sıradaki bölümde bir olaylar sınıfının, üzerinde olasılık ölçüsü tanımlanabilmesi için hangi kapalılık koşullarını sağlaması gerektiğini belirleyeceğiz: **Cebirler**, **σ -Cebirler** ve **Borel Cebri**.

2. Cebirler, σ -Cebirler ve Borel Cebri

Rastgele Deney, Örnek Uzak ve Olaylar bölümünde bir rastgele deneyin bütün olabilir sonuçlarını örnek uzay Ω 'da topladık ve Ω 'nın alt kümelerine olay dedik (Tanım 1.3). Olayları “veya”, “ve”, “değil” bağlaçlarıyla birleştirip yeni olaylar kurduk: $A \cup B$, $A \cap B$, A^c . Örnek uzayın bazı alt kümelerinden oluşan bir topluluğa sınıf adını verdik (Tanım 1.5); bütün alt kümelerin sınıfı da kuvvet kümesi $\mathcal{P}(\Omega)$ idi (Tanım 1.6). Şimdi şu soruyu soruyoruz: **bir deneyde “olay” diye adlandıracağımız alt kümelerin sınıfı nasıl bir sınıf olmalıdır?** En azından şunu isteriz: iki olaydan “veya”, “ve”, “değil” ile türetilen kümeler de olay olsun; yani sınıf bu işlemlere göre kapalı olsun. Bir madeni paranın tura gelinceye kadar atılması gibi sonsuz sonuçlu deneylerde ise “tura çift numaralı bir atışta gelir” gibi olaylar, sayılabilir sonsuz tane olayın birleşimidir; öyleyse sınıfın sayılabilir birleşime de kapalı olması gerekir. Bu iki isteğin kesin biçimi **cebri** ve **σ -cebri** kavramlarıdır.

Bu bölümde şunları öğreneceğiz: cebir ve σ -cebri tanımları ve aralarındaki fark; σ -cebrilerin kesişim, fark gibi işlemlere de kapalı olduğu; verilen bir sınıfı kapsayan en küçük σ -cebrin, yani **üretilen σ -cebrin** varlığı ve somut örneklerde nasıl hesaplandığı; bir olaya kısıtlanmış **iz σ -cebrin**; ve örnek uzay $\mathbb{R}$ olduğunda olasılık teorisinin standart olay sınıfı olan **Borel cebri**. Borel cebrini tanıttıktan sonra, örnek uzay $\mathbb{R}$ iken neden her alt kümeye olasılık verilemeyeceğini sezgisel olarak açıklayacağız. Bu yapı, bir sonraki bölümde olasılık ölçüsünün tanım kümesi olacaktır.

2.1 Cebirler

Bir olaylar sınıfından istediğimiz en temel şey, “veya” ve “değil” işlemlerine kapalı olmasıdır. “Ve” işlemini ayrıca istemeye gerek yoktur; birazdan göreceğimiz gibi De Morgan kuralı onu ötekilerden üretir. Bir de kesin olayı, yani Ω 'nın kendisini sınıfın içinde isteriz.

Tanım 2.1 (Cebri). $\Omega \neq \emptyset$ bir küme ve $\mathcal{U}$, Ω üzerinde bir sınıf olsun. $\mathcal{U}$ sınıfı

(i) $\Omega \in \mathcal{U}$,

(ii) $A \in \mathcal{U} \Rightarrow A^c \in \mathcal{U}$,

(iii) $A, B \in \mathcal{U} \Rightarrow A \cup B \in \mathcal{U}$

koşullarını sağlıyorsa $\mathcal{U}$ sınıfına Ω üzerinde bir **cebri** denir.

Koşul (i) kesin olayın sınıfta olduğunu, (ii) sınıfın tümleyen almaya, (iii) ise iki kümenin birleşimini almaya kapalı olduğunu söyler. Bu üç koşuldan başka kapalılıklar da hemen çıkar:

- $\emptyset = \Omega^c \in \mathcal{U}$; yani imkânsız olay sınıftadır.
- $A, B \in \mathcal{U}$ ise $A^c, B^c \in \mathcal{U}$ ve dolayısıyla $A^c \cup B^c \in \mathcal{U}$; De Morgan kuralı (Önerme 1.1) gereği $A \cap B = (A^c \cup B^c)^c \in \mathcal{U}$. Yani cebri kesişime de kapalıdır.
- $A \setminus B = A \cap B^c \in \mathcal{U}$.
- Tümevarımla: $B_1, \dots, B_m \in \mathcal{U}$ ise $B_1 \cup \dots \cup B_m \in \mathcal{U}$ ve $B_1 \cap \dots \cap B_m \in \mathcal{U}$. Gerçekten m küme için doğrusa $B_1 \cup \dots \cup B_{m+1} = (B_1 \cup \dots \cup B_m) \cup B_{m+1}$ yazılır ve (iii) uygulanır; kesişim için aynı yol izlenir.

Demek ki bir cebri, $\cup, \cap, \setminus$ ve tümleyen işlemlerinin **sonlu** kez uygulanmasına kapalıdır. Sonsuz sonuçlu deneylerde bu yetmez.

2.2 σ -Cebirler

Bir madeni paranın tura gelinceye kadar atılması deneyinde (Örnek 1.2) $\Omega = \{T, YT, YYT, YYYT, \dots\} = \{\omega_1, \omega_2, \omega_3, \dots\}$ idi; ω_n , $n - 1$ tane Y 'nin ardından T gelen sonuçtur. $A_n = \{\omega_n\}$, “tura n -inci atışta gelir” olayı olsun: $A_1 = \{T\}$, $A_2 = \{YT\}$, $A_3 = \{YYT\}$ ve böyle sürer. “Tura çift numaralı bir atışta gelir” olayı

$$A_2 \cup A_4 \cup A_6 \cup \dots = \bigcup_{k=1}^{\infty} A_{2k}$$

biçimindedir; yani sayılabilir sonsuz tane olayın birleşimidir. Böyle olayları da olay saymak istiyorsak, sınıfın sayılabilir birleşime kapalı olmasını istemeliyiz. Tanımdaki (iii) koşulunu bu yönde güçlendiriyoruz.

Tanım 2.2 (σ -Cebir). $\Omega \neq \emptyset$ bir küme ve $\mathcal{U}$, Ω üzerinde bir sınıf olsun. $\mathcal{U}$ sınıfı

(i) $\Omega \in \mathcal{U}$,

(ii) $A \in \mathcal{U} \Rightarrow A^c \in \mathcal{U}$,

(iii*) $A_n \in \mathcal{U}$ ($n = 1, 2, \dots$) bir küme dizisi ise $\bigcup_{n=1}^{\infty} A_n \in \mathcal{U}$

koşullarını sağlıyorsa $\mathcal{U}$ sınıfına Ω üzerinde bir **σ -cebri** (sigma cebri) denir. Bir σ -cebirin elemanı olan kümeler **olay** denir; $(\Omega, \mathcal{U})$ ikilisine de **ölçülebilir uzay** denir.

Bu tanımdan sonra “olay” sözcüğü daha dar bir anlam kazanır: Ω ’nın her alt kümesi değil, seçilen σ -cebir $\mathcal{U}$ ’nun elemanı olan alt kümeler olaydır. Sonlu ya da sayılabilir örnek uzaylarda $\mathcal{U} = \mathcal{P}(\Omega)$ alınır ve iki anlam çıkarır; farkın gerçekten önemli olduğu durum $\Omega = \mathbb{R}$ gibi sayılamayan örnek uzaylardır. Adındaki σ harfi sayılabilir birleşimi anımsatır. İlk soru, iki kavram arasındaki ilişkidir.

Önerme 2.1 (Her σ -Cebir Bir Cebirdir). Ω üzerindeki her σ -cebir bir cebirdir. Bunun tersi genel olarak doğru değildir: $\Omega = \mathbb{N}$ üzerinde, cebir olduğu hâlde σ -cebir olmayan bir sınıf vardır.

İspat.

Birinci kısım. $\mathcal{U}$ bir σ -cebir olsun. Cebir tanımındaki (i) ve (ii) koşulları σ -cebir tanımındakilerle aynıdır; yalnızca (iii) koşulunu göstermek gerekir. $A, B \in \mathcal{U}$ olsun. $A_1 = A$ ve $n \geq 2$ için $A_n = B$ alalım. (A_n) , $\mathcal{U}$ ’da bir küme dizisidir ve

$$\bigcup_{n=1}^{\infty} A_n = A \cup B \cup B \cup B \cup \dots = A \cup B$$

olur. (iii*) gereği $A \cup B \in \mathcal{U}$ ’dur. Demek ki $\mathcal{U}$ bir cebirdir.

İkinci kısım: bir karşı örnek. $\Omega = \mathbb{N} = \{1, 2, 3, \dots\}$ ve

$$\mathcal{U} = \{A \subseteq \mathbb{N} : A \text{ sonlu ya da } A^c \text{ sonlu}\}$$

olsun. Önce $\mathcal{U}$ ’nun bir cebir olduğunu gösterelim.

(i) $\mathbb{N}^c = \emptyset$ sonludur; öyleyse $\mathbb{N} \in \mathcal{U}$.

(ii) $A \in \mathcal{U}$ olsun. A sonluysa A^c ’nin tümleyeni $(A^c)^c = A$ sonludur, dolayısıyla $A^c \in \mathcal{U}$. A^c sonluysa A^c doğrudan tanım gereği $\mathcal{U}$ ’dadır.

(iii) $A, B \in \mathcal{U}$ olsun. İki durum vardır. Her ikisi de sonluysa $A \cup B$ sonludur ve $\mathcal{U}$ ’dadır. Değilse en az biri, diyelim A , sonsuzdur; $A \in \mathcal{U}$ olduğundan A^c sonludur ve

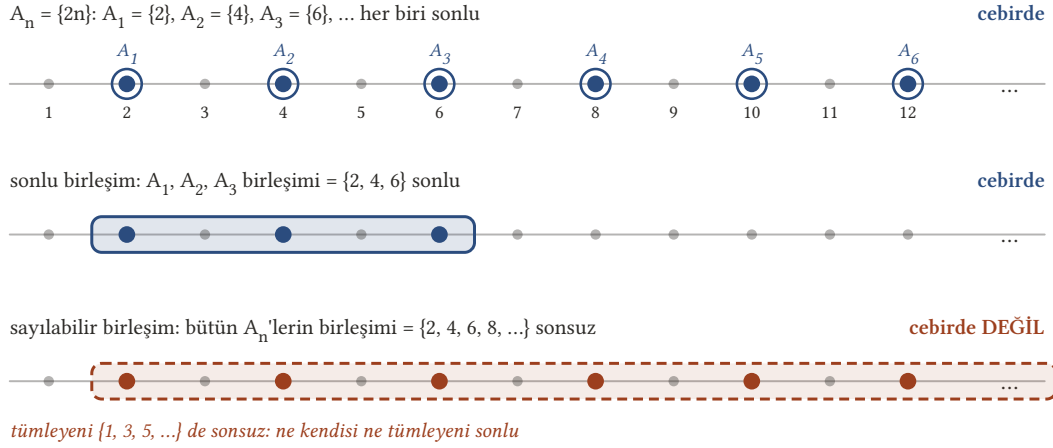
$$(A \cup B)^c = A^c \cap B^c \subseteq A^c$$

olduğundan $(A \cup B)^c$ sonlu bir kümenin alt kümesidir, dolayısıyla sonludur; yine $A \cup B \in \mathcal{U}$.

Şimdi $\mathcal{U}$ ’nun σ -cebir olmadığını gösterelim. $n = 1, 2, \dots$ için $A_n = \{2n\}$ olsun. Her A_n tek elemanlı, dolayısıyla sonludur; öyleyse $A_n \in \mathcal{U}$. Ama

$$\bigcup_{n=1}^{\infty} A_n = \{2, 4, 6, \dots\}$$

çift sayılar kümesidir. Bu küme sonsuzdur; tümleyeni olan tek sayılar kümesi $\{1, 3, 5, \dots\}$ de sonsuzdur. Öyleyse $\bigcup_{n=1}^{\infty} A_n \notin \mathcal{U}$ ve (iii*) sağlanmaz. $\mathcal{U}$ bir cebirdir ama σ -cebir değildir. Aynı kurgu, sayılabilir sonsuz bir alt küme içeren her sonsuz Ω üzerinde de çalışır.



Şekil 2.1: $\Omega = \mathbb{N}$ üzerinde "sonlu ya da tümleyeni sonlu" kümelerin sınıfı bir cebirdir ama σ -cebir değildir. $A_n = \{2n\}$ kümelerinin her biri sonludur ve sınıftadır; sonlu tanesinin birleşimi yine sonludur ve sınıfta kalır. Bütün A_n 'lerin sayılabilir birleşimi ise çift sayılar kümesidir: hem kendisi hem tümleyeni sonsuz olduğundan sınıfın dışına çıkar. Cebir ile σ -cebir arasındaki fark tam olarak bu sayılabilir birleşim adımıdır.

Karşı örnekteki sorun, dizideki sonsuz tane farklı kümenin birleşiminin sınıftan çıkmasıdır. Örnek uzay sonluysa bir dizide ancak sonlu tane farklı küme bulunabilir; bu yüzden fark ortadan kalkar.

Önerme 2.2 (Sonlu Örnek Uzayda Her Cebir Bir σ -Cebirdir). Ω sonlu bir küme ise Ω üzerindeki her cebir bir σ -cebirdir.

İspat.

$\mathcal{U}, \Omega$ üzerinde bir cebir olsun. (i) ve (ii) hazır; (iii*) koşulunu göstermeliyiz. $|\Omega| = k$ olsun. Ω 'nın tam 2^k tane alt kümesi vardır (Önerme 1.2); dolayısıyla $\mathcal{U}$ 'nun en çok 2^k elemanı vardır.

(A_n) , $\mathcal{U}$ 'da herhangi bir küme dizisi olsun. Dizinin terimleri $\mathcal{U}$ 'nun elemanları olduğundan, dizide **en çok 2^k tane farklı küme** geçer. Bu farklı kümeler $B_1, \dots, B_m$ olsun ($m \leq 2^k$). Her A_n bu kümelerden biridir ve her B_j dizide en az bir kez geçer. Öyleyse

$$\bigcup_{n=1}^{\infty} A_n = B_1 \cup B_2 \cup \dots \cup B_m$$

olur: soldaki birleşimin her elemanı bir A_n 'de, dolayısıyla bir B_j 'dedir; sağdaki birleşimin her elemanı bir B_j 'de, dolayısıyla bu B_j 'ye eşit olan bir A_n 'dedir. Sağ taraf $\mathcal{U}$ 'nun sonlu tane elemanının birleşimidir ve bir önceki alt bölümde gördüğümüz gibi cebirler sonlu birleşime kapalıdır. Demek ki $\bigcup_n A_n \in \mathcal{U}$; yani $\mathcal{U}$ bir σ -cebirdir.

Demek ki sonlu örnek uzaylarda "cebir" ile " σ -cebir" arasında fark yoktur; fark ancak sonsuz Ω 'da ortaya çıkar. Şimdi küçük bir örnek uzayda çeşitli sınıfları inceleyelim; amaç, tanımdaki koşulların somut bir sınıfta nasıl denetlendiğini ve hangilerinin bozulabildiğini görmek.

Örnek 2.1 (Dört Elemanlı Bir Örnek Uzayda Sınıflar). $\Omega = \{a, b, c, d\}$ olsun. Aşağıdaki sınıflardan hangileri Ω üzerinde bir σ -cebirdir?

$$\begin{aligned} \mathcal{U}_1 &= \{ \emptyset, \{a\}, \{b, c, d\}, \Omega \}, \mathcal{U}_2 = \{ \{a\}, \{b\}, \{c\}, \{d\} \}, \\ \mathcal{U}_3 &= \{ \emptyset, \Omega \}, \mathcal{U}_4 = \{ \emptyset, \{a\}, \{b\}, \{a, c, d\}, \{b, c, d\}, \Omega \}. \end{aligned}$$

Çözüm.

Ω sonlu olduğundan, **Önerme 2.2** gereği bir sınıfın σ -cebir olması için cebir olması yeterlidir; yani yalnızca (i), (ii) ve iki kümenin birleşimi için (iii) koşullarını denetlemek yeter.

$\mathcal{U}_1$ bir σ -cebirdir. (i) $\Omega \in \mathcal{U}_1$. (ii) Tümlenler: $\emptyset^c = \Omega$, $\{a\}^c = \{b, c, d\}$, $\{b, c, d\}^c = \{a\}$, $\Omega^c = \emptyset$; hepsi $\mathcal{U}_1$ 'dedir. (iii) $\emptyset \cup X = X$ ve $\Omega \cup X = \Omega$ olduğundan yalnızca $\{a\} \cup \{b, c, d\} = \Omega$ birleşimine bakmak yeter; o da $\mathcal{U}_1$ 'dedir. Bu σ -cebir, deneyde yalnızca “a geldi mi, gelmedi mi?” sorusunun yanıtlanabildiği duruma karşılık gelir.

$\mathcal{U}_2$ bir σ -cebir değildir. $\Omega \notin \mathcal{U}_2$ olduğundan (i) sağlanmaz. Ayrıca $\{a\}^c = \{b, c, d\} \notin \mathcal{U}_2$ ve $\{a\} \cup \{b\} = \{a, b\} \notin \mathcal{U}_2$; yani (ii) ve (iii) de bozulur. $\mathcal{U}_2$ bir cebir bile değildir.

$\mathcal{U}_3$ bir σ -cebirdir. (i) $\Omega \in \mathcal{U}_3$. (ii) $\emptyset^c = \Omega$ ve $\Omega^c = \emptyset$. (iii) $\emptyset \cup \emptyset = \emptyset$, $\emptyset \cup \Omega = \Omega$, $\Omega \cup \Omega = \Omega$. Bu, her örnek uzay üzerindeki en küçük σ -cebirdir ve **aşıkâr σ -cebir** diye anılır; deney hakkında hiçbir şey gözlenemediği duruma karşılık gelir.

$\mathcal{U}_4$ bir σ -cebir değildir. (i) sağlanır. (ii) de sağlanır: $\{a\}^c = \{b, c, d\}$, $\{b\}^c = \{a, c, d\}$, $\{a, c, d\}^c = \{b\}$, $\{b, c, d\}^c = \{a\}$ ve $\emptyset^c = \Omega$, $\Omega^c = \emptyset$. Ama $\{a\}, \{b\} \in \mathcal{U}_4$ olduğu hâlde $\{a\} \cup \{b\} = \{a, b\} \notin \mathcal{U}_4$; (iii) bozulur.

Özetle $\mathcal{U}_2$ üç koşulu birden, $\mathcal{U}_4$ ise yalnızca üçüncüsünü bozar. Yalnızca ikinci koşulu bozan bir sınıf da vardır: $\mathcal{U}_5 = \{\emptyset, \{a\}, \Omega\}$ sınıfında Ω vardır ve herhangi iki elemanın birleşimi yine sınıftadır, ama $\{a\}^c = \{b, c, d\} \notin \mathcal{U}_5$. Demek ki (ii) ile (iii) birbirinden bağımsızdır. Buna karşılık (i), boş olmayan bir sınıf için (ii) ve (iii)'ten çıkar: herhangi bir $A \in \mathcal{U}$ alınırsa $A^c \in \mathcal{U}$ ve $\Omega = A \cup A^c \in \mathcal{U}$ olur. Tanımda (i) koşulunu yazmak, boş sınıfı dışarıda bırakmaya ve Ω 'yı en baştan görünür kılmaya yarar.

■

Bir örnek uzay üzerindeki en büyük sınıf kuvvet kümesidir. Onun da bir σ -cebir olduğunu göstermek kolaydır; sonlu ve sayılabilir örnek uzaylarda hep bu σ -cebir kullanacağız.

Önerme 2.3 (Kuvvet Kümesi Bir σ -Cebirdir). $\Omega \neq \emptyset$ bir küme olsun. Kuvvet kümesi $\mathcal{P}(\Omega) = \{A : A \subseteq \Omega\}$, Ω üzerinde bir σ -cebirdir. Ayrıca Ω üzerindeki her σ -cebir $\mathcal{P}(\Omega)$ 'nin bir alt sınıfıdır; yani $\mathcal{P}(\Omega)$, Ω üzerindeki en büyük σ -cebirdir.

İspat.

(i) $\Omega \subseteq \Omega$ olduğundan $\Omega \in \mathcal{P}(\Omega)$.

(ii) $A \in \mathcal{P}(\Omega)$ olsun; yani $A \subseteq \Omega$. O zaman $A^c = \Omega \setminus A \subseteq \Omega$, dolayısıyla $A^c \in \mathcal{P}(\Omega)$.

(iii*) (A_n) , $\mathcal{P}(\Omega)$ 'da bir dizi olsun; yani her n için $A_n \subseteq \Omega$. $x \in \bigcup_{n=1}^{\infty} A_n$ ise en az bir n için $x \in A_n \subseteq \Omega$, dolayısıyla $x \in \Omega$. Demek ki $\bigcup_n A_n \subseteq \Omega$, yani $\bigcup_n A_n \in \mathcal{P}(\Omega)$.

Son iddia tanımdan çıkar: bir σ -cebir, tanım gereği Ω 'nın alt kümelerinden oluşan bir sınıftır; Ω 'nın her alt kümesi $\mathcal{P}(\Omega)$ 'da olduğundan $\mathcal{U} \subseteq \mathcal{P}(\Omega)$ 'dir.

■

Örneğin bir zarın atılması deneyinde $\Omega = \{1, 2, 3, 4, 5, 6\}$ ve $\mathcal{U} = \mathcal{P}(\Omega)$ alınır; bu σ -cebirde $2^6 = 64$ olay vardır.

2.3 σ -Cebirlerin Temel Özellikleri

σ -cebir tanımı yalnızca Ω 'yı, tümleniyi ve sayılabilir birleşimi anar. Oysa olaylarla çalışırken “ve” (kesişim), “ama değil” (fark) ve sonlu işlemleri de sürekli kullanırız. Aşağıdaki teorem, bütün bu işlemlerin bizi σ -cebirin dışına çıkarmadığını söyler; ispatın anahtarı, cebirlerde olduğu gibi, De Morgan kurallarıdır.

Teorem 2.1 (σ -Cebirlerin Temel Özellikleri). $\mathcal{U}, \Omega$ üzerinde bir σ -cebir olsun.

(a) $\emptyset \in \mathcal{U}$.

(b) (A_n) , $\mathcal{U}$ 'da bir küme dizisi ise $\bigcap_{n=1}^{\infty} A_n \in \mathcal{U}$.

(c) $A_1, \dots, A_n \in \mathcal{U}$ ise $\bigcup_{i=1}^n A_i \in \mathcal{U}$ ve $\bigcap_{i=1}^n A_i \in \mathcal{U}$.

(d) $A_1, A_2 \in \mathcal{U}$ ise $A_1 \setminus A_2 \in \mathcal{U}$.

İspat.

(a) (i) gereği $\Omega \in \mathcal{U}$; (ii) gereği $\Omega^c = \emptyset \in \mathcal{U}$.

(b) Her n için $A_n \in \mathcal{U}$ olsun. (ii) gereği her n için $A_n^c \in \mathcal{U}$; (iii*) gereği $\bigcup_{n=1}^{\infty} A_n^c \in \mathcal{U}$; yeniden (ii) gereği

$$\left(\bigcup_{n=1}^{\infty} A_n^c\right)^c \in \mathcal{U}$$

olur. Bu kümenin $\bigcap_n A_n$ 'ye eşit olduğunu görelim. Herhangi bir $x \in \Omega$ için

$$x \in \left(\bigcup_{n=1}^{\infty} A_n^c\right)^c \Leftrightarrow x \notin \bigcup_{n=1}^{\infty} A_n^c \Leftrightarrow \text{her } n \text{ için } x \notin A_n^c \Leftrightarrow \text{her } n \text{ için } x \in A_n \Leftrightarrow x \in \bigcap_{n=1}^{\infty} A_n.$$

Bu, **Önerme 1.1** (e)'deki De Morgan kuralının sayılabilir biçimidir. Demek ki $\bigcap_n A_n \in \mathcal{U}$.

(c) $A_1, \dots, A_n \in \mathcal{U}$ olsun. Sonlu birleşim için diziyi boş kümeyle uzatalım: $i > n$ için $A_i = \emptyset$ olsun. (a) gereği $\emptyset \in \mathcal{U}$, dolayısıyla $(A_i)_{i \geq 1}$, $\mathcal{U}$ 'da bir dizidir. Boş küme birleşime hiçbir eleman katmadığından

$$\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^n A_i$$

olur ve (iii*) gereği bu küme $\mathcal{U}$ 'dadır.

Sonlu kesişim için diziyi Ω ile uzatalım: $i > n$ için $A_i = \Omega$ olsun; (i) gereği $\Omega \in \mathcal{U}$. $\Omega \cap X = X$ olduğundan

$$\bigcap_{i=1}^{\infty} A_i = \bigcap_{i=1}^n A_i$$

olur ve (b) gereği bu küme $\mathcal{U}$ 'dadır.

(d) $A_1 \setminus A_2 = A_1 \cap A_2^c$ 'dir: $x \in A_1 \setminus A_2$ olması, $x \in A_1$ ve $x \notin A_2$, yani $x \in A_1$ ve $x \in A_2^c$ olması demektir. (ii) gereği $A_2^c \in \mathcal{U}$; (c) gereği $A_1 \cap A_2^c \in \mathcal{U}$.

■

Kısaca: σ -cebirlere $\cup, \cap, \setminus$ ve tümleyen işlemlerinin sonlu ya da sayılabilir sonsuz kez uygulanmasına göre kapalı sınıflardır. Örneğin simetrik fark $A \triangle B = (A \setminus B) \cup (B \setminus A)$ da σ -cebirdir.

💡 Tanımı kesişimle de verebilirdik

(i), (ii) ve "sayılabilir kesişime kapalılık" koşullarını sağlayan bir sınıf, aynı De Morgan hesabıyla sayılabilir birleşime de kapalıdır, yani bir σ -cebirdir. Tanımda birleşimi seçmek yalnızca bir gelenektir.

2.4 σ -Cebirlerin Arakesiti ve Üretilen σ -Cebir

Uygulamada işe genellikle bir σ -cebirle değil, olay olmasını istediğimiz kümelerin bir sınıfıyla başlarız: $\Omega = \mathbb{R}$ için aralıklar, $\Omega = \{a, b, c, d\}$ için "gözleyebildiğimiz" birkaç alt küme gibi. Bu sınıf çoğu zaman σ -cebir değildir. Doğal soru şudur: **verilen bir sınıfı kapsayan en küçük σ -cebir var mıdır?** Yanıt evettir ve yapı çok basittir: sınıfı kapsayan bütün σ -cebirlere arakesitini (kesişimini) alırız. Bunun işe yaraması için önce arakesitin σ -cebir olduğunu bilmeliyiz.

Teorem 2.2 (σ -Cebirlerin Arakesiti Bir σ -Cebirdir). $I \neq \emptyset$ bir indis kümesi ve her $i \in I$ için $\mathcal{U}_i$, Ω üzerinde bir σ -cebir olsun. O zaman

$$\mathcal{U} = \bigcap_{i \in I} \mathcal{U}_i = \{A \subseteq \Omega : \text{her } i \in I \text{ için } A \in \mathcal{U}_i\}$$

sınıfı da Ω üzerinde bir σ -cebirdir. (I sonlu, sayılabilir ya da sayılamaz olabilir.)

İspat.

Üç koşulu sırayla denetleyelim.

(i) Her $i \in I$ için $\mathcal{U}_i$ bir σ -cebir olduğundan $\Omega \in \mathcal{U}_i$ 'dir. Öyleyse Ω bütün $\mathcal{U}_i$ 'lerde, dolayısıyla arakesitte bulunur: $\Omega \in \mathcal{U}$.

(ii) $A \in \mathcal{U}$ olsun. O zaman her $i \in I$ için $A \in \mathcal{U}_i$ 'dir. $\mathcal{U}_i$ tümleyene kapalı olduğundan her i için $A^c \in \mathcal{U}_i$; yani $A^c \in \mathcal{U}$.

(iii*) (A_n) , $\mathcal{U}$ 'da bir dizi olsun. Her n ve her $i \in I$ için $A_n \in \mathcal{U}_i$ 'dir. i 'yi sabitleyelim: (A_n) , σ -cebir olan $\mathcal{U}_i$ 'de bir dizidir, dolayısıyla $\bigcup_n A_n \in \mathcal{U}_i$. Bu her $i \in I$ için doğru olduğundan $\bigcup_n A_n \in \mathcal{U}$.

■

Birleşim için benzer bir sonuç **yoktur**: iki σ -cebirin birleşimi genel olarak bir cebir bile değildir. Somut bir örneği bölüm sonundaki alıştırmalarda göreceğiz. Sebep şudur: birinden alınan bir küme ile ötekenden alınan bir kümenin birleşimi, iki sınıfın hiçbirinde bulunmak zorunda değildir.

Tanım 2.3 (Küçük ve Büyük σ -Cebir). $\mathcal{U}_1$ ve $\mathcal{U}_2$, Ω üzerinde iki σ -cebir (ya da iki cebir) olsun. $\mathcal{U}_1 \subseteq \mathcal{U}_2$ ise $\mathcal{U}_1$ 'e $\mathcal{U}_2$ 'den **küçük**, $\mathcal{U}_2$ 'ye $\mathcal{U}_1$ 'den **büyük** σ -cebir denir.

Bu bir kısmi sıralamadır; tam sıralama değildir: iki σ -cebirden biri ötekenden küçük olmak zorunda değildir. Bununla birlikte iki uç nokta her zaman vardır: aşıkâr σ -cebir $\{\emptyset, \Omega\}$ her σ -cebirden küçüktür (çünkü her σ -cebir $\emptyset$ ile Ω 'yı içerir) ve $\mathcal{P}(\Omega)$ her σ -cebirden büyüktür (**Önerme 2.3**).

Şimdi bölümün ana yapı taşına geliyoruz.

Teorem 2.3 (Üretilen σ -Cebir). $\mathcal{L}, \Omega$ üzerinde herhangi bir sınıf olsun. $\mathcal{L}$ 'yi kapsayan bütün σ -cebirlerin arakesiti $\sigma(\mathcal{L})$ ile gösterilsin. O zaman:

- (a) $\sigma(\mathcal{L})$, Ω üzerinde bir σ -cebirdir ve $\mathcal{L} \subseteq \sigma(\mathcal{L})$ 'dir.
- (b) $\mathcal{U}$, $\mathcal{L}$ 'yi kapsayan herhangi bir σ -cebir ise $\sigma(\mathcal{L}) \subseteq \mathcal{U}$ 'dur.
- (c) (a) ve (b) özelliklerini taşıyan tek σ -cebir $\sigma(\mathcal{L})$ 'dir.

$\sigma(\mathcal{L})$ 'ye $\mathcal{L}$ **sınıfının ürettiği σ -cebir** ya da $\mathcal{L}$ 'yi kapsayan **en küçük σ -cebir** denir; $\mathcal{L}$ 'ye de $\sigma(\mathcal{L})$ 'nin bir **üretici** denir.

İspat.

$\mathfrak{S}, \mathcal{L}$ 'yi kapsayan bütün σ -cebirlerin ailesi olsun:

$$\mathfrak{S} = \{\mathcal{U} : \mathcal{U}, \Omega \text{ üzerinde bir } \sigma\text{-cebir ve } \mathcal{L} \subseteq \mathcal{U}\}.$$

$\mathcal{L}, \Omega$ 'nın alt kümelerinden oluşan bir sınıf olduğundan $\mathcal{L} \subseteq \mathcal{P}(\Omega)$ 'dir; **Önerme 2.3** gereği $\mathcal{P}(\Omega)$ bir σ -cebirdir. Öyleyse $\mathcal{P}(\Omega) \in \mathfrak{S}$ ve $\mathfrak{S} \neq \emptyset$. Tanım gereği

$$\sigma(\mathcal{L}) = \bigcap_{\mathcal{U} \in \mathfrak{S}} \mathcal{U}.$$

(a) $\mathfrak{S}$ boş olmayan bir indis kümesi ve her elemanı bir σ -cebir olduğundan, **Teorem 2.2** gereği $\sigma(\mathcal{L})$ bir σ -cebirdir. Her $\mathcal{U} \in \mathfrak{S}$ için $\mathcal{L} \subseteq \mathcal{U}$ olduğundan $\mathcal{L}$, arakesitin de içindedir: $\mathcal{L} \subseteq \sigma(\mathcal{L})$.

(b) $\mathcal{U}$, $\mathcal{L}$ 'yi kapsayan bir σ -cebir olsun; yani $\mathcal{U} \in \mathfrak{S}$. Bir ailenin arakesiti ailenin her üyesinin içinde olduğundan $\sigma(\mathcal{L}) = \bigcap_{\mathcal{V} \in \mathfrak{S}} \mathcal{V} \subseteq \mathcal{U}$.

(c) $\mathcal{V}$ de (a) ve (b) özelliklerini taşıyan bir σ -cebir olsun: $\mathcal{L} \subseteq \mathcal{V}$ ve $\mathcal{L}$ 'yi kapsayan her σ -cebir $\mathcal{U}$ için $\mathcal{V} \subseteq \mathcal{U}$. $\mathcal{V}$ 'nin (b) özelliğini $\mathcal{U} = \sigma(\mathcal{L})$ için uygularsak $\mathcal{V} \subseteq \sigma(\mathcal{L})$ çıkar. $\sigma(\mathcal{L})$ 'nin (b) özelliğini $\mathcal{U} = \mathcal{V}$ için uygularsak $\sigma(\mathcal{L}) \subseteq \mathcal{V}$ çıkar. Öyleyse $\mathcal{V} = \sigma(\mathcal{L})$.

■

Teoremin verdiği tarif, “bütün σ -cebirlerin arakesiti”, pratikte hesap yapmaya elverişli değildir; ama (a) ve (b) birlikte çok kullanışlı bir **hesap yöntemi** sunar: $\sigma(\mathcal{L})$ için bir aday $\mathcal{V}$ bulunur, sonra iki şey gösterilir.

- $\mathcal{V}$ 'nin her elemanı $\mathcal{L}$ 'nin elemanlarından σ -cebir işlemleriyle elde edilir; böylece $\mathcal{V} \subseteq \sigma(\mathcal{L})$.
- $\mathcal{V}$ bir σ -cebirdir ve $\mathcal{L} \subseteq \mathcal{V}$ 'dir; böylece (b) gereği $\sigma(\mathcal{L}) \subseteq \mathcal{V}$.

İki kapsama birlikte $\sigma(\mathcal{L}) = \mathcal{V}$ verir. Ayrıca $\mathcal{L}$ kendisi bir σ -cebirse $\sigma(\mathcal{L}) = \mathcal{L}$ olur ve $\mathcal{K} \subseteq \mathcal{M}$ ise $\sigma(\mathcal{K}) \subseteq \sigma(\mathcal{M})$ 'dir; bunlar alıştırmalarda ispatlanmaktadır.

Örnek 2.2 (Üretilen σ -Cebirin Hesabı). $\Omega = \{a, b, c, d\}$ olsun. Aşağıdaki sınıfların ürettiği σ -cebirleri bulunuz.

- (a) $\mathcal{L} = \{\{a\}\}$.
- (b) $\mathcal{L} = \{\{a\}, \{b\}\}$.
- (c) $\mathcal{L} = \{\{a\}, \{b\}, \{c\}, \{d\}\}$.

Çözüm.

(a) Aday: $\mathcal{V} = \{ \emptyset, \{a\}, \{b, c, d\}, \Omega \}$; bu, [Örnek 2.1](#)'daki $\mathcal{U}_1$ sınıfıdır.

$\mathcal{V} \subseteq \sigma(\mathcal{L})$. $\sigma(\mathcal{L})$ bir σ -cebirdir ve $\{a\} \in \sigma(\mathcal{L})$ 'dir. (i) gereği $\Omega \in \sigma(\mathcal{L})$, [Teorem 2.1](#) (a) gereği $\emptyset \in \sigma(\mathcal{L})$, (ii) gereği $\{a\}^c = \{b, c, d\} \in \sigma(\mathcal{L})$. Böylece $\mathcal{V}$ 'nin dört elemanı da $\sigma(\mathcal{L})$ 'dedir.

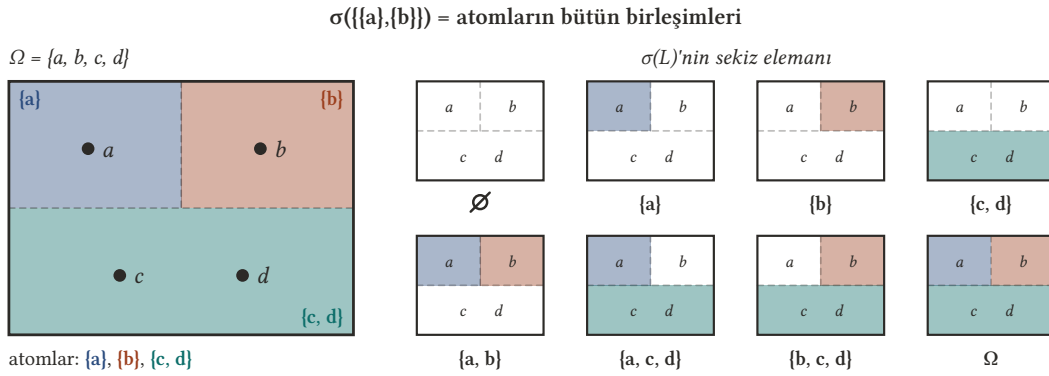
$\sigma(\mathcal{L}) \subseteq \mathcal{V}$. $\mathcal{V}$ 'nin bir σ -cebir olduğunu [Örnek 2.1](#)'da gösterdik ve $\mathcal{L} \subseteq \mathcal{V}$ 'dir. [Teorem 2.3](#) (b) gereği $\sigma(\mathcal{L}) \subseteq \mathcal{V}$.

Demek ki $\sigma(\{\{a\}\}) = \{ \emptyset, \{a\}, \{b, c, d\}, \Omega \}$.

(b) Şimdi $\{a\}$ ve $\{b\}$ birlikte veriliyor. $\{a\} \cup \{b\} = \{a, b\}$ ve bunun tümleyeni $\{c, d\}$ de $\sigma(\mathcal{L})$ 'de olmak zorundadır. $\{a\}$, $\{b\}$, $\{c, d\}$ kümeleri Ω 'yı parçalar; bunlara **atomlar** diyelim. Aday, atomların bütün birleşimlerinden oluşan sınıftır:

$$\mathcal{V} = \{ \emptyset, \{a\}, \{b\}, \{c, d\}, \{a, b\}, \{a, c, d\}, \{b, c, d\}, \Omega \}.$$

Üç atomun her alt kümesine bir birleşim karşılık geldiğinden $\mathcal{V}$ 'nin $2^3 = 8$ elemanı vardır.



Şekil 2.2: $\Omega = \{a, b, c, d\}$ üzerinde $L = \{\{a\}, \{b\}\}$ sınıfının ürettiği σ -cebir. $\{a\}$ ve $\{b\}$ verilince tümleyen ve birleşim işlemleri $\{c, d\}$ kümesini de zorunlu kılar; bu üç küme Ω 'yı parçalayan atomlardır. $\sigma(L)$ 'nin sekiz elemanı, atomların altkümelerinin birleşimleridir: hiçbir atomu almayan $\emptyset$ 'den üçünü de alan Ω 'ya kadar $2^3 = 8$ küme. c ile d hiçbir olayda birbirinden ayrılmaz.

$\mathcal{V} \subseteq \sigma(\mathcal{L})$. Her elemanı $\{a\}$ ve $\{b\}$ 'den σ -cebir işlemleriyle üretelim: $\{a, b\} = \{a\} \cup \{b\}$; $\{c, d\} = \{a, b\}^c$; $\{a, c, d\} = \{b\}^c$; $\{b, c, d\} = \{a\}^c$; $\emptyset$ ve Ω her σ -cebirdedir. Öyleyse sekiz küme de $\sigma(\mathcal{L})$ 'dedir. $\sigma(\mathcal{L}) \subseteq \mathcal{V}$. $\mathcal{V}$ 'nin bir σ -cebir olduğunu gösterelim. (i) $\Omega \in \mathcal{V}$. (ii) Tümleyen çiftleri $(\emptyset, \Omega)$, $(\{a\}, \{b, c, d\})$, $(\{b\}, \{a, c, d\})$, $(\{c, d\}, \{a, b\})$ biçimindedir ve her çiftin iki üyesi de $\mathcal{V}$ 'dedir. (iii) Atomların birleşimi olan iki kümenin birleşimi yine atomların bir birleşimidir; dolayısıyla $\mathcal{V}$ 'dedir. Ω sonlu olduğundan [Önerme 2.2](#) gereği $\mathcal{V}$ bir σ -cebirdir. $\mathcal{L} \subseteq \mathcal{V}$ olduğundan [Teorem 2.3](#) (b) gereği $\sigma(\mathcal{L}) \subseteq \mathcal{V}$.

Demek ki $\sigma(\{\{a\}, \{b\}\}) = \mathcal{V}$; sekiz olaylık bu σ -cebir, “ a geldi mi?” ve “ b geldi mi?” sorularının yanıtlarından karar verilebilen bütün olayları içerir; c ile d birbirinden ayırt edilemez.

(c) Ω 'nın her alt kümesi, elemanlarının tek nokta kümelerinin birleşimidir: $A \neq \emptyset$ için $A = \bigcup_{x \in A} \{x\}$, ve bu en çok dört kümenin birleşimidir. Tek nokta kümeleri $\mathcal{L} \subseteq \sigma(\mathcal{L})$ 'de olduğundan, [Teorem 2.1](#) (c) gereği boş olmayan her $A \subseteq \Omega$ için $A \in \sigma(\mathcal{L})$; $\emptyset$ da [Teorem 2.1](#) (a) gereği oradadır. Öyleyse $\mathcal{P}(\Omega) \subseteq \sigma(\mathcal{L})$. Öte yandan $\sigma(\mathcal{L})$, Ω 'nın alt kümelerinden oluştuğundan $\sigma(\mathcal{L}) \subseteq \mathcal{P}(\Omega)$. Demek ki

$$\sigma(\{\{a\}, \{b\}, \{c\}, \{d\}\}) = \mathcal{P}(\Omega),$$

on altı elemanlı kuvvet kümesi. Dikkat: $\mathcal{L}$, [Örnek 2.1](#)'daki $\mathcal{U}_2$ sınıfıdır; kendisi σ -cebir olmadığı hâlde en büyük σ -cebiri üretir.

■

Bir σ -cebiri bir olaya kısıtlamak da sık gereken bir işlemdir; örneğin “ B olayının gerçekleştiği biliniyor” durumunda deneyin örnek uzayı fiilen B 'ye küçülür (koşullu olasılık, [Tanım 6.1](#)). O zaman B üzerinde hangi σ -cebirle çalışılacağı sorusu doğar.

Önerme 2.4 (İz σ -Cebiri). $\mathcal{U}, \Omega$ üzerinde bir σ -cebir ve $\emptyset \neq B \subseteq \Omega$ olsun.

$$\mathcal{U}_B = \{A : A = B \cap C, C \in \mathcal{U}\} = \{B \cap C : C \in \mathcal{U}\}$$

sınıfı, B üzerinde bir σ -cebirdir. $\mathcal{U}_B$ 'ye $\mathcal{U}$ 'nin B üzerindeki **izi** ya da **iz σ -ceberi** denir. Ayrıca $B \in \mathcal{U}$ ise

$$\mathcal{U}_B = \{A \in \mathcal{U} : A \subseteq B\}$$

olur; yani iz, $\mathcal{U}$ 'nin B 'nin içinde kalan elemanlarından oluşur.

İspat.

Burada örnek uzay rolünü B oynar; dolayısıyla bir $A \subseteq B$ kümesinin tümleyeni $B \setminus A$ 'dır ve σ -cebir koşulları B 'ye göre denetlenir. Önce $\mathcal{U}_B$ 'nin gerçekten B 'nin alt kümelerinden oluştuğunu görelim: $B \cap C \subseteq B$ her zaman doğrudur.

(i) $\Omega \in \mathcal{U}$ olduğundan $B = B \cap \Omega \in \mathcal{U}_B$.

(ii) $A \in \mathcal{U}_B$ olsun; $A = B \cap C$ olacak biçimde bir $C \in \mathcal{U}$ vardır. A 'nın B 'ye göre tümleyeni $B \setminus A$ 'dır ve

$$B \setminus (B \cap C) = B \cap C^c$$

eşitliği geçerlidir: $x \in B \setminus (B \cap C)$ olması, $x \in B$ ve $x \notin B \cap C$ olması demektir; $x \in B$ olduğuna göre $x \notin B \cap C$ ancak $x \notin C$ ile mümkündür; yani $x \in B$ ve $x \in C^c$. Ters yön de aynı adımların tersidir. $C^c \in \mathcal{U}$ olduğundan $B \setminus A = B \cap C^c \in \mathcal{U}_B$.

(iii*) (A_n) , $\mathcal{U}_B$ 'de bir dizi olsun. Her n için $A_n = B \cap C_n$ olacak biçimde $C_n \in \mathcal{U}$ seçelim. Kesişimin birleşim üzerine dağılmasından

$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} (B \cap C_n) = B \cap \bigcup_{n=1}^{\infty} C_n$$

olur: x 'in soldaki kümede olması, bir n için $x \in B$ ve $x \in C_n$ olması; sağdakinde olması, $x \in B$ ve bir n için $x \in C_n$ olması demektir; bunlar aynı şeydir. $\mathcal{U}$ bir σ -cebir olduğundan $C = \bigcup_n C_n \in \mathcal{U}$ ve $\bigcup_n A_n = B \cap C \in \mathcal{U}_B$. Böylece $\mathcal{U}_B$, B üzerinde bir σ -cebirdir.

Son iddia. $B \in \mathcal{U}$ olsun. $A = B \cap C \in \mathcal{U}_B$ ise $B, C \in \mathcal{U}$ olduğundan [Teorem 2.1](#) (c) gereği $A \in \mathcal{U}$, ayrıca $A \subseteq B$. Tersine $A \in \mathcal{U}$ ve $A \subseteq B$ ise $A = B \cap A$ yazılır; $C = A \in \mathcal{U}$ alınarak $A \in \mathcal{U}_B$ bulunur. Demek ki $\mathcal{U}_B = \{A \in \mathcal{U} : A \subseteq B\}$.

■

2.5 Borel Cebri

Sıcaklık gözlemi gibi deneylerde örnek uzay $\Omega = \mathbb{R}$ ya da $\mathbb{R}$ 'nin bir aralığıdır ve doğal olaylar aralıklardır: “sıcaklık 15 ile 25 derece arasındadır” olayı $(15, 25)$ aralığıdır. Açık aralıklar sınıfı

$$\mathcal{L}_1 = \{(a, b) : a, b \in \mathbb{R}, a < b\}$$

ve yarı açık aralıklar sınıfı

$$\mathcal{L}_2 = \{(a, b] : a, b \in \mathbb{R}, a < b\}$$

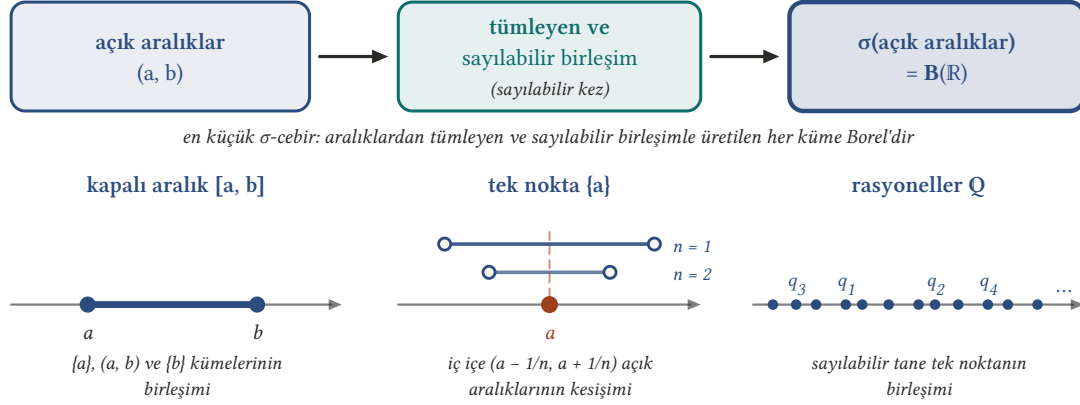
σ -cebir değildir; hatta cebir bile değildir. Gerçekten $\mathbb{R}$ sınırlı bir aralık olmadığından $\mathbb{R} \notin \mathcal{L}_1$; ayrıca $(a, b)^c = (-\infty, a] \cup [b, \infty)$ bir aralık değildir ve $(0, 1) \cup (2, 3)$ de bir aralık değildir; yani (i), (ii), (iii) koşullarının hiçbiri sağlanmaz. $\mathcal{L}_2$ için de aynı gerekçeler geçerlidir. Ama [Teorem 2.3](#) gereği bu sınıfları kapsayan en küçük bir σ -cebir vardır. Olasılık teorisinde $\mathbb{R}$ üzerinde hep bu σ -cebirle çalışılır.

Tanım 2.4 (Borel Cebri). $\mathbb{R}$ üzerindeki açık aralıklar sınıfı $\mathcal{L}_1 = \{(a, b) : a, b \in \mathbb{R}, a < b\}$ olsun. $\mathcal{L}_1$ 'in ürettiği σ -cebre $\mathbb{R}$ üzerindeki **Borel cebri** denir ve $\mathcal{B}$ ya da $\mathcal{B}(\mathbb{R})$ ile gösterilir:

$$\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{L}_1) = \sigma(\{(a, b) : a < b\}).$$

$\mathcal{B}(\mathbb{R})$ 'nin elemanlarına **Borel kümeleri** denir.

Tanım gereği her açık aralık bir Borel kümesidir ve Borel kümelerinden tümleyen, sayılabilir birleşim ve sayılabilir kesişimle elde edilen her küme yine Borel kümesidir. Aşağıdaki önerme en önemli örnekleri toplar; ispatın araçları σ -cebir işlemleri, Arşimet özelliği (her pozitif gerçel sayı için ondan küçük bir $\frac{1}{n}$ vardır) ve son şıkta rasyonel sayıların yoğunluğu ile sayılabilirliği.



Şekil 2.3: Borel cebirinin kuruluşu. Açık aralıklar sınıfından başlanır; tümleyen ve sayılabilir birleşim işlemleri (dolayısıyla sayılabilir kesişim de) sayılabilir kez uygulanarak ulaşılan her küme Borel kümesidir ve bu kümelerin sınıfı, açık aralıkları kapsayan en küçük σ -cebir olan $\mathbf{B}(\mathbb{R})$ 'dir. Altta üç örnek: kapalı aralık $[a, b]$ iki tek nokta ile bir açık aralığın birleşimi, tek nokta $\{a\}$ daralan açık aralıkların kesişimi, $\mathbb{Q}$ ise sayılabilir tane tek noktanın birleşimi olarak Borel'dir.

Önerme 2.5 (Borel Kümelerine Örnekler). $a, b \in \mathbb{R}$ ve $a < b$ olsun.

- (a) Tek nokta kümesi $\{a\}$ bir Borel kümesidir. Dolayısıyla $\mathbb{R}$ 'nin her sonlu ya da sayılabilir sonsuz alt kümesi (örneğin $\{a, b\}, \mathbb{N}, \mathbb{Z}, \mathbb{Q}$) Borel kümesidir.
- (b) $[a, b], [a, b)$ ve $(a, b]$ Borel kümeleridir.
- (c) $(-\infty, a], (-\infty, a), (a, \infty), [a, \infty)$ ve $\mathbb{R}$ Borel kümeleridir.
- (d) $\mathbb{R}$ 'nin her açık alt kümesi ve her kapalı alt kümesi Borel kümesidir.

İspat.

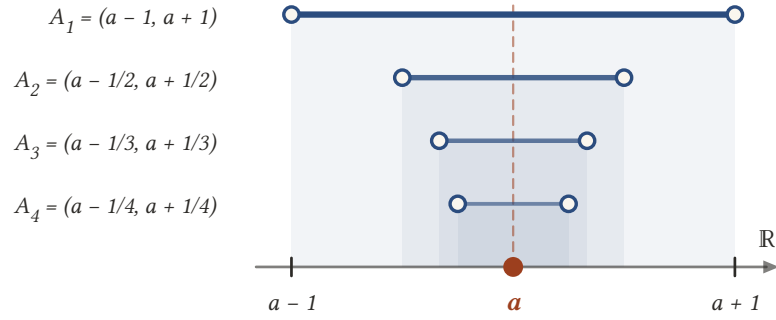
(a) $n = 1, 2, \dots$ için

$$A_n = \left(a - \frac{1}{n}, a + \frac{1}{n} \right)$$

olsun. Her A_n bir açık aralık olduğundan $A_n \in \mathcal{B}(\mathbb{R})$ 'dir. İddia:

$$\bigcap_{n=1}^{\infty} A_n = \{a\}.$$

$\supseteq$: her n için $a - \frac{1}{n} < a < a + \frac{1}{n}$, yani $a \in A_n$; öyleyse a kesişimdedir. $\subseteq$: x kesişimde olsun; o zaman her n için $|x - a| < \frac{1}{n}$ 'dir. $x \neq a$ olsaydı $|x - a| > 0$ olur ve Arşimet özelliği gereği $\frac{1}{n} < |x - a|$ olan bir n bulunurdu; bu, $x \in A_n$ ile çelişir. Demek ki $x = a$.



$$\bigcap_n A_n = \bigcap_n (a - 1/n, a + 1/n) = \{a\}$$

Şekil 2.4: Tek nokta kümesi $\{a\}$, iç içe daralan açık aralıkların kesişimidir. Her $A_n = (a - 1/n, a + 1/n)$ aralığı a 'yı içerir; a 'dan farklı her nokta ise Arşimet özelliği gereği yeterince büyük n için aralığın dışında kalır. Böylece kesişimde yalnızca a kalır ve $\{a\}$ bir Borel kümesidir.

Böylece $\{a\}$, Borel kümelerinden oluşan bir dizinin kesişimidir; Teorem 2.1 (b) gereği $\{a\} \in \mathcal{B}(\mathbb{R})$.

$D = \{d_1, d_2, d_3, \dots\}$ boş olmayan sayılabilir bir küme ise $D = \bigcup_n \{d_n\}$, Borel kümelerinin sayılabilir birleşimidir; (iii*) gereği $D \in \mathcal{B}(\mathbb{R})$. D sonluysa birleşim sonludur ve Teorem 2.1 (c) kullanılır; $D = \emptyset$ ise Teorem 2.1 (a) yeter.

(b) $[a, b] = \{a\} \cup (a, b) \cup \{b\}$: $x \in [a, b]$ ise ya $x = a$, ya $a < x < b$, ya da $x = b$ 'dir; tersi açıktır. Sağ taraf üç Borel kümesinin birleşimidir, (a) ve Teorem 2.1 (c) gereği Borel'dir. Benzer biçimde $[a, b) = \{a\} \cup (a, b)$ ve $(a, b] = (a, b) \cup \{b\}$ Borel kümeleridir.

(c) İddia:

$$(-\infty, a] = \bigcup_{n=1}^{\infty} [a - n, a].$$

$\supseteq$: her n için $[a - n, a] \subseteq (-\infty, a]$. $\subseteq$: $x \leq a$ olsun. Arşimet özelliği gereği $n > a - x$ olan bir $n \in \mathbb{N}$ vardır; o zaman $a - n < x \leq a$, yani $x \in [a - n, a]$. Her $[a - n, a]$ kapalı aralığı (b) gereği Borel'dir; sayılabilir birleşim de Borel'dir. Aynı yolla $(-\infty, a) = \bigcup_{n=1}^{\infty} (a - n, a)$ açık aralıkların birleşimi olarak Borel'dir. Tümleyen alarak $(a, \infty) = (-\infty, a]^c$ ve $[a, \infty) = (-\infty, a)^c$ Borel kümeleridir. $\mathbb{R}$ ise (i) gereği her σ -cebirdedir.

(d) $G \subseteq \mathbb{R}$ açık bir küme olsun. $G = \emptyset$ ise iddia (a) gereği açıktır; $G \neq \emptyset$ olsun. Açıklık tanımı gereği her $x \in G$ için $(x - \varepsilon, x + \varepsilon) \subseteq G$ olacak biçimde bir $\varepsilon > 0$ vardır. Rasyonel sayılar $\mathbb{Q}$ 'de yoğun olduğundan $x - \varepsilon < p < x < q < x + \varepsilon$ olan $p, q \in \mathbb{Q}$ seçilebilir; o zaman $x \in (p, q) \subseteq G$. Demek ki

$$G = \bigcup \{(p, q) : p, q \in \mathbb{Q}, p < q, (p, q) \subseteq G\}.$$

Sağdaki birleşimdeki aralıklar rasyonel sayı çiftleriyle indislenir; $\mathbb{Q} \times \mathbb{Q}$ sayılabilir olduğundan bu, sayılabilir tane açık aralığın birleşimidir ve (iii*) gereği Borel'dir. F kapalı ise F^c açıktır, dolayısıyla Borel'dir; (ii) gereği $F = (F^c)^c$ de Borel'dir.

■

Önermenin (d) şıkkı, Borel cebirinin **açık kümelerin ürettiği σ -cebri** olarak da tanımlanabileceğini gösterir: açık kümeler sınıfı $\mathcal{G}$ olsun; her açık aralık açık küme olduğundan $\mathcal{L}_1 \subseteq \mathcal{G}$ ve dolayısıyla $\sigma(\mathcal{L}_1) \subseteq \sigma(\mathcal{G})$; (d) gereği $\mathcal{G} \subseteq \sigma(\mathcal{L}_1)$ ve dolayısıyla $\sigma(\mathcal{G}) \subseteq \sigma(\mathcal{L}_1)$. Ayrıca alıştırılarda göreceğimiz gibi $\{(-\infty, a] : a \in \mathbb{R}\}$ sınıfı da Borel cebirini üretir; bu, rastgele değişkenleri incelerken $(X \leq a)$ biçimindeki olayların neden yeterli olduğunu açıklayacaktır.

Buraya kadar Borel cebirinin ne kadar zengin olduğunu gördük. Doğal soru şudur: madem neredeyse her küme Borel'dir, neden $\mathcal{P}(\mathbb{R})$ ile çalışmıyoruz?

! Neden her alt küme olay olamaz?

$\Omega = [0, 1]$ aralığından “rastgele bir nokta seçme” deneyini düşünelim. Doğal olasılık, bir aralığın uzunluğudur: $P((a, b)) = b - a$. Bu P 'yi $[0, 1]$ 'in olabildiğince çok alt kümesine, şu üç doğal isteği koruyarak genişletmek isteriz: $P([0, 1]) = 1$; ayrık kümelerden oluşan bir dizinin birleşiminin olasılığı olasılıkların toplamıdır (sayılabilir toplamsallık); bir kümeyi kaydırmak olasılığını değiştirmez (öteleme değişmezliği).

1905'te Giuseppe Vitali, bu üç isteği birden sağlayan ve $[0, 1]$ 'in **bütün** alt kümelerinde tanımlı bir P 'nin **var olamayacağını** gösterdi. Fikir şudur: $[0, 1]$ 'deki iki sayıyı, farkları rasyonel ise “aynı sınıfta” sayalım; bu, $[0, 1]$ 'i sayılamayan çoklukta sınıfa böler. Her sınıftan bir temsilci seçip (seçim aksiyomu) bunları bir V kümesinde toplayalım. V 'nin rasyonel sayılar kadar, yani sayılabilir çoklukta ötelenmiş (1'i aşan kısım başa sarılarak) $[0, 1]$ 'i ayrık parçalar hâlinde örter ve öteleme değişmezliği gereği her parçanın olasılığı $P(V)$ 'dir. $P(V) = 0$ olsaydı bu parçaların olasılıkları toplamı 0, $P(V) > 0$ olsaydı toplam ∞ olurdu; oysa toplam, örtülen $[0, 1]$ kümesinin olasılığına, yani 1'e eşit olmalıdır. Öyleyse V 'ye tutarlı bir olasılık verilemez; V **ölçülemeyen** bir kümedir.

Bu yüzden olasılık ölçüsü (Tanım 3.1) $\mathcal{P}(\Omega)$ üzerinde değil, bir σ -cebri üzerinde tanımlanır. $\mathcal{B}(\mathbb{R})$ bu iş için tam uygundur: aralıkları, açık ve kapalı kümeleri, sayılabilir kümeleri ve bunlardan sayılabilir işlemlerle türeyen her şeyi içerir; ama Vitali kümesi gibi tuhaf kümeleri dışarıda bırakır. Sonlu ya da sayılabilir örnek uzaylarda böyle bir sorun çıkmaz ve $\mathcal{P}(\Omega)$ rahatça kullanılır.

Birden çok sayının birlikte gözlemlendiği deneylerde, örneğin iki zarın atılması ya da bir noktanın düzlemde seçilmesi durumunda, örnek uzay $\mathbb{R}^2$ ya da genel olarak $\mathbb{R}^n$ 'dir. Aralığın çok boyutlu karşılığı açık dikkdörtgendir ve Borel cebri aynı kalıpla kurulur.

Tanım 2.5 (Çok Boyutlu Uzayda Borel Cebri). $n \in \mathbb{N}$ olsun. $\underline{a} = (a_1, a_2, \dots, a_n)$ ve $\underline{b} = (b_1, b_2, \dots, b_n) \in \mathbb{R}^n$, her i için $a_i < b_i$ olacak biçimde verilsin.

$$(\underline{a}, \underline{b}) = \{(x_1, x_2, \dots, x_n) \in \mathbb{R}^n : a_i < x_i < b_i, i = 1, 2, \dots, n\} = (a_1, b_1) \times (a_2, b_2) \times \dots \times (a_n, b_n)$$

kümesine $\mathbb{R}^n$ 'de bir **açık dikkdörtgen** (açık kutu) denir. Açık dikkdörtgenler sınıfı

$$\mathcal{J}_n = \{(\underline{a}, \underline{b}) : \underline{a}, \underline{b} \in \mathbb{R}^n, a_i < b_i\}$$

olmak üzere, $\mathcal{J}_n$ 'in ürettiği σ -cebire $\mathbb{R}^n$ üzerindeki **Borel cebri** denir ve $\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{J}_n)$ ile gösterilir. Özel olarak $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ üzerinde

$$\mathcal{B}(\mathbb{R}^2) = \sigma(\{(a, b) \times (c, d) : a < b, c < d\})$$

olup üreteçler düzlemdeki açık dikkdörtgensel bölgelerdir.

$n = 1$ için açık dikkdörtgenler açık aralıklardır ve tanım Tanım 2.4'ye indirgenir. $\mathbb{R}^2$ 'de de bir boyuttaki bütün sonuçların karşılıkları geçerlidir: kenarları kapalı ya da yarı açık dikkdörtgenler, doğru parçaları, tek noktalar, açık ve kapalı diskler Borel kümeleridir; örneğin düzlemdeki bir açık küme, içinde kalan rasyonel köşeli açık dikkdörtgenlerin sayılabilir birleşimidir. Bu σ -cebri, rastgele vektörlerin dağılımlarını incelerken kullanılacaktır.

2.6 Alıştırmalar

Alıştırma 2.1 (Cebirler, σ -Cebirler ve Borel Cebri). (a) $\Omega = \{a, b\}$ üzerindeki bütün σ -cebirleri belirleyiniz.

(b) $\Omega = \{a, b, c\}$ ve $\mathcal{U} = \{\emptyset, \{a\}, \{b\}, \{a, b\}, \{c\}, \Omega\}$ olsun. $\mathcal{U}$ bir σ -cebri midir? $\sigma(\mathcal{U})$ 'yu bulunuz.

(c) $\Omega = \mathbb{N}$ ve $\mathcal{L} = \{\{n\} : n \in \mathbb{N}\}$ olsun. $\sigma(\mathcal{L}) = \mathcal{P}(\mathbb{N})$ olduğunu gösteriniz.

(d) $\Omega = \mathbb{R}$ olsun ve “sayılabilir” sözcüğü “sonlu ya da sayılabilir sonsuz” anlamında kullanılsın.

$$\mathcal{U} = \{A \subseteq \mathbb{R} : A \text{ sayılabilir ya da } A^c \text{ sayılabilir}\}$$

sınıfının bir σ -cebri olduğunu ve $\mathcal{U} \subsetneq \mathcal{B}(\mathbb{R})$ olduğunu gösteriniz.

(e) $\Omega = \{a, b, c\}$, $\mathcal{U}_1 = \{\emptyset, \{a\}, \{b, c\}, \Omega\}$ ve $\mathcal{U}_2 = \{\emptyset, \{b\}, \{a, c\}, \Omega\}$ olsun. $\mathcal{U}_1$ ve $\mathcal{U}_2$ 'nin σ -cebri olduğunu, $\mathcal{U}_1 \cup \mathcal{U}_2$ 'nin ise cebri bile olmadığını gösteriniz. $\sigma(\mathcal{U}_1 \cup \mathcal{U}_2)$ nedir?

(f) $\mathcal{L} = \{(-\infty, c] : c \in \mathbb{R}\}$ olsun. $\sigma(\mathcal{L}) = \mathcal{B}(\mathbb{R})$ olduğunu gösteriniz.

(g) $\mathcal{K} \subseteq \mathcal{M}$ ise $\sigma(\mathcal{K}) \subseteq \sigma(\mathcal{M})$ olduğunu; her sınıf $\mathcal{L}$ için $\sigma(\sigma(\mathcal{L})) = \sigma(\mathcal{L})$ olduğunu; ayrıca $\mathcal{U}$ bir σ -cebir ise $\sigma(\mathcal{U}) = \mathcal{U}$ olduğunu gösteriniz.

Çözüm.

(a) Her σ -cebir $\emptyset$ ve Ω 'yı içerir. Ω 'nın öteki alt kümeleri $\{a\}$ ve $\{b\}$ 'dir ve bunlar birbirinin tümleyesidir. Bir σ -cebir $\{a\}$ 'yi içeriyorsa (ii) gereği $\{b\}$ 'yi de içerir ve $\mathcal{P}(\Omega)$ 'ya eşit olur; $\{b\}$ 'yi içeriyorsa aynı şey olur. İkisini de içermiyorsa $\{\emptyset, \Omega\}$ 'dir. Bu iki sınıfın gerçekten σ -cebir olduğunu biliyoruz (Örnek 2.1'daki $\mathcal{U}_3$ ile aynı gerekçe ve Önerme 2.3). Demek ki $\Omega = \{a, b\}$ üzerinde tam iki σ -cebir vardır:

$$\{\emptyset, \Omega\} \quad \text{ve} \quad \mathcal{P}(\Omega) = \{\emptyset, \{a\}, \{b\}, \Omega\}.$$

(b) $\{a\} \in \mathcal{U}$ ama $\{a\}^c = \{b, c\} \notin \mathcal{U}$; (ii) bozulur, $\mathcal{U}$ bir σ -cebir değildir. $\sigma(\mathcal{U})$, $\{a\}$, $\{b\}$ ve $\{c\}$ tek nokta kümelerini içerir; Ω 'nın boş olmayan her alt kümesi bu tek nokta kümelerinden bazılarının birleşimi olduğundan, Örnek 2.2 (c)'deki gerekçeyle $\mathcal{P}(\Omega) \subseteq \sigma(\mathcal{U}) \subseteq \mathcal{P}(\Omega)$. Öyleyse $\sigma(\mathcal{U}) = \mathcal{P}(\Omega)$, sekiz elemanlı kuvvet kümesi.

(c) $\sigma(\mathcal{L}) \subseteq \mathcal{P}(\mathbb{N})$ açıktır. Tersini için $A \subseteq \mathbb{N}$ herhangi bir küme olsun. $A = \emptyset$ ise $A \in \sigma(\mathcal{L})$ 'dir. $A \neq \emptyset$ ise A sonlu ya da sayılabilir sonsuzdur; elemanlarını $A = \{n_1, n_2, \dots\}$ diye sıralayalım. O zaman

$$A = \bigcup_k \{n_k\}$$

sonlu ya da sayılabilir tane tek nokta kümesinin birleşimidir. Her $\{n_k\} \in \mathcal{L} \subseteq \sigma(\mathcal{L})$ olduğundan, (iii*) ya da Teorem 2.1 (c) gereği $A \in \sigma(\mathcal{L})$. Demek ki $\sigma(\mathcal{L}) = \mathcal{P}(\mathbb{N})$.

(d) (i) $\mathbb{R}^c = \emptyset$ sayılabilir; $\mathbb{R} \in \mathcal{U}$. (ii) $A \in \mathcal{U}$ olsun. A sayılabilirse $(A^c)^c = A$ sayılabilir, dolayısıyla $A^c \in \mathcal{U}$; A^c sayılabilirse A^c doğrudan $\mathcal{U}$ 'dadır. (iii*) (A_n) , $\mathcal{U}$ 'da bir dizi olsun. Bütün A_n 'ler sayılabilirse $\bigcup_n A_n$, sayılabilir tane sayılabilir kümenin birleşimi olarak sayılabilir ve $\mathcal{U}$ 'dadır. Değilse en az bir A_m sayılamaz; $A_m \in \mathcal{U}$ olduğundan A_m^c sayılabilir ve

$$\left(\bigcup_{n=1}^{\infty} A_n\right)^c = \bigcap_{n=1}^{\infty} A_n^c \subseteq A_m^c$$

sayılabilir bir kümenin alt kümesi olarak sayılabilir; yine $\bigcup_n A_n \in \mathcal{U}$. Öyleyse $\mathcal{U}$ bir σ -cebirdir.

$\mathcal{U} \subseteq \mathcal{B}(\mathbb{R})$: A sayılabilirse Önerme 2.5 (a) gereği Borel'dir; A^c sayılabilirse A^c Borel'dir ve $A = (A^c)^c$ de Borel'dir. Eşitlik yoktur: $(0, 1) \in \mathcal{B}(\mathbb{R})$ ama $(0, 1)$ sayılamaz ve tümleyeni $(-\infty, 0] \cup [1, \infty)$ da sayılamaz; yani $(0, 1) \notin \mathcal{U}$. Demek ki $\mathcal{U} \subsetneq \mathcal{B}(\mathbb{R})$.

(e) $\mathcal{U}_1$ için: $\Omega \in \mathcal{U}_1$; tümleyen çiftleri $(\emptyset, \Omega)$ ve $(\{a\}, \{b, c\})$; birleşimler $\{a\} \cup \{b, c\} = \Omega$ dışında hep elemanlardan birine eşittir. Ω sonlu olduğundan Önerme 2.2 gereği $\mathcal{U}_1$ bir σ -cebirdir. $\mathcal{U}_2$ için a ile b 'nin rolleri değişir; aynı gerekçe geçerlidir.

$\mathcal{U}_1 \cup \mathcal{U}_2 = \{\emptyset, \{a\}, \{b\}, \{b, c\}, \{a, c\}, \Omega\}$ sınıfında $\{a\}$ ve $\{b\}$ vardır ama $\{a\} \cup \{b\} = \{a, b\}$ yoktur; (iii) bozulur, sınıf bir cebir değildir.

$\sigma(\mathcal{U}_1 \cup \mathcal{U}_2)$, $\{a\}$ ve $\{b\}$ 'yi, dolayısıyla $\{a, b\}^c = \{c\}$ 'yi içerir. Üç tek nokta kümesini içeren bir σ -cebir, (b) şıkkındaki gerekçeyle $\mathcal{P}(\Omega)$ 'dir: $\sigma(\mathcal{U}_1 \cup \mathcal{U}_2) = \mathcal{P}(\Omega)$.

(f) $\sigma(\mathcal{L}) \subseteq \mathcal{B}(\mathbb{R})$: Önerme 2.5 (c) gereği her $(-\infty, c]$ Borel'dir; yani $\mathcal{L} \subseteq \mathcal{B}(\mathbb{R})$. $\mathcal{B}(\mathbb{R})$ bir σ -cebir olduğundan Teorem 2.3 (b) gereği $\sigma(\mathcal{L}) \subseteq \mathcal{B}(\mathbb{R})$.

$\mathcal{B}(\mathbb{R}) \subseteq \sigma(\mathcal{L})$: Her açık aralığın $\sigma(\mathcal{L})$ 'de olduğunu göstermek yeter; o zaman $\sigma(\mathcal{L})$, $\mathcal{L}_1$ 'i kapsayan bir σ -cebir olur ve Teorem 2.3 (b) gereği $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{L}_1) \subseteq \sigma(\mathcal{L})$. $a < b$ olsun. Önce

$$(a, b] = (-\infty, b] \setminus (-\infty, a]$$

yazılır: $x \leq b$ ve $x \not\leq a$ olması, $a < x \leq b$ olması demektir. Teorem 2.1 (d) gereği $(a, b] \in \sigma(\mathcal{L})$. Sonra

$$(a, b) = \bigcup_{n=1}^{\infty} \left(a, b - \frac{1}{n}\right]$$

eşitliğini gösterelim; burada $b - \frac{1}{n} \leq a$ olan n 'ler için sağdaki küme boştur ve birleşime katkısı yoktur. $\supseteq$: $a < x \leq b - \frac{1}{n} < b$ ise $x \in (a, b)$. $\subseteq$: $x \in (a, b)$ ise $b - x > 0$; Arşimet özelliği gereği $\frac{1}{n} < b - x$ olan bir n vardır ve o zaman $a < x < b - \frac{1}{n}$, yani $x \in (a, b - \frac{1}{n}]$. Her $(a, b - \frac{1}{n}]$ kümesi $\sigma(\mathcal{L})$ 'de (boş olanlar Teorem 2.1 (a) gereği) olduğundan (iii*) gereği $(a, b) \in \sigma(\mathcal{L})$. Demek ki $\sigma(\mathcal{L}) = \mathcal{B}(\mathbb{R})$. Aynı iki adım, yarı açık aralıklar sınıfı $\mathcal{L}_2 = \{(a, b] : a < b\}$ için de $\sigma(\mathcal{L}_2) = \mathcal{B}(\mathbb{R})$ verir: her $(a, b]$,

Önerme 2.5 (b) gereği Borel'dir, dolayısıyla $\sigma(\mathcal{L}_2) \subseteq \mathcal{B}(\mathbb{R})$; tersine her açık aralık, yukarıdaki eşitlik gereği $\mathcal{L}_2$ 'deki kümelerin sayılabilir birleşimidir, dolayısıyla $\mathcal{B}(\mathbb{R}) \subseteq \sigma(\mathcal{L}_2)$.

(g) $\mathcal{K} \subseteq \mathcal{M}$ olsun. $\sigma(\mathcal{M})$, $\mathcal{M}$ 'yi, dolayısıyla $\mathcal{K}$ 'yi kapsayan bir σ -cebirdir; **Teorem 2.3** (b) gereği $\sigma(\mathcal{K}) \subseteq \sigma(\mathcal{M})$.

İkinci iddia için: $\sigma(\mathcal{L})$, kendisini kapsayan bir σ -cebirdir; (b) gereği $\sigma(\sigma(\mathcal{L})) \subseteq \sigma(\mathcal{L})$. Öte yandan (a) gereği her sınıf ürettiği σ -cebirin içindedir: $\sigma(\mathcal{L}) \subseteq \sigma(\sigma(\mathcal{L}))$. İki kapsama eşitliği verir.

Son iddia, ikinci iddianın genel hâlidir ve aynı iki adımla doğrudan çıkar: $\mathcal{U}$ bir σ -cebirse, $\mathcal{U}$ 'yu kapsayan σ -cebirlerden biri kendisidir; (b) gereği $\sigma(\mathcal{U}) \subseteq \mathcal{U}$, (a) gereği $\mathcal{U} \subseteq \sigma(\mathcal{U})$ 'dur.

■

Artık olayların hangi sınıfta yaşadığını biliyoruz; sırada bu olaylara sayı atayan fonksiyonun, olasılık ölçüsünün kuralları var: **Olasılık Ölçüsü ve Olasılık Uzayı**.

3. Olasılık Ölçüsü ve Olasılık Uzayı

Cebirler, σ -Cebirler ve Borel Cebri bölümünde bir rastgele deneyin olası sonuçlarını örnek uzay Ω ile, “olasılığı sorulabilecek” olayları ise Ω üzerindeki bir σ -cebir $\mathcal{U}$ ile modelledik. Böylece elimizde $(\Omega, \mathcal{U})$ ikilisi var; ama henüz hiçbir olaya bir sayı atamadık. Bu bölümde eksik kalan üçüncü bileşeni, **olasılık ölçüsü** P ’yi tanımlıyoruz.

Olasılığı tanımlamanın eski yolları — eş olasılıklı sonuçları saymak ya da bir deneyin çok sayıda tekrarında gözlenen göreceli sıklığa bakmak — belli durumlarda iyi çalışır, ama genel bir kuram için yetersizdir. 1933’te Kolmogorov, olasılığı “ne olduğu” ile değil “hangi kurallara uyduğu” ile tanımlamayı önerdi: olasılık, olaylar σ -cebri üzerinde tanımlı ve üç basit aksiyomu sağlayan bir fonksiyondur. Bu bakış açısı hem sayma yoluyla hem de göreceli sıklıkla elde edilen olasılıkları kapsar ve olasılık teorisini ölçü teorisinin sağlam zeminine oturtur.

Bölümde önce aksiyomları ve olasılık uzayı kavramını vereceğiz. Ardından yalnızca üç aksiyomdan yola çıkarak olasılık hesaplarında her gün kullanılan kuralları — tümleyen kuralı, monotonluk, toplama kuralı, içerme-dışarma formülü, Boole eşitsizliği — tam ispatlarıyla türeteceğiz. Bölümün sonunda, sayılabilir toplamsallık aksiyomunun asıl gücünü gösteren **olasılığın sürekliliği** özelliğini ispatlayacağız.

3.1 Kolmogorov Aksiyomları

Bir olasılık ölçüsünden ne bekleriz? Birincisi, hiçbir olayın olasılığı negatif olmamalı. İkincisi, kesin olay Ω ’nın olasılığı 1 olmalı; olasılıkları 0 ile 1 arasına yerleştiren ölçek budur. Üçüncüsü, birbirini dışlayan olaylardan “en az birinin” gerçekleşme olasılığı, tek tek olasılıkların toplamı olmalı. Bu üçüncü kuralı yalnızca iki olay için değil, sayılabilir çoklukta olay için istemek, kuramın limit işlemleriyle uyumlu olmasını sağlar; nedenini bölümün sonunda göreceğiz.

Tanım 3.1 (Olasılık Ölçüsü (Kolmogorov Aksiyomları)). $\Omega \neq \emptyset$ bir örnek uzay ve $\mathcal{U}$, Ω üzerinde bir σ -cebir olsun.

$$P : \mathcal{U} \rightarrow \mathbb{R}, \quad A \mapsto P(A)$$

fonksiyonu aşağıdaki üç koşulu sağlıyorsa P ’ye $\mathcal{U}$ üzerinde bir **olasılık ölçüsü** denir:

(A1) Her $A \in \mathcal{U}$ için $P(A) \geq 0$.

(A2) $P(\Omega) = 1$.

(A3) $\mathcal{U}$ ’daki ikişer ikişer ayrık olayların her $(A_n)_{n \geq 1}$ dizisi için (yani $i \neq j$ iken $A_i \cap A_j = \emptyset$ olduğunda)

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n).$$

$P(A)$ sayısına A olayının **olasılığı** denir. (A3) koşuluna **sayılabilir toplamsallık** (σ -toplamsallık) denir.

Tanımda birkaç noktaya dikkat edelim. $\mathcal{U}$ bir σ -cebir olduğundan (A3)’teki sayılabilir birleşim yine $\mathcal{U}$ ’dadır; yani eşitliğin sol tarafı anlamlıdır. Sağ taraf ise negatif olmayan terimli bir seridir; (A3) bu serinin yakınsak ve toplamının tam olarak $P\left(\bigcup A_n\right)$ olduğunu söyler. Ayrıca tanımda P ’nin değerlerinin $[0, 1]$ aralığında olduğu **varsayılmamıştır**; bunu aksiyomlardan çıkaracağız.

$P(A)$ sayısının yorumu şudur: deney bir kez yapıldığında A olayının gerçekleşme “şansı” $P(A)$ ’dır. $P(A)$ sayısı 1’e ne kadar yakınsa A ’nın gerçekleşmesini o kadar çok, 0’a ne kadar yakınsa o kadar az bekleriz. Aksiyomlar bu sezgiyi, deneyin çok sayıda tekrarındaki göreceli sıklıkların uyduğu kurallardan alır: göreceli sıklık negatif olamaz, kesin olayın göreceli sıklığı 1’dir ve ayrık olayların göreceli sıklıkları toplanır.

Tanım 3.2 (Olasılık Uzayı). $\Omega \neq \emptyset$ bir örnek uzay, $\mathcal{U}$, Ω üzerinde bir σ -cebir ve P , $\mathcal{U}$ üzerinde bir olasılık ölçüsü olsun. $(\Omega, \mathcal{U}, P)$ üçlüsüne **olasılık uzayı** denir.

Olasılık uzayının üç bileşeni bir rastgele deneyin üç yönünü modeller: Ω deneyin **olası sonuçlarını**, $\mathcal{U}$ deney hakkında **sorabileceğimiz soruları** (olayları), P ise bu soruların **yanıtlarını** (olasılıkları) verir. Bundan sonra “olay” dediğimizde her zaman $\mathcal{U}$ ’nun bir elemanını kastedeceğiz; $\mathcal{U}$ dışındaki bir Ω alt kümesinin olasılığından söz edilemez.

i Sonlu örnek uzaylarda olasılık ölçüsü

$\Omega = \{\omega_1, \dots, \omega_k\}$ sonlu bir küme, $\mathcal{U} = \mathcal{P}(\Omega)$ ve $p_1, \dots, p_k$ toplamı 1 olan negatif olmayan sayılar olsun. Her $A \subset \Omega$ için

$$P(A) = \sum_{\omega_i \in A} p_i$$

tanımlansın. (A1) açıktır; (A2) $P(\Omega) = p_1 + \dots + p_k = 1$ eşitliğidir. (A3) için ikişer ikişer ayrık bir (A_n) dizisi alalım. Boş olmayan ayrık kümeler Ω ’nın farklı elemanlarını içermek zorunda olduğundan dizide en çok k tane boş olmayan küme vardır; ötekiler $\emptyset$ ’dir ve $P(\emptyset) = 0$ ’dır (boş toplam). Boş olmayan ayrık kümelerin birleşimi üzerindeki toplam, her bir kümenin toplamının toplamıdır; dolayısıyla (A3) sağlanır. Demek ki sonlu bir örnek uzayda olasılık ölçüsü vermek, her sonuca toplamı 1 olan bir “ağırlık” vermekten ibarettir. Bir para atışı için $\Omega = \{Y, T\}$ (yazı ve tura), $p_1 = p_2 = \frac{1}{2}$ alınır; hileli bir para için $p_1 = p$, $p_2 = 1 - p$ alınabilir. Bu tür uzayların ayrıntılı incelemesi sonraki bölümdedir.

3.2 Olasılık Ölçüsünün Temel Özellikleri

Aksiyomlar $P(\emptyset)$ ’nin değerini, A^c ’nin olasılığını, hatta $P(A) \leq 1$ olup olmadığını doğrudan söylemez. Bütün bunlar aksiyomlardan çıkar; aşağıdaki teorem bu temel sonuçları toplar. Her hesapta bunları kullanacağımız için ispatları dikkatle izlemek gerekir.

Teorem 3.1 (Olasılık Ölçüsünün Temel Özellikleri). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı olsun.

a) $P(\emptyset) = 0$.

b) **(Sonlu toplamsallık)** $A_1, A_2, \dots, A_n \in \mathcal{U}$ ikişer ikişer ayrık olaylar ise

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i).$$

c) **(Tümleyen kuralı)** Her $A \in \mathcal{U}$ için $P(A^c) = 1 - P(A)$.

d) **(Monotonluk ve fark kuralı)** $A, B \in \mathcal{U}$ ve $A \subset B$ ise

$$P(A) \leq P(B) \quad \text{ve} \quad P(B \setminus A) = P(B) - P(A).$$

e) Her $A \in \mathcal{U}$ için $0 \leq P(A) \leq 1$.

İspat.

a) $A_1 = \Omega$ ve $n \geq 2$ için $A_n = \emptyset$ alalım. Bu dizi ikişer ikişer ayrıktır: $i \neq j$ iken kümelerden en az biri boştur, dolayısıyla kesişim boştur. Birleşimleri Ω ’dır. (A3) gereği

$$1 = P(\Omega) = P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n) = P(\Omega) + \sum_{n=2}^{\infty} P(\emptyset) = 1 + \sum_{n=2}^{\infty} P(\emptyset).$$

Sağdaki serinin bütün terimleri aynı $P(\emptyset) \geq 0$ sayısıdır. $P(\emptyset) > 0$ olsaydı kısmi toplamlar sınırsız büyür ve seri iraksardı; oysa (A3) serinin yakınsak olduğunu söyler. Demek ki $P(\emptyset) = 0$ ’dır. Burada doğrudan $1 = P(\Omega \cup \emptyset) = P(\Omega) + P(\emptyset)$ yazmak yeterli olmazdı: iki ayrık olay için toplamsallık henüz elimizde değildir; onu (b)’de, (a)’yı kullanarak ispatlayacağız. Bu yüzden (a) doğrudan (A3)’ten çıkarılmalıdır.

b) $A_1, \dots, A_n$ ikişer ikişer ayrık olsun. Diziyi boş kümelerle sonsuza uzatalım: n ’den büyük her i için $A_i = \emptyset$ olsun. Yeni $(A_i)_{i \geq 1}$ dizisi de ikişer ikişer ayrıktır ve

$$\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^n A_i$$

olur. (A3) ve (a) gereği

$$P\left(\bigcup_{i=1}^n A_i\right) = P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) = \sum_{i=1}^n P(A_i) + \sum_{i=n+1}^{\infty} P(\emptyset) = \sum_{i=1}^n P(A_i).$$

c) $A \cup A^c = \Omega$ ve $A \cap A^c = \emptyset$ olduğundan (b) ve (A2) gereği

$$1 = P(\Omega) = P(A \cup A^c) = P(A) + P(A^c),$$

yani $P(A^c) = 1 - P(A)$ 'dır.

d) $A \subset B$ olsun. B 'nin her elemanı ya A 'dadır ya da A 'da değildir; öyleyse

$$B = A \cup (B \cap A^c) = A \cup (B \setminus A)$$

ve bu iki küme ayrıktır. $B \setminus A = B \cap A^c \in \mathcal{U}$ olduğunu biliyoruz. (b) gereği

$$P(B) = P(A) + P(B \setminus A).$$

(A1) gereği $P(B \setminus A) \geq 0$ olduğundan $P(B) \geq P(A)$ elde edilir; eşitliği yeniden düzenlersek $P(B \setminus A) = P(B) - P(A)$ bulunur.

e) Her $A \in \mathcal{U}$ için $\emptyset \subset A \subset \Omega$ 'dır. (d)'deki monotonluk, (a) ve (A2) ile

$$0 = P(\emptyset) \leq P(A) \leq P(\Omega) = 1$$

olur.

■

Teoremin (d) şıkkındaki fark kuralında $A \subset B$ koşulu gereklidir; genel A, B için $P(B \setminus A) = P(B) - P(A \cap B)$ yazılır, çünkü $B \setminus A = B \setminus (A \cap B)$ ve $A \cap B \subset B$ 'dir. Bu biçimi aşağıdaki örneklerde sık kullanacağız.

⚠ Sıfır olasılık imkânsızlık demek değildir

$P(\emptyset) = 0$ ve $P(\Omega) = 1$ olduğunu gösterdik. Bunların tersleri doğru değildir: $P(A) = 0$ olan bir olay boş küme olmak zorunda değildir ve $P(A) = 1$ olan bir olay Ω 'ya eşit olmak zorunda değildir. Sonraki bölümde $[0, 1]$ aralığından rastgele bir nokta seçme deneyini inceleyeceğiz; orada tek bir noktadan oluşan her olayın olasılığı 0 olacaktır, oysa deney her seferinde bir nokta üretir. Bu yüzden $P(A) = 0$ olan olaylara “imkânsız” değil, **hemen hemen imkânsız**; $P(A) = 1$ olan olaylara da **hemen hemen kesin** denir.

3.3 Toplama Kuralı ve İçerme–Dışarma Formülü

Sonlu toplamsallık yalnızca ayrık olaylar için geçerlidir. A ile B ortak sonuçlar içeriyorsa $P(A) + P(B)$ toplamı $A \cap B$ 'yi iki kez sayar; bu fazlalığı geri almak gerekir.

Teorem 3.2 (Toplama Kuralı). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı ve $A, B \in \mathcal{U}$ olsun. O zaman

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Özel olarak her $A, B \in \mathcal{U}$ için $P(A \cup B) \leq P(A) + P(B)$ 'dir.

İspat.

$A \cup B$ 'nin bir elemanı ya A 'dadır ya da A 'da olmayıp B 'dedir. Öyleyse

$$A \cup B = A \cup (A^c \cap B)$$

ve sağdaki iki küme ayrıktır (A ile $A^c \cap B$ kesişemez). [Teorem 3.1](#) (b) gereği

$$P(A \cup B) = P(A) + P(A^c \cap B).$$

Şimdi B 'yi de A 'ya göre parçalayalım: B 'nin her elemanı ya A 'dadır ya da değildir, yani

$$B = (A \cap B) \cup (A^c \cap B)$$

ve bu iki küme ayrıktır. Yine sonlu toplamsallıkla

$$P(B) = P(A \cap B) + P(A^c \cap B), \quad \text{yani} \quad P(A^c \cap B) = P(B) - P(A \cap B).$$

Bunu ilk eşitlikte yerine koyarsak

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

elde edilir. Son olarak (A1) gereği $P(A \cap B) \geq 0$ olduğundan $P(A \cup B) \leq P(A) + P(B)$ 'dir.

■

$A \cap B = \emptyset$ ise $P(A \cap B) = 0$ olur ve toplama kuralı sonlu toplamsallığa indirgenir; yani toplama kuralı, iki olay için sonlu toplamsallığın ayrık olmayan olaylara genellemesidir. Şimdi bu kurallarla neler hesaplanabildiğini görelim.

Örnek 3.1 (İki Olay Üzerine Olasılık Hesapları). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı, $A, B \in \mathcal{U}$ ve

$$P(A) = 0,6, \quad P(B) = 0,5, \quad P(A \cap B) = 0,3$$

olsun. Aşağıdaki olasılıkları hesaplayınız:

a) $P(A \cup B)$; b) $P(A^c)$ ve $P(B^c)$; c) $P(A \setminus B)$ ve $P(B \setminus A)$; d) $P(A^c \cap B^c)$; e) $P(A^c \cup B^c)$; f) A ile B 'den **tam olarak birinin** gerçekleşme olasılığı; g) $P(A^c \cup B)$.

Çözüm.

a) Toplama kuralı (Teorem 3.2) ile

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0,6 + 0,5 - 0,3 = 0,8.$$

b) Tümlen kuralı (Teorem 3.1 c) ile $P(A^c) = 1 - 0,6 = 0,4$ ve $P(B^c) = 1 - 0,5 = 0,5$.

c) $A \setminus B = A \setminus (A \cap B)$ ve $A \cap B \subset A$ olduğundan fark kuralı (Teorem 3.1 d) ile

$$P(A \setminus B) = P(A) - P(A \cap B) = 0,6 - 0,3 = 0,3, \quad P(B \setminus A) = P(B) - P(A \cap B) = 0,5 - 0,3 = 0,2.$$

d) De Morgan kuralı ile $A^c \cap B^c = (A \cup B)^c$ 'dir. Tümlen kuralı ve (a) ile

$$P(A^c \cap B^c) = 1 - P(A \cup B) = 1 - 0,8 = 0,2.$$

Bu, “ A da B de gerçekleşmez” olayının olasılığıdır.

e) Yine De Morgan ile $A^c \cup B^c = (A \cap B)^c$, dolayısıyla

$$P(A^c \cup B^c) = 1 - P(A \cap B) = 1 - 0,3 = 0,7.$$

Aynı sonuca toplama kuralıyla da varılır: $P(A^c) + P(B^c) - P(A^c \cap B^c) = 0,4 + 0,5 - 0,2 = 0,7$.

f) “Tam olarak biri gerçekleşir” olayı $(A \setminus B) \cup (B \setminus A)$ ayrık birleşimdir (A 'nın B dışındaki kısmı ile B 'nin A dışındaki kısmı kesişemez). Sonlu toplamsallık ve (c) ile

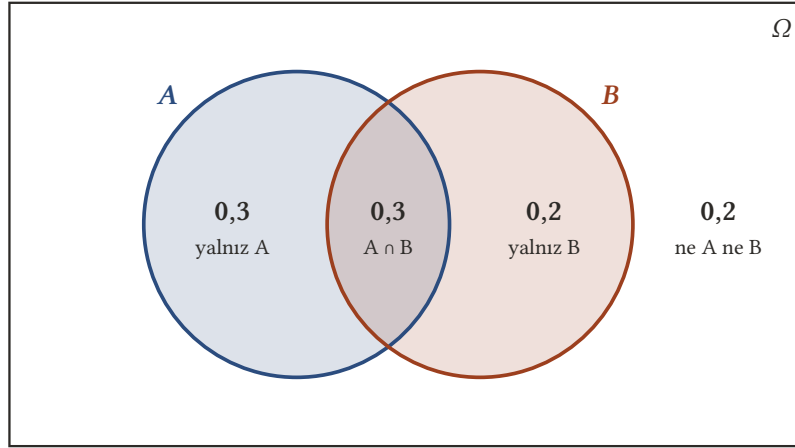
$$P((A \setminus B) \cup (B \setminus A)) = 0,3 + 0,2 = 0,5.$$

Buna A ile B 'nin **simetrik farkı** $A \triangle B$ denir; genel olarak $P(A \triangle B) = P(A \cup B) - P(A \cap B)$ 'dir (Alıştırma 3.1, (c) şıkkı).

g) $(A^c \cup B)^c = A \cap B^c = A \setminus B$ olduğundan

$$P(A^c \cup B) = 1 - P(A \setminus B) = 1 - 0,3 = 0,7.$$

Kontrol: Ω , dört ayrık parçaya ayrılır: $A \cap B$, $A \setminus B$, $B \setminus A$ ve $A^c \cap B^c$. Olasılıkları $0,3 + 0,3 + 0,2 + 0,2 = 1 = P(\Omega)$ 'dir; hesaplar tutarlıdır. Bütün soruların yanıtı bu dört parçanın olasılıklarından toplama ile okunabilir.



$$P(\Omega) = 0,3 + 0,3 + 0,2 + 0,2 = 1$$

Şekil 3.1: $P(A) = 0,6$, $P(B) = 0,5$, $P(A \cap B) = 0,3$ olan iki olayın Venn şeması. Dikdörtgen örnek uzay Ω 'yı gösterir ve dört ayırık parçaya ayrılır: yalnız A ($A \setminus B$), hem A hem B ($A \cap B$), yalnız B ($B \setminus A$) ve ikisi de değil ($A^c \cap B^c$). Parçaların olasılıkları 0,3; 0,3; 0,2 ve 0,2'dir; toplamı 1 = $P(\Omega)$ eder. Örnekteki her soru, bu dört sayıdan uygun olanları toplayarak yanıtlanır.

■

Toplama kuralını üç, dört ya da daha çok olaya genişletirken aynı düşünce geçerlidir: fazladan sayılan kesişimleri çıkarır, fazla çıkardıklarımızı geri ekleriz. Sonuç **içerme-dışarma formülüdür**.

Teorem 3.3 (İçerme-Dışarma Formülü). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı ve $A_1, A_2, \dots, A_n \in \mathcal{U}$ ayırık olmaları gerekmeyen olaylar olsun. O zaman

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i \cap A_j) + \sum_{1 \leq i < j < k \leq n} P(A_i \cap A_j \cap A_k) - \dots + (-1)^{n-1} P(A_1 \cap A_2 \cap \dots \cap A_n).$$

Kısaca, $k = 1, \dots, n$ için k 'lı kesişimlerin olasılıkları toplamı

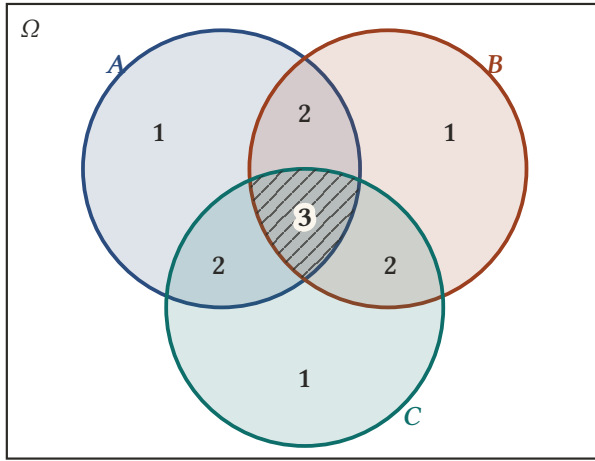
$$S_k = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k})$$

ile gösterilirse

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{k=1}^n (-1)^{k-1} S_k.$$

Özel olarak $n = 3$ için, $A, B, C \in \mathcal{U}$ olmak üzere

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C).$$



sayılar: $P(A) + P(B) + P(C)$ 'de kaç kez sayıldığı

$$P(A \text{ veya } B \text{ veya } C) =$$

$$+ P(A) + P(B) + P(C) \\ - P(A \cap B) - P(A \cap C) - P(B \cap C) \\ + P(A \cap B \cap C)$$

her bölge tam bir kez sayılır:

$$\text{tek olayda: } 1 - 0 + 0 = 1 \text{ kez}$$

$$\text{iki olayda: } 2 - 1 + 0 = 1 \text{ kez}$$

$$\text{üç olayda: } 3 - 3 + 1 = 1 \text{ kez}$$

(taralı bölge üç ikili kesişimde de çıkarılır,
üçlü kesişimde bir kez geri eklenir)

Şekil 3.2: Üç olay için içme–dışarma formülünün Venn şeması. Örnek uzay Ω içindeki A, B, C daireleri yedi ayrık bölge oluşturur; her bölgedeki sayı, o bölgenin $P(A) + P(B) + P(C)$ toplamında kaç kez sayıldığı gösterir. İkili kesişimler çıkarılınca iki olaya ait bölgeler $2 - 1 = 1$ kez, ortadaki taralı bölge ise $3 - 3 = 0$ kez sayılmış olur; üçlü kesişim geri eklenince o da 1'e tamamlanır. Böylece formül, birleşimin her parçasını tam bir kez sayar.

İspat.

n üzerinde tümevarım yapalım. $n = 1$ için formül $P(A_1) = S_1 = P(A_1)$ der; doğrudur. $n = 2$ için formül $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$ biçimini alır; bu da Teorem 3.2'dir.

Şimdi formülün herhangi n olay için doğru olduğunu varsayalım ve $A_1, \dots, A_n, A_{n+1} \in \mathcal{U}$ olsun. $B = \bigcup_{i=1}^n A_i$ diyelim; $B \in \mathcal{U}$ 'dur. Toplama kuralı ile

$$P\left(\bigcup_{i=1}^{n+1} A_i\right) = P(B \cup A_{n+1}) = P(B) + P(A_{n+1}) - P(B \cap A_{n+1}).$$

Sağdaki üç terimden ilkinde ve sonuncusuna tümevarım varsayımını uygulayacağız.

$P(B)$ terimi. $A_1, \dots, A_n$ için tümevarım varsayımı $P(B) = \sum_{k=1}^n (-1)^{k-1} S_k$ verir; burada S_k , $A_1, \dots, A_n$ olaylarının k 'li kesişim toplamıdır.

$P(B \cap A_{n+1})$ terimi. Kesişimin birleşim üzerine dağılma özelliği ile

$$B \cap A_{n+1} = \bigcup_{i=1}^n (A_i \cap A_{n+1}).$$

$C_i = A_i \cap A_{n+1}$ ($i = 1, \dots, n$) olaylarına tümevarım varsayımını uygulayalım. $C_{i_1} \cap \dots \cap C_{i_k} = A_{i_1} \cap \dots \cap A_{i_k} \cap A_{n+1}$ olduğundan

$$P(B \cap A_{n+1}) = \sum_{k=1}^n (-1)^{k-1} T_k, \quad T_k = \sum_{1 \leq i_1 < \dots < i_k \leq n} P(A_{i_1} \cap \dots \cap A_{i_k} \cap A_{n+1}).$$

Yani T_k , A_{n+1} 'i içeren $(k+1)$ 'li kesişimlerin olasılıkları toplamıdır.

Toplamların karşılaştırılması. $A_1, \dots, A_{n+1}$ olaylarının k 'li kesişim toplamını S'_k ile gösterelim. S'_k 'deki her terim $\{1, \dots, n+1\}$ 'den seçilmiş k indisin kesişimidir. Bu terimleri ikiye ayıralım: indisler arasında $n+1$ yoksa terim S'_k 'nin bir terimidir; $n+1$ varsa terim, $\{1, \dots, n\}$ 'den seçilmiş $k-1$ indis ile A_{n+1} 'in kesişimidir, yani T_{k-1} 'in bir terimidir. Böylece

$$S'_1 = S_1 + P(A_{n+1}), \quad S'_k = S_k + T_{k-1} \quad (2 \leq k \leq n), \quad S'_{n+1} = T_n$$

olur; sonuncusu, $n+1$ indisin hepsini almanın tek yolu olduğundan gelir.

Şimdi toplama kuralından gelen eşitliği yazalım ve ikinci toplamda k yerine $m = k+1$ koyalım:

$$\begin{aligned}
P\left(\bigcup_{i=1}^{n+1} A_i\right) &= \sum_{k=1}^n (-1)^{k-1} S_k + P(A_{n+1}) - \sum_{k=1}^n (-1)^{k-1} T_k \\
&= \sum_{k=1}^n (-1)^{k-1} S_k + P(A_{n+1}) + \sum_{m=2}^{n+1} (-1)^{m-1} T_{m-1} \\
&= [S_1 + P(A_{n+1})] + \sum_{m=2}^n (-1)^{m-1} [S_m + T_{m-1}] + (-1)^n T_n \\
&= S'_1 + \sum_{m=2}^n (-1)^{m-1} S'_m + (-1)^n S'_{n+1} \\
&= \sum_{m=1}^{n+1} (-1)^{m-1} S'_m.
\end{aligned}$$

Bu, $n+1$ olay için içerme-dışarma formülüdür. Tümevarım ilkesi gereği formül her n için doğrudur. $n=3$ için $S_1 = P(A) + P(B) + P(C)$, $S_2 = P(A \cap B) + P(A \cap C) + P(B \cap C)$ ve $S_3 = P(A \cap B \cap C)$ yazılarak üç olaylı biçim elde edilir.

■

Formülün adı yöntemi anlatır: tek tek olasılıkları toplarken ikili kesişimler iki kez içerilir, bunlar çıkarılarak dışarıda bırakılır; ama bu kez üçlü kesişimler fazla çıkarılmış olur, geri eklenir ve böyle sürer. Üç olaylı biçimin sayısal bir uygulaması [Alıştırma 3.1](#)'nin (b) şikkındadır.

3.4 Boole Eşitsizliği

Çoğu uygulamada kesişimlerin olasılıkları bilinmez; yalnızca tek tek olasılıklar elimizdedir. Bu durumda birleşimin olasılığını tam hesaplayamayız ama yukarıdan sınırlayabiliriz. Toplama kuralının $P(A \cup B) \leq P(A) + P(B)$ sonucunu sonlu sayıda olaya genişletelim.

Teorem 3.4 (Boole Eşitsizliği). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı ve $A_1, A_2, \dots, A_n \in \mathcal{U}$ ayrık olmaları gerekmeyen olaylar olsun. O zaman

$$P\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n P(A_i).$$

İspat.

Birleşimi ayrık parçalara ayıracağız. $B_1 = A_1$ ve $i = 2, \dots, n$ için

$$B_i = A_1^c \cap A_2^c \cap \dots \cap A_{i-1}^c \cap A_i$$

olsun; B_i , “ $A_1, \dots, A_{i-1}$ gerçekleşmez, A_i gerçekleşir” olayıdır. Dört şeyi gösterelim.

(i) Her $B_i \in \mathcal{U}$ 'dur: $\mathcal{U}$ tümleyen ve sonlu kesişim altında kapalıdır.

(ii) $B_i \subset A_i$ 'dir: B_i, A_i ile başka kümelerin kesişimidir.

(iii) $B_1, \dots, B_n$ ikişer ikişer ayrıktır: $i < j$ olsun. $B_j \subset A_i^c$ (çünkü $i \leq j-1$) ve $B_i \subset A_i$ olduğundan $B_i \cap B_j \subset A_i \cap A_i^c = \emptyset$ 'dir.

(iv) $\bigcup_{i=1}^n B_i = \bigcup_{i=1}^n A_i$ 'dir: (ii) gereği soldaki küme sağdakinin içindedir. Tersine, $\omega \in \bigcup A_i$ olsun ve $i_0, \omega \in A_{i_0}$ olan **en küçük** indis olsun. O zaman $i < i_0$ için $\omega \notin A_i$, yani $\omega \in A_i^c$ 'dir; dolayısıyla $\omega \in B_{i_0}$ 'dir.

Şimdi [Teorem 3.1](#) (b) ve (d) ile

$$P\left(\bigcup_{i=1}^n A_i\right) = P\left(\bigcup_{i=1}^n B_i\right) = \sum_{i=1}^n P(B_i) \leq \sum_{i=1}^n P(A_i).$$

■

İspattaki B_i kümelerine birleşimin **ayrıklaştırılması** denir; bu yapıyı olasılığın sürekliliğini ispatlarken yeniden kullanacağız. Boole eşitsizliği tümleyenlere uygulandığında kesişimlerin olasılığı için bir alt sınır verir.

Örnek 3.2 (Kesişimin Olasılığı için Alt Sınır). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı ve $A, B \in \mathcal{U}$ olsun.

$$P(A \cap B) \geq 1 - P(A^c) - P(B^c)$$

olduğunu gösteriniz. Bunu kullanarak, her biri 0,9 olasılıkla çalışan iki parçanın ikisinin birden çalışma olasılığının en az 0,8 olduğunu gösteriniz.

Çözüm.

De Morgan kuralı ile $(A \cap B)^c = A^c \cup B^c$ 'dir. Tümleyen kuralı ve toplama kuralı ile

$$P(A \cap B) = 1 - P((A \cap B)^c) = 1 - P(A^c \cup B^c) = 1 - P(A^c) - P(B^c) + P(A^c \cap B^c).$$

(A1) gereği $P(A^c \cap B^c) \geq 0$ olduğundan

$$P(A \cap B) \geq 1 - P(A^c) - P(B^c)$$

elde edilir. Aynı sonuca, son adımda toplama kuralı yerine doğrudan Boole eşitsizliği $P(A^c \cup B^c) \leq P(A^c) + P(B^c)$ uygulanarak da varılır. Tümleyen kuralıyla eşitsizlik

$$P(A \cap B) \geq P(A) + P(B) - 1$$

biçiminde de yazılabilir.

Uygulama: A birinci parçanın, B ikinci parçanın çalışması olayı olsun. $P(A) = P(B) = 0,9$ ise $P(A^c) = P(B^c) = 0,1$ ve

$$P(A \cap B) \geq 1 - 0,1 - 0,1 = 0,8$$

olur. Bu alt sınır, A ile B arasındaki ilişki hakkında hiçbir bilgi kullanmadan elde edilmiştir; iki parçanın arızaları ne kadar iç içe olursa olsun geçerlidir. ■

Bu eşitsizliğin n olaya genellemesi olan $P\left(\bigcap_{i=1}^n A_i\right) \geq 1 - \sum_{i=1}^n P(A_i^c)$ eşitsizliği [Alıştırma 3.1](#)'nin (d) şıkkındadır.

3.5 Olasılığın Sürekliliği

Şimdiye kadarki her sonuç, (A3) yerine yalnızca sonlu toplamsallık istenseydi de ispatlanabilirdi. Peki sayılabilir toplamsallık ne getirir? Yanıt: olasılık ölçüsü, kümelerin limitleriyle uyumlu çalışır. Bunu söyleyebilmek için önce bir olay dizisinin limitinin ne demek olduğunu, en basit durum olan monoton diziler için belirleyelim.

Tanım 3.3 (Monoton Olay Dizisi ve Limiti). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı ve $(A_n)_{n \geq 1}$, $\mathcal{U}$ 'da bir olay dizisi olsun.

- $A_1 \subset A_2 \subset \dots \subset A_n \subset \dots$ ise diziye **artan** denir ve limiti $\lim_{n \rightarrow \infty} A_n = \bigcup_{n=1}^{\infty} A_n$ olarak tanımlanır.
- $A_1 \supset A_2 \supset \dots \supset A_n \supset \dots$ ise diziye **azalan** denir ve limiti $\lim_{n \rightarrow \infty} A_n = \bigcap_{n=1}^{\infty} A_n$ olarak tanımlanır.

$\mathcal{U}$ bir σ -cebiri olduğundan her iki limit de $\mathcal{U}$ 'dadır.

Bu tanım doğal bir "limit" kavramıdır: artan bir dizide $\omega \in \bigcup A_n$ olması, ω 'nın bir yerden sonra bütün A_n 'lerde bulunması demektir; azalan bir dizide $\omega \in \bigcap A_n$ olması, ω 'nın hiçbir zaman dizinin dışına çıkmaması demektir. Örneğin $\Omega = \mathbb{R}$ üzerinde $A_n = (-\infty, a - \frac{1}{n}]$ artan bir dizidir ve limiti $(-\infty, a)$ 'dir; $A_n = (-\infty, a + \frac{1}{n}]$ azalan bir dizidir ve limiti $(-\infty, a]$ 'dir.

Gerçek değişkenli bir fonksiyonda "limit ile fonksiyonun yer değiştirebilmesi" sürekliliğin ta kendisidir. Aşağıdaki teorem olasılık ölçüsünün bu anlamda sürekli olduğunu söyler: monoton bir olay dizisinin limitinin olasılığı, olasılıkların limitidir.

Teorem 3.5 (Olasılığın Sürekliliği). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı ve $(A_n)_{n \geq 1}$, $\mathcal{U}$ 'da bir olay dizisi olsun.

a) (A_n) artan ise $\lim_{n \rightarrow \infty} P(A_n) = P\left(\bigcup_{n=1}^{\infty} A_n\right) = P(\lim_{n \rightarrow \infty} A_n)$.

b) (A_n) azalan ise $\lim_{n \rightarrow \infty} P(A_n) = P\left(\bigcap_{n=1}^{\infty} A_n\right) = P(\lim_{n \rightarrow \infty} A_n)$.

Her iki durumda da soldaki limit vardır.

İspat.

a) (A_n) artan olsun. Birleşimi ayrık "halkalara" ayıralım: $B_1 = A_1$ ve $n \geq 2$ için

$$B_n = A_n \setminus A_{n-1} = A_n \cap A_{n-1}^c$$

olsun. Dizi artan olduğundan $A_1^c \cap \dots \cap A_{n-1}^c = A_{n-1}^c$ 'dir; yani B_n 'ler, Boole eşitsizliğinin ispatındaki ayrıklaştırmanın bu özel durumdaki biçimidir. Şunları gösterelim.

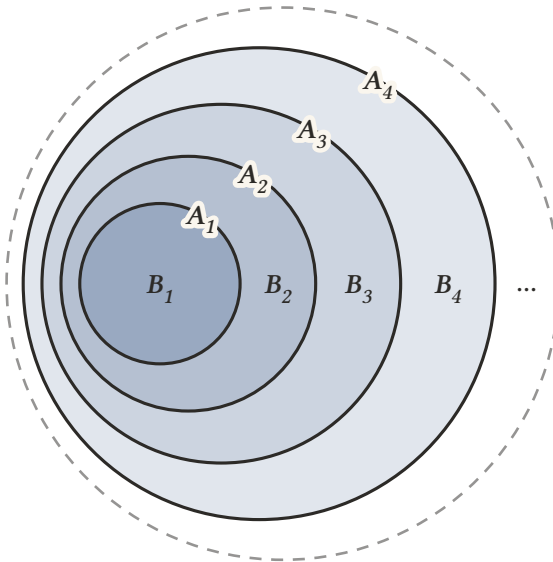
(i) B_n 'ler ikiye ikiye ayrıktır: $m < n$ ise $B_m \subset A_m \subset A_{n-1}$ ve $B_n \subset A_{n-1}^c$ olduğundan $B_m \cap B_n = \emptyset$ 'dir.

(ii) Her n için $\bigcup_{i=1}^n B_i = A_n$ 'dir: $n = 1$ için açıktır. $\bigcup_{i=1}^{n-1} B_i = A_{n-1}$ ise

$$\bigcup_{i=1}^n B_i = A_{n-1} \cup (A_n \setminus A_{n-1}) = A_n$$

olur, çünkü $A_{n-1} \subset A_n$ 'dir.

(iii) $\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$ 'dir: $B_n \subset A_n$ olduğundan soldaki sağdakinin içindedir; (ii) gereği her A_n soldakinin içindedir.



$$B_1 = A_1, B_n = A_n \cap A_{n-1}^c$$

halkalar ikiye ikiye ayrık

$$B_1, \dots, B_n \text{ 'nin birleşimi} = A_n$$

lim A_n = bütün A_n 'lerin birleşimi

$$P(A_n) = P(B_1) + \dots + P(B_n)$$

$$P(B_i) = P(A_i) - P(A_{i-1})$$

teleskopik toplam: $P(A_n)$ 'ye sadeleşir

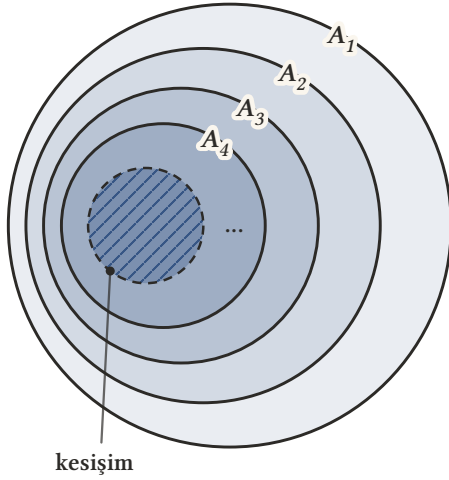
Şekil 3.3: Artan bir olay dizisi $A_1 \subset A_2 \subset A_3 \subset A_4 \subset \dots$ ve onu oluşturan halkalar. En içteki bölge $B_1 = A_1$, sonraki halkalar $B_2 = A_2 \setminus A_1$, $B_3 = A_3 \setminus A_2$, $B_4 = A_4 \setminus A_3$ 'tür. Halkalar ikiye ikiye ayrıktır ve ilk n halkanın birleşimi A_n 'yi verir; bütün halkaların birleşimi ise dizinin limiti $\cup A_n$ 'dir. Bu yüzden $P(A_n) = P(B_1) + \dots + P(B_n)$ olur ve $n \rightarrow \infty$ alınınca süreklilik teoremi çıkar.

Şimdi sırasıyla (iii), (A3), serinin tanımı, sonlu toplamsallık ve (ii) uygulanırsa

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = P\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{i=1}^{\infty} P(B_i) = \lim_{n \rightarrow \infty} \sum_{i=1}^n P(B_i) = \lim_{n \rightarrow \infty} P\left(\bigcup_{i=1}^n B_i\right) = \lim_{n \rightarrow \infty} P(A_n)$$

elde edilir. Buradaki üçüncü eşitlik serinin tanımıdır: bir serinin toplamı, kısmi toplamlar dizisinin limitidir. Bu zincir aynı zamanda $\lim P(A_n)$ limitinin var olduğunu da gösterir; nitekim $P(B_i) = P(A_i) - P(A_{i-1})$ (fark kuralı) olduğundan kısmi toplam

$$\sum_{i=1}^n P(B_i) = P(A_1) + \sum_{i=2}^n (P(A_i) - P(A_{i-1})) = P(A_n)$$



A_n azalıyor: her A_{n+1} , A_n 'nin içinde

daireler küçülür, tonlar koyulaşır

kesişim = $\lim A_n$ = bütün A_n 'lerin ortak kısmı (taralı bölge)

$P(A_n) \rightarrow P(\text{kesişim})$

tümleyenleri artan bir dizidir: A_n^c büyür

$1 - P(A_n) = P(A_n^c) \rightarrow 1 - P(\text{kesişim})$

(a) şıkkı tümleyenlere uygulanır

Şekil 3.4: Azalan bir olay dizisi: A_1, A_2 'yi, o da A_3 'ü kapsar ve dizi böyle sürer. Daireler küçüldükçe tonlar koyulaşır; ortadaki taralı bölge bütün A_n 'lerin ortak kısmı, yani dizinin limiti $\cap A_n$ 'dir. Süreklilik teoreminin (b) şıkkı $P(A_n)$ 'nin bu limitin olasılığına yaklaştığını söyler. İspat için tümleyenlere geçilir: A_n^c artan bir dizidir, (a) şıkkı ona uygulanır ve $P(A_n) = 1 - P(A_n^c)$ eşitliği sonucu verir.

biçiminde açılır (teleskopik toplam) ve $(P(A_n))$ dizisi monotonluk gereği arttığından ve Teorem 3.1 (e) gereği 1 ile sınırlı olduğundan zaten yakınsaktır.

b) (A_n) azalan olsun. Tümleyenlere geçelim: $A_n \supset A_{n+1}$ ise $A_n^c \subset A_{n+1}^c$ 'dir, yani (A_n^c) artan bir dizidir. De Morgan kuralı ile

$$\bigcup_{n=1}^{\infty} A_n^c = \left(\bigcap_{n=1}^{\infty} A_n \right)^c.$$

Şimdi (a) şıkkını (A_n^c) dizisine uygulayalım ve iki yanda tümleyen kuralını kullanalım:

$$\lim_{n \rightarrow \infty} (1 - P(A_n)) = \lim_{n \rightarrow \infty} P(A_n^c) = P\left(\bigcup_{n=1}^{\infty} A_n^c\right) = 1 - P\left(\bigcap_{n=1}^{\infty} A_n\right).$$

Soldaki limit var olduğundan $\lim P(A_n)$ de vardır ve

$$\lim_{n \rightarrow \infty} P(A_n) = P\left(\bigcap_{n=1}^{\infty} A_n\right)$$

olur.

■

! Azalan dizide sonluluk gereklidir

Teoremin (b) şıkkının ispatı $P(A_n^c) = 1 - P(A_n)$ eşitliğine, yani $P(\Omega) = 1$ olmasına dayanır. Toplam ölçüsü sonsuz olan ölçülerde (b) bozulabilir: $\mathbb{R}$ üzerinde uzunluk ölçüsünü ve $A_n = [n, \infty)$ azalan dizisini düşünelim. Her A_n 'nin uzunluğu sonsuzdur, ama $\bigcap A_n = \emptyset$ 'dir ve uzunluğu 0'dır. Olasılık ölçüsünde bu sorun çıkmaz, çünkü bütün değerler $[0, 1]$ içindedir. (a) şıkkı ise $P(\Omega) = 1$ koşulundan bağımsızdır: ispatın ana zincirinde yalnızca sayılabilir toplamsallık ve $P(\emptyset) = 0$ kullanılmıştır; bu yüzden uzunluk ölçüsü gibi toplam ölçüsü sonsuz olan ölçülerde de geçerlidir.

Süreklilik teoremi, (A3)'ün sonlu toplamsallığa göre fazlasını tam olarak açıklar: sonlu toplamsallık ile artan dizilerde süreklilik birlikte (A3)'e denktir (A1ıştırma 3.1, (f) şıkkı). Şimdi teoremin ilk sonucunu, Boole eşitsizliğinin sonsuz sürümünü verelim.

Sonuç 3.1 (Sayılabilir Alt Toplamsallık). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı ve $(A_n)_{n \geq 1}$, $\mathcal{U}$ 'da ayrık olmaları gerekmeyen herhangi bir olay dizisi olsun. O zaman

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} P(A_n).$$

Özel olarak, olasılığı 0 olan sayılabilir çoklukta olayın birleşiminin olasılığı 0'dır.

İspat.

Sağdaki serinin terimleri negatif olmadığından seri ya yakınsar ya da $+\infty$ 'a ıraksar; ıraksıyorsa eşitsizlik apaçık doğrudur. Yakınsadığını varsayalım ve

$$C_n = \bigcup_{i=1}^n A_i, \quad n = 1, 2, \dots$$

olsun. (C_n) artan bir olay dizisidir ve $\bigcup_{n=1}^{\infty} C_n = \bigcup_{n=1}^{\infty} A_n$ 'dir. Her n için Boole eşitsizliği (Teorem 3.4) ile

$$P(C_n) \leq \sum_{i=1}^n P(A_i) \leq \sum_{i=1}^{\infty} P(A_i)$$

olur; ikinci eşitsizlik terimlerin negatif olmamasından gelir. Olasılığın sürekliliği (Teorem 3.5 a) ile

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} P(C_n) \leq \sum_{i=1}^{\infty} P(A_i).$$

Her n için $P(A_n) = 0$ ise sağ taraf 0'dır; (A1) ile birleşimin olasılığı 0 olur.

■

Bu sonuç, sonraki bölümlerde sürekli olasılık modellerinde sık kullanılacaktır: tek bir noktanın olasılığı 0 olan bir modelde sayılabilir bir nokta kümesinin, örneğin bütün rasyonel sayıların, olasılığı da 0'dır.

3.6 A1ıştırma

A1ıştırma 3.1 (Olasılık Ölçüsü Üzerine). Bütün şıklarda $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı, adı geçen bütün olaylar ise $\mathcal{U}$ 'nun elemanlarıdır.

a) $P(A) = 0,6$ ve $P(B) = 0,5$ olsun; $P(A \cap B)$ bilinmiyor. $0,1 \leq P(A \cap B) \leq 0,5$ ve $0,6 \leq P(A \cup B) \leq 1$ olduğunu gösteriniz. Sınırların her birine ulaşabileceğini bir örnekle gösteriniz.

b) $P(A) = 0,5$, $P(B) = 0,4$, $P(C) = 0,3$, $P(A \cap B) = 0,2$, $P(A \cap C) = 0,15$, $P(B \cap C) = 0,1$ ve $P(A \cap B \cap C) = 0,05$ olsun. Üç olaydan en az birinin, hiçbirinin, tam olarak birinin ve tam olarak ikisinin gerçekleşme olasılıklarını bulunuz.

c) $A \triangle B = (A \setminus B) \cup (B \setminus A)$ simetrik farkı için $P(A \triangle B) = P(A) + P(B) - 2P(A \cap B) = P(A \cup B) - P(A \cap B)$ olduğunu gösteriniz.

d) $A_1, \dots, A_n \in \mathcal{U}$ için $P\left(\bigcap_{i=1}^n A_i\right) \geq 1 - \sum_{i=1}^n P(A_i^c)$ olduğunu gösteriniz. Her biri en az 0,99 olasılıkla çalışan beş parçanın hepsinin birden çalışma olasılığının en az 0,95 olduğunu çıkarınız.

- e) Her n için $P(A_n) = 1$ ise $P\left(\bigcap_{n=1}^{\infty} A_n\right) = 1$ olduğunu gösteriniz.
- f) $Q : \mathcal{U} \rightarrow \mathbb{R}$ fonksiyonu (A1) ve (A2)'yi sağlasın, sonlu toplamsal olsun (ayrık A, B için $Q(A \cup B) = Q(A) + Q(B)$) ve artan her (C_n) dizisi için $Q(\bigcup C_n) = \lim Q(C_n)$ olsun. Q 'nun (A3)'ü sağladığını, yani bir olasılık ölçüsü olduğunu gösteriniz.
- g) $\Omega = \mathbb{N} = \{1, 2, 3, \dots\}$ ve $\mathcal{U} = \mathcal{P}(\mathbb{N})$ olsun. Her n için $P(\{n\}) = c$ olacak biçimde (yani “her doğal sayı eşit olasılıklı” olacak biçimde) bir olasılık ölçüsünün var olmadığını gösteriniz.

Çözüm.

a) $A \cap B \subset B$ olduğundan monotonlukla $P(A \cap B) \leq P(B) = 0,5$ 'tir. Alt sınır için Örnek 3.2'ndeki eşitsizliği kullanalım: $P(A \cap B) \geq P(A) + P(B) - 1 = 0,1$. Birleşim için $A \subset A \cup B$ olduğundan $P(A \cup B) \geq 0,6$ 'dır; Teorem 3.1 (e) ile $P(A \cup B) \leq 1$. (Toplama kuralı ile $P(A \cup B) = 1,1 - P(A \cap B)$ olduğundan bu iki sınır çifti birbirine denktir.)

Ulaşılabirlik: $\Omega = \{1, 2, \dots, 10\}$ ve her sonucun ağırlığı $\frac{1}{10}$ olsun (bölüm başındaki nota göre bu bir olasılık ölçüsü verir). $A = \{1, \dots, 6\}$, $B = \{1, \dots, 5\}$ alınrsa $B \subset A$ olur, $P(A \cap B) = 0,5$ ve $P(A \cup B) = 0,6$ 'dır. $A = \{1, \dots, 6\}$, $B = \{6, \dots, 10\}$ alınrsa $A \cap B = \{6\}$ olur, $P(A \cap B) = 0,1$ ve $P(A \cup B) = 1$ 'dir.

b) İçerme-dışarma (Teorem 3.3, $n = 3$):

$$P(A \cup B \cup C) = 0,5 + 0,4 + 0,3 - 0,2 - 0,15 - 0,1 + 0,05 = 0,8.$$

Hiçbiri gerçekleşmez: $P(A^c \cap B^c \cap C^c) = 1 - P(A \cup B \cup C) = 0,2$.

Öteki iki soru için Ω 'yı üç olayın belirlediği ayrık parçalara bölelim. “Yalnız A ” olayı $A \setminus (B \cup C)$ 'dir; fark kuralı ve toplama kuralı ile

$$P(A \setminus (B \cup C)) = P(A) - P(A \cap (B \cup C)) = P(A) - [P(A \cap B) + P(A \cap C) - P(A \cap B \cap C)] = 0,5 - 0,2 - 0,15 + 0,05 = 0,2.$$

Aynı biçimde “yalnız B ” için $0,4 - 0,2 - 0,1 + 0,05 = 0,15$ ve “yalnız C ” için $0,3 - 0,15 - 0,1 + 0,05 = 0,1$ bulunur. Tam olarak biri: $0,2 + 0,15 + 0,1 = 0,45$.

Tam olarak ikisi: “ A ve B ama C değil” olayı $(A \cap B) \setminus (A \cap B \cap C)$ 'dir ve olasılığı $0,2 - 0,05 = 0,15$ 'tir; benzer biçimde A, C için $0,15 - 0,05 = 0,1$ ve B, C için $0,1 - 0,05 = 0,05$. Toplam $0,15 + 0,1 + 0,05 = 0,3$.

Kontrol: tam olarak biri, tam olarak ikisi ve üçü birden olayları ayrıktır ve birleşimleri $A \cup B \cup C$ 'dir: $0,45 + 0,3 + 0,05 = 0,8$. Tutarlıdır.

c) $A \setminus B$ ile $B \setminus A$ ayrıktır; $A \cap B \subset A$ ve $A \cap B \subset B$ olduğundan fark kuralı ile

$$P(A \triangle B) = P(A \setminus B) + P(B \setminus A) = [P(A) - P(A \cap B)] + [P(B) - P(A \cap B)] = P(A) + P(B) - 2P(A \cap B).$$

Toplama kuralına göre $P(A) + P(B) = P(A \cup B) + P(A \cap B)$ olduğundan bu ifade $P(A \cup B) - P(A \cap B)$ 'ye eşittir.

d) De Morgan kuralı ile $\left(\bigcap_{i=1}^n A_i\right)^c = \bigcup_{i=1}^n A_i^c$ 'dir. Tümlen kuralı ve Boole eşitsizliği ile

$$P\left(\bigcap_{i=1}^n A_i\right) = 1 - P\left(\bigcup_{i=1}^n A_i^c\right) \geq 1 - \sum_{i=1}^n P(A_i^c).$$

A_i , i 'inci parçanın çalışması olayı olsun. $P(A_i) \geq 0,99$ ise $P(A_i^c) \leq 0,01$ ve $n = 5$ için

$$P(A_1 \cap \dots \cap A_5) \geq 1 - 5 \cdot 0,01 = 0,95.$$

e) $P(A_n) = 1$ ise tümlen kuralı ile $P(A_n^c) = 0$ 'dır. Sonuç 3.1 gereği $P(\bigcup A_n^c) \leq \sum P(A_n^c) = 0$, yani $P(\bigcup A_n^c) = 0$. De Morgan ve tümlen kuralı ile

$$P\left(\bigcap_{n=1}^{\infty} A_n\right) = 1 - P\left(\bigcup_{n=1}^{\infty} A_n^c\right) = 1.$$

f) (A_n) , $\mathcal{U}$ 'da ikiye ikiye ayrık bir dizi olsun. $C_n = \bigcup_{i=1}^n A_i$ diyelim. (C_n) artan bir dizidir ve $\bigcup C_n = \bigcup A_n$ 'dir. Sonlu toplamsallık, tümevarımla n ayrık kümeye genişler: $C_n = C_{n-1} \cup A_n$ ayrık bir birleşimdir, dolayısıyla $Q(C_n) = Q(C_{n-1}) + Q(A_n)$ ve buradan $Q(C_n) = \sum_{i=1}^n Q(A_i)$. Süreklilik varsayımı ile

$$Q\left(\bigcup_{n=1}^{\infty} A_n\right) = Q\left(\bigcup_{n=1}^{\infty} C_n\right) = \lim_{n \rightarrow \infty} Q(C_n) = \lim_{n \rightarrow \infty} \sum_{i=1}^n Q(A_i) = \sum_{i=1}^{\infty} Q(A_i).$$

Bu (A3)'tür. Demek ki sonlu toplamsallık ile artan dizilerde süreklilik birlikte, sayılabilir toplamsallığa denktir; ters yön [Teorem 3.1](#) (b) ve [Teorem 3.5](#) (a)'dır.

g) Böyle bir P olduğunu varsayalım. $\{1\}, \{2\}, \{3\}, \dots$ tek elemanlı kümeleri ikiye ikiye ayrıştır ve birleřimleri $\mathbb{N} = \Omega$ 'dır. (A2) ve (A3) gereęi

$$1 = P(\mathbb{N}) = \sum_{n=1}^{\infty} P(\{n\}) = \sum_{n=1}^{\infty} c.$$

$c = 0$ ise saę taraf 0'dır; $c > 0$ ise seri ıraksar. Her iki durumda da eřitlik saęlanamaz; çeliřki. Demek ki "rastgele bir doęal sayı seçmek" ifadesinin bütün doęal sayıları eřit olasılıklı yapan bir modeli yoktur. Aynı hesap, $\mathcal{P}(\mathbb{N})$ üzerindeki her olasılık ölçüsünün $p_n = P(\{n\})$ sayıları ile belirlendięini ve bu sayıların $\sum_{n=1}^{\infty} p_n = 1$ kořulunu saęlaması gerektięini gösterir.

■

Bu bölümde olasılığı üç aksiyomla tanımladık ve hesaplarda kullanılan bütün temel kuralları bu aksiyomlardan türettik. Sıradaki bölümde bu soyut yapıyı somut modellerle dolduracaęız: [Olasılık Uzaı Örnekleri](#) ve [Geometrik Olasılık](#).

4. Olasılık Uzayı Örnekleri ve Geometrik Olasılık

Olasılık Ölçüsü ve Olasılık Uzayı bölümünde olasılık ölçüsünü Kolmogorov aksiyomlarıyla (Tanım 3.1) tanımladık ve bir olasılık uzayının (Tanım 3.2) genel özelliklerini aksiyomlardan çıkardık. O bölümde P soyut bir nesneydi. Bu bölümde soruyu tersine çeviriyoruz: **somut bir rastgele deney verildiğinde onu modelleyen olasılık uzayı nasıl kurulur?**

Yanıt, örnek uzayın büyüklüğüne göre üç biçim alır. Ω sonlu ya da sayılabilir sonsuz ise her sonuca bir ağırlık vermek yeter; olasılık ölçüsü bu ağırlıkların toplanmasıyla elde edilir. Ağırlıkların eşit olduğu özel durum bizi Laplace olasılık uzayına ve klasik “uygun durum sayısı bölü olası durum sayısı” formülüne götürür. $\Omega = \mathbb{R}$ olduğunda ise noktalara ağırlık vermek işe yaramaz; bunun yerine olasılığın $\mathbb{R}$ üzerine nasıl yığıldığını tek bir fonksiyonla, **dağılım fonksiyonuyla** betimleriz. Dağılım fonksiyonunun üç temel özelliğini olasılığın sürekliliğinden yararlanarak ispatlayacağız; bu teorem ilerleyen bölümlerde rastgele değişkenler için sürekli kullanılacaktır.

Son kısımda uzunluk, alan ve hacimle kurulan **geometrik olasılık** uzaylarını ele alacağız. Burada “rastgele bir nokta seçmek” sözünün ne anlama geldiğini belirlemek gerekir: para-silindir örneğinde aynı deneyi iki biçimde modelleyip iki farklı yanıt alacağız ve hangisinin doğru olduğunu tartışacağız.

4.1 Sonlu Örnek Uzaylar

En basit durum Ω ’nın sonlu olmasıdır. Bu durumda σ -cebiri olarak kuvvet kümesi (Tanım 1.6) alınır: Ω ’nın her alt kümesi bir olaydır. Olasılık ölçüsü kurmak için her sonuca negatif olmayan bir ağırlık vermek ve ağırlıkların toplamını 1 yapmak yeter; aşağıdaki önerme bunu söyler.

Önerme 4.1 (Sonlu Örnek Uzayda Ağırlıklarla Olasılık Ölçüsü). $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ sonlu bir örnek uzay ve $\mathcal{U} = \mathcal{P}(\Omega)$ olsun. Her ω_i sonucuna

- (1) $p_i \geq 0, i = 1, 2, \dots, n$,
- (2) $\sum_{i=1}^n p_i = 1$

koşullarını sağlayan bir p_i sayısı karşılık getirilsin. Bu durumda

$$P : \mathcal{U} \rightarrow \mathbb{R}, \quad P(A) = \sum_{\omega_i \in A} p_i$$

fonksiyonu $\mathcal{U}$ üzerinde bir olasılık ölçüsüdür. Burada $A = \emptyset$ için toplam boştur ve 0 sayılır; yani $P(\emptyset) = 0$ ’dır. Özel olarak her i için $P(\{\omega_i\}) = p_i$ ’dir.

İspat.

Kolmogorov aksiyomlarını (Tanım 3.1) tek tek doğrulayalım.

(A1) Her $A \in \mathcal{U}$ için $P(A)$, negatif olmayan sayıların sonlu bir toplamıdır; dolayısıyla $P(A) \geq 0$ ’dır.

(A2) Ω ’nın bütün elemanları toplama girer: $P(\Omega) = \sum_{i=1}^n p_i = 1$.

(A3) $(A_k)_{k \geq 1}$, $\mathcal{U}$ içinde ikişer ikişer ayrık olayların bir dizisi ve $A = \bigcup_{k=1}^{\infty} A_k$ olsun. Bu dizide en çok n tane boş olmayan terim vardır: boş olmayan A_k ’ların her birinden birer eleman seçilirse, kümeler ayrık olduğundan seçilenler birbirinden farklıdır ve Ω ’nın yalnızca n elemanı vardır. Boş terimlerin katkısı 0 olduğundan $\sum_{k=1}^{\infty} P(A_k)$ aslında sonlu bir toplamdır.

A ’nın her ω_i elemanı, kümeler ayrık olduğundan, A_k ’lardan **tam olarak birine** aittir. Öyleyse A üzerindeki toplam, A_k ’lar üzerindeki toplamlara gruplanabilir:

$$P(A) = \sum_{\omega_i \in A} p_i = \sum_{k=1}^{\infty} \sum_{\omega_i \in A_k} p_i = \sum_{k=1}^{\infty} P(A_k).$$

Sonlu bir toplamın terimlerini yeniden gruplamak toplamı değiştirmez; bu yüzden ortadaki eşitlik geçerlidir. Böylece (A3) sağlanır.

■

Önermenin tersi de doğrudur: sonlu Ω üzerindeki **her** olasılık ölçüsü bu biçimdedir. Gerçekten, $p_i = P(\{\omega_i\})$ alınırsa $A = \bigcup_{\omega_i \in A} \{\omega_i\}$ sonlu ayrık birleşimi ve sonlu toplamsallık (Teorem 3.1) $P(A) = \sum_{\omega_i \in A} p_i$ verir. Demek ki sonlu bir örnek uzayda olasılık ölçüsü kurmak ile ağırlık vektörü $(p_1, \dots, p_n)$ seçmek aynı şeydir.

Örnek 4.1 (Altı Sonuçlu Ağırlıklı Uzay). $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$, $\mathcal{U} = \mathcal{P}(\Omega)$ ve sonuçlara karşılık getirilen ağırlıklar sırasıyla

$$p_1 = 0, 1, \quad p_2 = 0, 2, \quad p_3 = 0, 3, \quad p_4 = 0, 2, \quad p_5 = 0, 1, \quad p_6 = 0, 1$$

olsun. $A = \{\omega_1, \omega_3, \omega_5\}$ ve $B = \{\omega_5, \omega_6\}$ olayları için $P(A)$, $P(B)$, $P(A \cap B)$, $P(A \cup B)$ ve $P(A^c)$ değerlerini bulunuz.

Çözüm.

Ağırlıkların hiçbirisi negatif değildir ve toplamları

$$0, 1 + 0, 2 + 0, 3 + 0, 2 + 0, 1 + 0, 1 = 1$$

olduğundan Önerme 4.1 uygulanır: $P(A) = \sum_{\omega_i \in A} p_i$ bir olasılık ölçüsüdür. Buna göre

$$P(A) = p_1 + p_3 + p_5 = 0, 1 + 0, 3 + 0, 1 = 0, 5,$$

$$P(B) = p_5 + p_6 = 0, 1 + 0, 1 = 0, 2,$$

$$P(A \cap B) = P(\{\omega_5\}) = p_5 = 0, 1,$$

$$P(A \cup B) = P(\{\omega_1, \omega_3, \omega_5, \omega_6\}) = 0, 1 + 0, 3 + 0, 1 + 0, 1 = 0, 6$$

bulunur. Son değer toplama kuralıyla (Teorem 3.2) da tutarlıdır: $0, 5 + 0, 2 - 0, 1 = 0, 6$. Tümleney için $A^c = \{\omega_2, \omega_4, \omega_6\}$ ve $P(A^c) = 0, 2 + 0, 2 + 0, 1 = 0, 5 = 1 - P(A)$ olur (Teorem 3.1). İleride tanımlanacak koşullu olasılık (Tanım 6.1) için de not edelim: $P(A | B) = \frac{P(A \cap B)}{P(B)} = \frac{0, 1}{0, 2} = \frac{1}{2}$.

■

Bir olasılık uzayının matematiksel olarak tutarlı olması ile bir deneyi **doğru** modellemesi ayrı şeylerdir; ağırlıkların nereden geldiği bir modelleme sorusudur.

Örnek 4.2 (Torba Modeli ve Eşit Ağırlıklar). a) Örnek 4.1'deki olasılık uzayı hangi deneyi modellemek için kullanılabilir?

b) Bir torbada altı farklı renkten **eşit sayıda** top bulunuyor; bir top çekilip rengi gözleniyor. Bu deney için bir olasılık uzayı kurunuz.

Çözüm.

a) İçinde 1 beyaz, 2 siyah, 3 mavi, 2 yeşil, 1 sarı ve 1 kırmızı olmak üzere 10 top bulunan bir torbadan bir top çekilip renginin gözlenmesi deneyini düşünelim. Sonuçlar renklerdir:

$$\Omega = \{\text{beyaz, siyah, mavi, yeşil, sarı, kırmızı}\},$$

ve $\omega_1, \dots, \omega_6$ bu renklere sırasıyla karşılık gelsin. Toplar iyice karıştırılmışsa her topun çekilme şansı aynıdır; bir rengin olasılığı, o renkteki top sayısının toplam top sayısına oranıdır:

$$p_1 = \frac{1}{10}, \quad p_2 = \frac{2}{10}, \quad p_3 = \frac{3}{10}, \quad p_4 = \frac{2}{10}, \quad p_5 = \frac{1}{10}, \quad p_6 = \frac{1}{10}.$$

Bunlar tam olarak Örnek 4.1'deki ağırlıklardır. Örneğin $A = \{\omega_1, \omega_3, \omega_5\}$, “çekilen topun beyaz, mavi ya da sarı olması” olayıdır ve olasılığı $0, 5$ 'tir.

b) Her renkten r top varsa toplam $6r$ top vardır ve her rengin olasılığı $\frac{r}{6r} = \frac{1}{6}$ olur. Öyleyse

$$\Omega = \{\omega_1, \dots, \omega_6\}, \quad \mathcal{U} = \mathcal{P}(\Omega), \quad p_1 = p_2 = \dots = p_6 = \frac{1}{6}$$

alınır ve **Önerme 4.1** gereği

$$P(A) = \sum_{\omega_i \in A} \frac{1}{6} = \frac{n(A)}{6}$$

bir olasılık ölçüsüdür; burada $n(A)$, A 'nın eleman sayısıdır. Aynı uzay, $\Omega = \{1, \dots, 6\}$ alınarak düzgün bir zarın atılması deneyini de modeller.

■

Eşit ağırlıklar durumu o kadar sık karşımıza çıkar ki ayrı bir ad taşır.

Tanım 4.1 (Laplace Olasılık Uzayı). Ω sonlu bir örnek uzay, $\mathcal{U} = \mathcal{P}(\Omega)$ ve her $\omega \in \Omega$ için

$$P(\{\omega\}) = \frac{1}{n(\Omega)}$$

olacak biçimde kurulan $(\Omega, \mathcal{U}, P)$ olasılık uzayına **Laplace olasılık uzayı** denir. Burada $n(\Omega)$ (kimi zaman $|\Omega|$ ile de gösterilir), Ω 'nın eleman sayısıdır. Bu P ölçüsüne Ω üzerindeki **düzgün dağılım** da denir.

Tanımdaki P gerçekten bir olasılık ölçüsüdür: **Önerme 4.1**'de $p_i = \frac{1}{n(\Omega)}$ alınırsa $p_i \geq 0$ ve $\sum p_i = 1$ olur. Laplace uzayında olasılık hesabı bir **sayma** işine indirgenir; bunu söyleyen formül olasılık teorisinin tarihsel başlangıç noktasıdır.

Önerme 4.2 (Laplace Formülü). $(\Omega, \mathcal{U}, P)$ bir Laplace olasılık uzayı ve $A \subset \Omega$ olsun. Bu durumda

$$P(A) = \frac{n(A)}{n(\Omega)},$$

yani $P(A)$, “ A için uygun durumların sayısı” bölü “olası durumların sayısı”dır.

İspat.

$A = \emptyset$ ise iki taraf da 0'dır. $A \neq \emptyset$ ise A , tek elemanlı olayların sonlu ve ayrık birleşimidir:

$$A = \bigcup_{\omega \in A} \{\omega\}.$$

Sonlu toplamsallık (**Teorem 3.1**) ve tanım gereği

$$P(A) = \sum_{\omega \in A} P(\{\omega\}) = \sum_{\omega \in A} \frac{1}{n(\Omega)} = \frac{n(A)}{n(\Omega)}$$

elde edilir.

■

⚠ Laplace formülü yalnızca eşit olasılıklı sonlu uzaylarda geçerlidir

İki düzgün zar atılıp toplama bakıldığında $\Omega = \{2, 3, \dots, 12\}$ alıp “toplamanın 7 olması olasılığı $\frac{1}{11}$ ’dir” demek yanlıştır; bu 11 sonuç eşit olasılıklı değildir. Doğru Laplace uzayı, 36 sıralı ikiliden oluşan $\Omega = \{(x, y) : x, y \in \{1, \dots, 6\}\}$ uzayıdır; toplamı 7 yapan $(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)$ ikilileri 6 tane olduğundan olasılık $\frac{6}{36} = \frac{1}{6}$ ’dır. Hangi sonuçların eşit olasılıklı olduğu deneyin fiziksel yapısından okunur.

4.2 Sayılabilir Örnek Uzaylar

$\Omega = \{\omega_1, \omega_2, \dots, \omega_n, \dots\}$ sayılabilir sonsuz olduğunda da aynı fikir işler; ancak artık toplamlar birer seridir. Ağırlıkların negatif olmaması burada belirleyici rol oynar: terimleri negatif olmayan bir serinin toplamı terimlerin sıralanışına bağlı değildir ve indis kümesinin sonlu alt kümeleri üzerinden alınan sonlu toplamların supremumuna eşittir. İspatta yalnızca bu gerçeği kullanacağız.

Önerme 4.3 (Sayılabilir Örnek Uzayda Ağırlıklarla Olasılık Ölçüsü). $\Omega = \{\omega_1, \omega_2, \dots, \omega_n, \dots\}$ sayılabilir sonsuz bir örnek uzay ve $\mathcal{U} = \mathcal{P}(\Omega)$ olsun. Her ω_i sonucuna

- (1) $p_i \geq 0, i = 1, 2, \dots,$
- (2) $\sum_{i=1}^{\infty} p_i = 1$

koşullarını sağlayan bir p_i sayısı karşılık getirilsin. Bu durumda

$$P : \mathcal{U} \rightarrow \mathbb{R}, \quad P(A) = \sum_{\omega_i \in A} p_i$$

fonksiyonu $\mathcal{U}$ üzerinde bir olasılık ölçüsüdür. (A sonsuz ise toplam, $\omega_i \in A$ olan p_i 'lerin artan indis sırasıyla oluşturduğu seridir; boş toplam 0 sayılır.)

İspat.

$B \subset \Omega$ için $S(B) = \sum_{\omega_i \in B} p_i$ yazalım. Terimler negatif olmadığından $S(B)$, B 'nin sonlu alt kümeleri F üzerinden alınan $\sum_{\omega_i \in F} p_i$ toplamlarının supremumudur ve $S(B) \leq S(\Omega) = 1$ olur. Özel olarak her A için $P(A) = S(A)$ iyi tanımlı bir gerçel sayıdır ve $0 \leq P(A) \leq 1$ 'dir.

(A1) Yukarıdaki gözlem gereği $P(A) \geq 0$ 'dır.

(A2) $P(\Omega) = \sum_{i=1}^{\infty} p_i = 1$.

(A3) $(A_k)_{k \geq 1}$, $\mathcal{U}$ içinde ikiye ikiye ayrık olayların bir dizisi ve $A = \bigcup_{k=1}^{\infty} A_k$ olsun. $P(A) = \sum_{k=1}^{\infty} P(A_k)$ eşitliğini iki eşitsizlikle göstereceğiz.

Birinci eşitsizlik: $P(A) \leq \sum_k P(A_k)$. $F \subset A$ sonlu olsun. F 'nin her elemanı A_k 'lerden tam olarak birine aittir ve F sonlu olduğundan F yalnızca sonlu sayıda A_k ile kesişir; diyelim $k \leq N$ olanlarla. O zaman

$$\sum_{\omega_i \in F} p_i = \sum_{k=1}^N \sum_{\omega_i \in F \cap A_k} p_i \leq \sum_{k=1}^N P(A_k) \leq \sum_{k=1}^{\infty} P(A_k).$$

Sol taraf F üzerinde supremum alınarak $P(A)$ 'ya çıkar; öyleyse $P(A) \leq \sum_k P(A_k)$.

İkinci eşitsizlik: $\sum_k P(A_k) \leq P(A)$. $N \in \mathbb{N}$ ve $\varepsilon > 0$ verilsin. Supremum tanımı gereği her $k \leq N$ için

$$\sum_{\omega_i \in F_k} p_i \geq P(A_k) - \frac{\varepsilon}{N}$$

olacak biçimde sonlu bir $F_k \subset A_k$ vardır. $F = F_1 \cup \dots \cup F_N$ kümesi A 'nın sonlu bir alt kümesidir ve parçaları ayrıktır. Dolayısıyla

$$P(A) \geq \sum_{\omega_i \in F} p_i = \sum_{k=1}^N \sum_{\omega_i \in F_k} p_i \geq \sum_{k=1}^N P(A_k) - \varepsilon.$$

Önce $\varepsilon \rightarrow 0$, sonra $N \rightarrow \infty$ alınırsa $P(A) \geq \sum_{k=1}^{\infty} P(A_k)$ bulunur.

İki eşitsizlik birlikte (A3)'ü verir.

■

Sonlu durumdaki gibi burada da tersi doğrudur: sayılabilir Ω üzerindeki her olasılık ölçüsü, $p_i = P(\{\omega_i\})$ ağırlıklarıyla bu biçimde yazılır; çünkü $A = \bigcup_{\omega_i \in A} \{\omega_i\}$ sayılabilir ayrık bir birleşimdir ve (A3) doğrudan istenen eşitliği verir.

Örnek 4.3 (Ağırlıkları Geometrik Dizi Olan Uzay). $\Omega = \{\omega_1, \omega_2, \omega_3, \dots\}$, $\mathcal{U} = \mathcal{P}(\Omega)$ ve ağırlıklar

$$p_1 = \frac{1}{2}, \quad p_2 = \frac{1}{2^2}, \quad p_3 = \frac{1}{2^3}, \quad \dots, \quad p_i = \frac{1}{2^i}, \quad \dots$$

olsun. $A = \{\omega_1, \omega_3, \omega_5, \dots\}$ ve $B = \{\omega_5, \omega_6\}$ olayları için $P(A)$, $P(B)$, $P(A \cap B)$, $P(A \cup B)$ ve $P(A^c)$ değerlerini bulunuz. Bu olasılık uzayı hangi deneyi modeller?

Çözüm.

Ağırlıklar pozitif ve geometrik seri formülüyle

$$\sum_{i=1}^{\infty} \frac{1}{2^i} = \frac{1/2}{1 - 1/2} = 1$$

olduğundan **Önerme 4.3** uygulanır. A , tek indisli sonuçlardan oluşur; ilk terimi $\frac{1}{2}$, oranı $\frac{1}{4}$ olan geometrik seriyi toplayarak

$$P(A) = \frac{1}{2} + \frac{1}{2^3} + \frac{1}{2^5} + \dots = \frac{1}{2} \cdot \frac{1}{1 - \frac{1}{4}} = \frac{1}{2} \cdot \frac{4}{3} = \frac{2}{3}$$

bulunur. Öteki değerler:

$$P(B) = \frac{1}{2^5} + \frac{1}{2^6} = \frac{2+1}{64} = \frac{3}{64}, \quad P(A \cap B) = P(\{\omega_5\}) = \frac{1}{2^5} = \frac{1}{32}.$$

Toplama kuralı (**Teorem 3.2**) ile

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{2}{3} + \frac{3}{64} - \frac{1}{32} = \frac{128}{192} + \frac{9}{192} - \frac{6}{192} = \frac{131}{192}.$$

Bunu doğrudan da doğrulayabiliriz: $A \cup B = A \cup \{\omega_6\}$ ayrık bir birleşimdir, dolayısıyla $P(A \cup B) = \frac{2}{3} + \frac{1}{64} = \frac{128+3}{192} = \frac{131}{192}$. Tümleyen $A^c = \{\omega_2, \omega_4, \omega_6, \dots\}$ çift indisli sonuçlardan oluşur:

$$P(A^c) = \frac{1}{2^2} + \frac{1}{2^4} + \dots = \frac{1/4}{1 - 1/4} = \frac{1}{3} = 1 - P(A).$$

İleride tanımlanacak koşullu olasılık (**Tanım 6.1**) için de not edelim:

$$P(A | B) = \frac{P(A \cap B)}{P(B)} = \frac{1/32}{3/64} = \frac{64}{96} = \frac{2}{3}.$$

Modellenen deney. Düzgün bir para, ilk kez tura gelinceye kadar atılıyor ve üste gelen yüzler gözleniyor. Yazıyı Y , turayı T ile gösterirsek sonuçlar

$$\Omega = \{T, YT, YYT, YYYT, \dots\}$$

kümesini oluşturur; ω_i , “ilk turaanın i -inci atışta gelmesi” sonucudur. Bu sonuca $\frac{1}{2^i}$ ağırlığı vermek doğaldır: i atışlık 2^i tane Y - T dizisi düzgün para için eşit olasılıklıdır ve içlerinden yalnızca biri ($YY \dots YT$) ω_i 'ye karşılık gelir; Laplace formülü (**Önerme 4.2**) $\frac{1}{2^i}$ verir. Ağırlıkların toplamının 1 çıkması, paranın hiç tura gelmemesi olasılığının 0 olduğunu söyler; bu yüzden $YYY \dots$ sonsuz dizisini Ω 'ya almaya gerek yoktur. A , “turaanın tek sayılı bir atışta gelmesi”, B ise “ilk turaanın beşinci ya da altıncı atışta gelmesi” olayıdır.

■

4.3 Gerçel Sayılar Üzerinde Olasılık Ölçüleri

Ölçüm sonucu bir gerçel sayı olan deneylerde (bir ampulün ömrü gibi) örnek uzay $\Omega = \mathbb{R}$ ya da $\mathbb{R}$ 'nin bir alt kümesidir. Bu durumda σ -cebiri olarak kuvvet kümesi yerine Borel cebri $\mathcal{B}(\mathbb{R})$ (**Tanım 2.4**) alınır; Borel cebri her türden aralığı içerir ve ilgilendiğimiz olaylar zaten aralık biçimindedir.

Ağırlık yöntemi burada çöker: $\mathbb{R}$ sayılamaz sonsuzdur; bütün noktalara aynı pozitif ağırlığı vermek istesek, sayılabilir bir nokta kümesinin bile olasılığı 1'i aşardı. Bunun yerine olasılığın $\mathbb{R}$ üzerine nasıl yığıldığını tek bir fonksiyonla betimleyeceğiz.

Tanım 4.2 (Borel Olasılık Ölçüsü (Olasılık Dağılımı)). $\Omega = \mathbb{R}$ ve $\mathcal{U} = \mathcal{B}(\mathbb{R})$ olsun. Borel cebri üzerinde tanımlı bir

$$P : \mathcal{B}(\mathbb{R}) \rightarrow \mathbb{R}, \quad A \mapsto P(A)$$

olasılık ölçüsüne $\mathbb{R}$ üzerinde bir **Borel olasılık ölçüsü** ya da kısaca $\mathbb{R}$ üzerinde bir **olasılık dağılımı** denir. $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P)$ bir olasılık uzayıdır ve $P(\mathbb{R}) = 1$ 'dir: bir birim olasılık gerçel sayılar üzerine dağıtılmıştır.

Örnek uzay olarak bir Borel kümesi $\Omega \subset \mathbb{R}$ alınırsa $\mathcal{U} = \{A \in \mathcal{B}(\mathbb{R}) : A \subset \Omega\}$ alınır; her $B \in \mathcal{B}(\mathbb{R})$ için $\tilde{P}(B) = P(B \cap \Omega)$ konularak bu ölçü, Ω^c kümesine 0 olasılığı veren, $\mathbb{R}$ üzerinde bir Borel olasılık ölçüsü gibi görülebilir.

Tanım 4.3 (Olasılık Ölçüsünün Dağılım Fonksiyonu). $P, \mathcal{B}(\mathbb{R})$ üzerinde bir olasılık ölçüsü olsun.

$$F : \mathbb{R} \rightarrow [0, 1], \quad F(x) = P((-\infty, x])$$

fonksiyonuna P olasılık ölçüsüne karşılık gelen **dağılım fonksiyonu** (birikimli dağılım fonksiyonu) denir.

Tanım anlamlıdır: $(-\infty, x]$ kapalı bir yarı doğru olduğundan Borel cebirindedir (Tanım 2.4). $F(x)$, olasılığın x noktasına kadar birikmiş kısmıdır. Peki P 'den gelen böyle bir F hangi özellikleri taşımak zorundadır? Aşağıdaki teorem üç özellik verir; ispatı olasılığın sürekliliğine (Teorem 3.5) dayanır ve ilerleyen bölümlerde rastgele değişkenlerin dağılım fonksiyonları için de aynen kullanılacaktır.

Teorem 4.1 (Dağılım Fonksiyonunun Özellikleri). $P, \mathcal{B}(\mathbb{R})$ üzerinde bir olasılık ölçüsü ve $F(x) = P((-\infty, x])$ ona karşılık gelen dağılım fonksiyonu olsun. Bu durumda

- a) F azalmayandır: $x_1 < x_2 \Rightarrow F(x_1) \leq F(x_2)$;
- b) F sağdan süreklidir: her $x \in \mathbb{R}$ için $\lim_{h \rightarrow 0^+} F(x+h) = F(x)$;
- c) $\lim_{x \rightarrow -\infty} F(x) = 0$ ve $\lim_{x \rightarrow +\infty} F(x) = 1$.

İspat.

a) $x_1 < x_2$ ise $(-\infty, x_1] \subset (-\infty, x_2]$ 'dir. Olasılık ölçüsünün monotonluğu (Teorem 3.1) gereği $F(x_1) = P((-\infty, x_1]) \leq P((-\infty, x_2]) = F(x_2)$ olur.

b) $x \in \mathbb{R}$ sabit olsun. $n \in \mathbb{N}$ için $A_n = (-\infty, x + \frac{1}{n}]$ olayını tanımlayalım. $\frac{1}{n}$ azaldığından $A_1 \supset A_2 \supset A_3 \supset \dots$ azalan bir dizidir. Kesişimini bulalım: $y \leq x$ ise her n için $y \leq x + \frac{1}{n}$, yani $y \in A_n$ 'dir; tersine y bütün A_n 'lerde ise her n için $y \leq x + \frac{1}{n}$ 'dir ve $n \rightarrow \infty$ alınarak $y \leq x$ bulunur. Demek ki

$$\bigcap_{n=1}^{\infty} A_n = (-\infty, x].$$

Azalan diziler için olasılığın sürekliliği (Teorem 3.5) gereği

$$\lim_{n \rightarrow \infty} F(x + \frac{1}{n}) = \lim_{n \rightarrow \infty} P(A_n) = P\left(\bigcap_{n=1}^{\infty} A_n\right) = F(x).$$

Bu, limitin yalnızca $h = \frac{1}{n}$ dizisi boyunca alındığını söyler; keyfi $h \rightarrow 0^+$ için (a) şıkkını kullanacağız. $\varepsilon > 0$ verilsin. Yukarıdaki limit gereği $F(x + \frac{1}{N}) < F(x) + \varepsilon$ olacak biçimde bir N vardır. $0 < h < \frac{1}{N}$ ise $x < x+h < x + \frac{1}{N}$ olduğundan (a) gereği

$$F(x) \leq F(x+h) \leq F(x + \frac{1}{N}) < F(x) + \varepsilon,$$

yani $|F(x+h) - F(x)| < \varepsilon$ 'dur. Öyleyse $\lim_{h \rightarrow 0^+} F(x+h) = F(x)$.

c) Önce $-\infty$ 'daki limit. $B_n = (-\infty, -n]$ olsun; $B_1 \supset B_2 \supset \dots$ azalandır ve kesişimi boştur: her n için $y \leq -n$ olan bir y gerçel sayısı yoktur. Azalan diziler için süreklilik ve $P(\emptyset) = 0$ (Teorem 3.1) gereği

$$\lim_{n \rightarrow \infty} F(-n) = \lim_{n \rightarrow \infty} P(B_n) = P(\emptyset) = 0.$$

$\varepsilon > 0$ verilsin; $F(-N) < \varepsilon$ olacak biçimde N seçelim. $x < -N$ ise (a) gereği $0 \leq F(x) \leq F(-N) < \varepsilon$ 'dur. Bu, $\lim_{x \rightarrow -\infty} F(x) = 0$ demektir.
 Şimdi $+\infty$ 'daki limit. $C_n = (-\infty, n]$ olsun; $C_1 \subset C_2 \subset \dots$ artandır ve birleşimi $\mathbb{R}$ 'dir: her y gerçel sayısı, $y \leq n$ olan bir n için C_n 'dedir. Artan diziler için süreklilik gereği

$$\lim_{n \rightarrow \infty} F(n) = \lim_{n \rightarrow \infty} P(C_n) = P(\mathbb{R}) = 1.$$

$\varepsilon > 0$ için $F(N) > 1 - \varepsilon$ olacak biçimde N seçilirse, $x > N$ olan her x için $1 \geq F(x) \geq F(N) > 1 - \varepsilon$ olur. Bu da $\lim_{x \rightarrow +\infty} F(x) = 1$ demektir.

i Neden sağdan süreklilik

F azalmayan ve sınırlı olduğundan her x noktasında soldan limiti $F(x^-) = \lim_{h \rightarrow 0^+} F(x - h)$ de vardır; ama teorem $F(x^-) = F(x)$ demez. Birazdan göreceğimiz gibi $F(x) - F(x^-)$ farkı tam olarak $P(\{x\})$ 'tir. Sağdan sürekliliğin nedeni, tanımda $(-\infty, x]$ yarı doğrusunun sağ ucunun kapalı alınmasıdır.

Teoremin tersi de doğrudur: bu üç özelliği taşıyan **her** fonksiyon bir Borel olasılık ölçüsünün dağılım fonksiyonudur ve bu ölçü tektir. Sonucun ispatı ölçü teorisinin genişleme teoremlerini gerektirdiğinden burada yalnızca ifadesini veriyoruz.

Teorem 4.2 (Borel--Stieltjes Teoremi). $F : \mathbb{R} \rightarrow [0, 1]$ fonksiyonu azalmayan ve sağdan sürekli olsun; ayrıca $\lim_{x \rightarrow -\infty} F(x) = 0$ ve $\lim_{x \rightarrow +\infty} F(x) = 1$ olsun. Bu durumda $\mathcal{B}(\mathbb{R})$ üzerinde, her $x \in \mathbb{R}$ için

$$P((-\infty, x]) = F(x)$$

olan **bir ve yalnız bir** P olasılık ölçüsü vardır. Bu P 'ye F 'nin belirlediği **Borel--Stieltjes ölçüsü** denir.

Teorem 4.1 ile **Teorem 4.2** birlikte şunu söyler: dağılım fonksiyonları ile Borel olasılık ölçüleri arasında $F \leftrightarrow P$ bire bir eşlemesi vardır. $\mathbb{R}$ üzerinde bir olasılık dağılımı vermek için F 'yi vermek yeter; yarı doğrulara aynı olasılığı veren iki ölçü bütün Borel kümelerinde aynıdır. Öyleyse F 'den yola çıkarak aralıkların ve tek noktaların olasılığını hesaplayabilmeliyiz; aşağıdaki önerme bu hesabın kurallarını toplar.

Önerme 4.4 (Aralıkların ve Tek Noktaların Olasılığı). P bir Borel olasılık ölçüsü, F onun dağılım fonksiyonu ve $F(a^-) = \lim_{h \rightarrow 0^+} F(a - h)$ olsun. $a < b$ gerçel sayıları için

- $P((a, b]) = F(b) - F(a)$,
 - $P((a, \infty)) = 1 - F(a)$,
 - $P(\{a\}) = F(a) - F(a^-)$,
 - $P((-\infty, a)) = F(a^-)$,
 - $P([a, b]) = F(b) - F(a^-)$, $P((a, b)) = F(b^-) - F(a)$, $P([a, b)) = F(b^-) - F(a^-)$
- eşitlikleri geçerlidir. Özel olarak, $P(\{a\}) = 0$ olması için gerek ve yeter koşul F 'nin a noktasında sürekli olmasıdır.

İspat.

F azalmayan ve sınırlı olduğundan $F(a^-)$ soldan limiti vardır ve $F(a^-) = \sup\{F(y) : y < a\}$ 'dir; özel olarak $F(a - \frac{1}{n}) \rightarrow F(a^-)$ 'dir.

a) $(-\infty, a] \subset (-\infty, b]$ ve $(a, b] = (-\infty, b] \setminus (-\infty, a]$ 'dir. Fark kuralı (**Teorem 3.1**) gereği $P((a, b]) = F(b) - F(a)$.

b) $(a, \infty) = ((-\infty, a])^c$ olduğundan tümleyen kuralı $P((a, \infty)) = 1 - F(a)$ verir.

c) $D_n = (a - \frac{1}{n}, a]$ olsun. $D_1 \supset D_2 \supset \dots$ azalandır ve kesişimi $\{a\}$ 'dir: bütün D_n 'lerde bulunan bir y için her n 'de $a - \frac{1}{n} < y \leq a$ olur, bu da $y = a$ demektir. Azalan diziler için süreklilik (**Teorem 3.5**) ve (a) gereği

$$P(\{a\}) = \lim_{n \rightarrow \infty} P(D_n) = \lim_{n \rightarrow \infty} \left[F(a) - \left(F(a - \frac{1}{n}) \right) \right] = F(a) - F(a^-).$$

d) $\{a\} \subset (-\infty, a]$ ve $(-\infty, a) = (-\infty, a] \setminus \{a\}$ olduğundan fark kuralı ve (c) ile $P((-\infty, a)) = F(a) - (F(a) - F(a^-)) = F(a^-)$.

e) Ayrık birleşimler ve farklar üzerinden, (a) ve (c) kullanılarak:

$$P([a, b]) = P(\{a\}) + P((a, b]) = (F(a) - F(a^-)) + (F(b) - F(a)) = F(b) - F(a^-),$$

$$P((a, b)) = P((a, b]) - P(\{b\}) = (F(b) - F(a)) - (F(b) - F(b^-)) = F(b^-) - F(a),$$

$$P([a, b)) = P(\{a\}) + P((a, b)) = (F(a) - F(a^-)) + (F(b^-) - F(a)) = F(b^-) - F(a^-).$$

Son iddia: F her noktada sağdan sürekli olduğundan (Teorem 4.1), F 'nin a 'da sürekli olması $F(a^-) = F(a)$ olmasıyla, bu da (c) gereği $P(\{a\}) = 0$ olmasıyla eşdeğerdir.

■

Önermenin (c) şıkkı grafiği okumanın anahtarıdır: F 'nin a noktasındaki **sıçraması** tam olarak a noktasının olasılığıdır; F sürekli olduğunda aralığın uçlarının açık ya da kapalı olması olasılığı değiştirmez. Şimdi iki temel örnek görelim; ilki sürekli, ikincisi basamaklı bir dağılım fonksiyonudur.

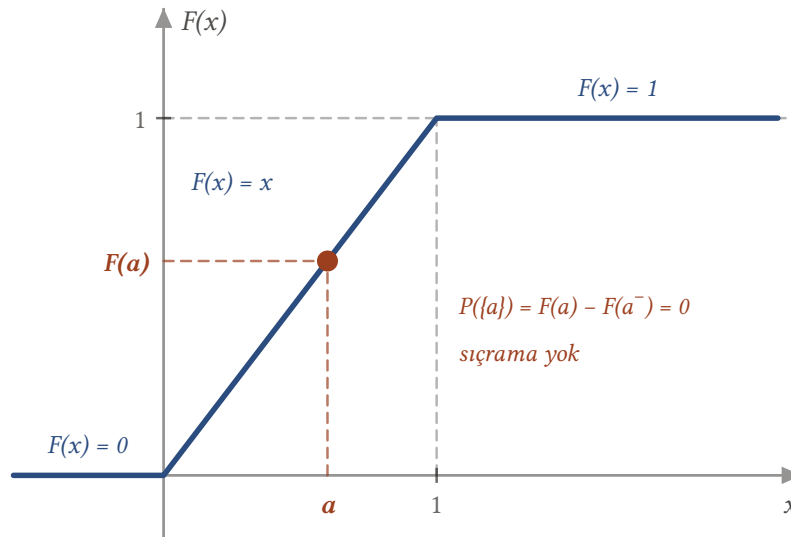
Örnek 4.4 (Birim Aralık Üzerinde Düzgün Dağılım). $F : \mathbb{R} \rightarrow [0, 1]$ fonksiyonu

$$F(x) = \begin{cases} 0, & x < 0 \\ x, & 0 \leq x < 1 \\ 1, & x \geq 1 \end{cases}$$

ile tanımlansın. F 'nin bir dağılım fonksiyonu olduğunu gösteriniz ve F 'nin belirlediği P olasılık ölçüsü için $P(\{\frac{1}{2}\})$, $P((-\frac{1}{2}, \frac{1}{2}])$, $P((\frac{1}{3}, \frac{1}{2}])$, $P([\frac{1}{3}, 3])$, $P((\frac{1}{3}, \infty))$ değerlerini ve $0 \leq a < b \leq 1$ için $P((a, b])$ 'yi hesaplayınız.

Çözüm.

Dağılım fonksiyonu özellikleri. F üç parçanın her birinde azalmayandır (sabit 0, eğimi 1 olan doğru, sabit 1) ve parçaların birleştiği noktalarda değerler uyuşur: $x \rightarrow 0^-$ iken $F(x) \rightarrow 0 = F(0)$, $x \rightarrow 1^-$ iken $F(x) \rightarrow 1 = F(1)$. Dolayısıyla F bütün $\mathbb{R}$ üzerinde azalmayan ve **sürekli** bir fonksiyondur; özel olarak sağdan sürekli. $x < 0$ için $F(x) = 0$ ve $x \geq 1$ için $F(x) = 1$ olduğundan limit koşulları da sağlanır. Teorem 4.2 gereği F , $\mathcal{B}(\mathbb{R})$ üzerinde tek bir P olasılık ölçüsü belirler.



Şekil 4.1: $[0, 1]$ üzerindeki düzgün dağılımın dağılım fonksiyonu: $x < 0$ için 0, $0 \leq x \leq 1$ için x , $x > 1$ için 1. Fonksiyon azalmayan, sağdan sürekli ve limit koşullarını sağlayan bir dağılım fonksiyonudur; üstelik her yerde sürekli. Bu yüzden a noktasında sıçrama yoktur: $P(\{a\}) = F(a) - F(a^-) = 0$ ve bir aralığın olasılığı yalnızca uzunluğudur.

Hesaplar. Önerme 4.4'ni kullanıyoruz. F sürekli olduğundan her $a \in \mathbb{R}$ için $P(\{a\}) = F(a) - F(a^-) = 0$ 'dır; özel olarak $P(\{\frac{1}{2}\}) = \frac{1}{2} - \frac{1}{2} = 0$. Ayrıca

$$P\left(\left(-2, \frac{1}{2}\right]\right) = F\left(\frac{1}{2}\right) - F(-2) = \frac{1}{2} - 0 = \frac{1}{2},$$

$$P\left(\left(\frac{1}{3}, \frac{1}{2}\right]\right) = F\left(\frac{1}{2}\right) - F\left(\frac{1}{3}\right) = \frac{1}{2} - \frac{1}{3} = \frac{1}{6},$$

$$P\left(\left[\frac{1}{3}, 3\right]\right) = F(3) - F\left(\frac{1}{3^-}\right) = 1 - \frac{1}{3} = \frac{2}{3},$$

$$P\left(\left(\frac{1}{3}, \infty\right)\right) = 1 - F\left(\frac{1}{3}\right) = 1 - \frac{1}{3} = \frac{2}{3}.$$

$0 \leq a < b \leq 1$ için $P((a, b]) = F(b) - F(a) = b - a$, yani aralığın **uzunluğudur**; uç noktaların olasılığı 0 olduğundan $[a, b]$, (a, b) ve $[a, b)$ için de olasılık $b - a$ 'dır. Öte yandan $P((-\infty, 0)) = F(0^-) = 0$ ve $P((1, \infty)) = 1 - F(1) = 0$ 'dır: bütün olasılık $[0, 1]$ aralığındadır ve orada düzgün yayılmıştır. Bu ölçüye $[0, 1]$ üzerindeki **düzgün dağılım** denir.

■

Örnek 4.5 (Basamaklı Bir Dağılım Fonksiyonu). $F : \mathbb{R} \rightarrow [0, 1]$ fonksiyonu

$$F(x) = \begin{cases} 0, & x < 1 \\ 1 - \frac{1}{n+1}, & n \leq x < n+1, \quad n = 1, 2, \dots \end{cases}$$

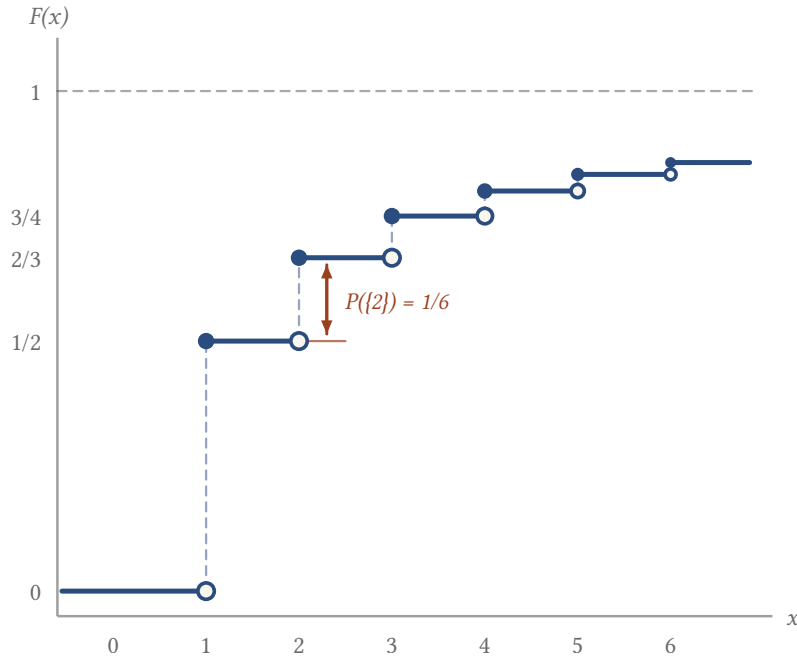
ile tanımlansın. F 'nin bir dağılım fonksiyonu olduğunu gösteriniz ve F 'nin belirlediği P için $P((-\infty, 1])$, $P(\{n\})$ ($n = 1, 2, \dots$), $P(\mathbb{Z}^+)$, $P((-\infty, 2))$ ve $P((1, 3])$ değerlerini hesaplayınız.

Çözüm.

Dağılım fonksiyonu özellikleri. F 'nin değerleri $[1, 2)$ üzerinde $\frac{1}{2}$, $[2, 3)$ üzerinde $\frac{2}{3}$, $[3, 4)$ üzerinde $\frac{3}{4}$, genel olarak $[n, n+1)$ üzerinde $1 - \frac{1}{n+1}$ 'dir. Her parçada sabit olduğundan ve $1 - \frac{1}{n+1} < 1 - \frac{1}{n+2}$ olduğundan F azalmayıp artmaktadır. Sağdan süreklilik: $x \in [n, n+1)$ ise $F, [x, n+1)$ aralığında sabittir, dolayısıyla $h \rightarrow 0^+$ iken $F(x+h) = F(x)$ 'tir; $x < 1$ için de $F, [x, 1)$ üzerinde 0'dır. Ancak F **sürekli değildir**: $n \geq 1$ tam sayısında

$$F(n^-) = 1 - \frac{1}{n}, \quad F(n) = 1 - \frac{1}{n+1}$$

olur ($n = 1$ için $F(1^-) = 0 = 1 - \frac{1}{1}$ olduğuna dikkat ediniz; formül $n = 1$ 'de de geçerlidir). Limitler: $x < 1$ için $F(x) = 0$ olduğundan $\lim_{x \rightarrow -\infty} F(x) = 0$; $x \geq n$ için $1 - \frac{1}{n+1} \leq F(x) \leq 1$ olduğundan ve $n \rightarrow \infty$ iken sol taraf 1'e gittiğinden $\lim_{x \rightarrow +\infty} F(x) = 1$. Böylece F bir dağılım fonksiyonudur ve Teorem 4.2 gereği tek bir P belirler.



Şekil 4.2: Basamaklı dağılım fonksiyonu: $x < 1$ için $F(x) = 0$, $n \leq x < n + 1$ için $F(x) = 1 - 1/(n + 1)$. Her pozitif tam sayıda bir sıçrama vardır ve sıçramanın yüksekliği o noktanın olasılığıdır: $P(\{n\}) = 1/(n(n + 1))$, örneğin $P(\{2\}) = 1/6$. Fonksiyon sağdan süreklidir (her basamağın sol ucu dolu, sağ ucu boştur); basamaklar küçülerek kesikli çizgiyle gösterilen $y = 1$ düzeyine yaklaşır.

Hesaplar. Önerme 4.4 ile:

$$P((-\infty, 1]) = F(1) = \frac{1}{2}.$$

Her $n = 1, 2, \dots$ için tek noktanın olasılığı, F 'nin n 'deki sıçramasıdır:

$$P(\{n\}) = F(n) - F(n^-) = \left(1 - \frac{1}{n+1}\right) - \left(1 - \frac{1}{n}\right) = \frac{1}{n} - \frac{1}{n+1} = \frac{1}{n(n+1)}.$$

$\mathbb{Z}^+ = \bigcup_{n=1}^{\infty} \{n\}$ sayılabilir ayrık bir birleşim olduğundan (A3) ve teleskopik toplam ile

$$P(\mathbb{Z}^+) = \sum_{n=1}^{\infty} \frac{1}{n(n+1)} = \sum_{n=1}^{\infty} \left(\frac{1}{n} - \frac{1}{n+1}\right) = \lim_{N \rightarrow \infty} \left(1 - \frac{1}{N+1}\right) = 1.$$

Demek ki $P(\mathbb{R} \setminus \mathbb{Z}^+) = 0$: bütün olasılık pozitif tam sayılar üzerinde, n noktasına $\frac{1}{n(n+1)}$ ağırlığı düşecek biçimde toplanmıştır; bu, Önerme 4.3 durumunun $\mathbb{R}$ içine gömülmüş hâlidir. Ayrıca

$$P((-\infty, 2)) = F(2^-) = 1 - \frac{1}{2} = \frac{1}{2};$$

aynı sonuç $P((-\infty, 2]) - P(\{2\}) = \frac{2}{3} - \frac{1}{6} = \frac{1}{2}$ biçiminde de bulunur. Son olarak

$$P((1, 3]) = F(3) - F(1) = \frac{3}{4} - \frac{1}{2} = \frac{1}{4} = P(\{2\}) + P(\{3\}),$$

çünkü $(1, 3]$ aralığındaki tam sayılar yalnızca 2 ve 3'tür.

■

❓ İki örnek, iki tür dağılım

İlk örnekte F süreklidir ve olasılık bir aralık üzerine yayılmıştır; ikincisinde F parça parça sabittir ve olasılığın tamamı sayılabilir bir nokta kümesinde, sıçramalar olarak toplanmıştır. Bu iki davranış ilerleyen bölümlerde **süreklili** ve **kesikli** rastgele değişkenler olarak yeniden karşımıza çıkacaktır.

4.4 Ölçü ve Geometrik Olasılık

Uzunluk, alan ve hacim olasılık ölçüsüyle aynı yapısal özellikleri taşır: negatif değildirler, boş kümede sıfırdır ve ayrık parçaların birleşiminin büyüklüğü parçaların büyüklüklerinin toplamıdır. Tek fark, toplam büyüklüğün 1 olmak zorunda olmaması, hatta sonsuz olabilmesidir. Bu ortak yapı **ölçü** kavramıdır.

Tanım 4.4 (Ölçü). $\Omega \neq \emptyset$ ve $\mathcal{U}$, Ω üzerinde bir σ -cebiri olsun. $[0, \infty] = \{t \in \mathbb{R} : t \geq 0\} \cup \{\infty\}$ olmak üzere

$$m : \mathcal{U} \rightarrow [0, \infty], \quad A \mapsto m(A)$$

fonksiyonu

- (1) her $A \in \mathcal{U}$ için $m(A) \geq 0$,
- (2) $m(\emptyset) = 0$,
- (3) (A_n) , $\mathcal{U}$ içinde ikişer ikişer ayrık kümelerin bir dizisi ise $m\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} m(A_n)$ (iki taraf da ∞ olabilir)

özelliklerini sağlıyorsa m 'ye $\mathcal{U}$ üzerinde bir **ölçü**, $(\Omega, \mathcal{U}, m)$ üçlüsüne bir **ölçü uzayı** denir. $m(\Omega) = 1$ olan bir ölçü, tam olarak bir olasılık ölçüsüdür.

Olasılık aksiyomlarında $P(\emptyset) = 0$ ayrıca istenmemiştir; $P(\Omega) = 1$ ile birlikte (A3)'ten çıkıyordu (Teorem 3.1). Genel ölçüde $m(\Omega) = \infty$ olabildiğinden bu koşul ayrıca konur; aksi hâlde her kümeye ∞ veren fonksiyon da (1) ve (3)'ü sağlardı.

Bizim için önemli üç örnek: $\mathcal{B}(\mathbb{R})$ üzerinde her aralığa uzunluğunu veren **uzunluk ölçüsü**; $\mathcal{B}(\mathbb{R}^2)$ üzerinde her dikdörtgene kenarlarının çarpımını veren **alan ölçüsü**; $\mathcal{B}(\mathbb{R}^3)$ üzerinde **hacim ölçüsü**. Bu ölçülerin bütün Borel kümelerine tek biçimde genişletilebildiği ölçü teorisinin temel sonuçlarındandır (Lebesgue ölçüsü); burada bu gerçeği kabul ediyoruz.

Bir ölçüden olasılık ölçüsü üretmenin en doğal yolu, ölçüyü toplam büyüklüğe bölerek normalleştirmektir.

Tanım 4.5 (Geometrik Olasılık). $k \in \{1, 2, 3\}$ ve m , $\mathcal{B}(\mathbb{R}^k)$ üzerinde sırasıyla uzunluk, alan ya da hacim ölçüsü olsun. $\Omega \in \mathcal{B}(\mathbb{R}^k)$ kümesi $0 < m(\Omega) < \infty$ koşulunu sağlasın ve $\mathcal{U} = \{A \in \mathcal{B}(\mathbb{R}^k) : A \subset \Omega\}$ olsun.

$$P : \mathcal{U} \rightarrow \mathbb{R}, \quad P(A) = \frac{m(A)}{m(\Omega)}$$

ile tanımlanan $(\Omega, \mathcal{U}, P)$ olasılık uzayına bir **geometrik olasılık uzayı**, P 'ye de Ω üzerindeki **düzgün dağılım** denir. A kümesinin uzunluğunu, alanını ve hacmini sırasıyla uzunluk (A), alan (A), hacim (A) ile gösterirsek açık olarak:

- (1) $k = 1$: $c < d$ olmak üzere $\Omega = [c, d]$ ve $A = [a, b] \subset \Omega$ bir aralık ise $P(A) = \frac{\text{uzunluk}(A)}{\text{uzunluk}(\Omega)} = \frac{b-a}{d-c}$;
- (2) $k = 2$: $\Omega \subset \mathbb{R}^2$ sonlu alanlı bir küme ise $P(A) = \frac{\text{alan}(A)}{\text{alan}(\Omega)}$;
- (3) $k = 3$: $\Omega \subset \mathbb{R}^3$ sonlu hacimli bir küme ise $P(A) = \frac{\text{hacim}(A)}{\text{hacim}(\Omega)}$.

Tanımdaki P gerçekten bir olasılık ölçüsüdür: Ω Borel olduğundan $\mathcal{U}$, Ω üzerinde bir σ -cebri; $m(A) \geq 0$ ve $m(\Omega) > 0$ olduğundan $P(A) \geq 0$ (A1); ayrıca $A \subset \Omega$ için $\Omega = A \cup (\Omega \setminus A)$ ayrık birleşimi ve (3) gereği $m(A) \leq m(\Omega)$, yani $P(A) \leq 1$ olur; $P(\Omega) = 1$ (A2); m 'nin (3) özelliği $m(\Omega)$ 'ya bölünerek (A3) elde edilir. Bu ölçüler, bir birim olasılığın Ω üzerine **düzgün** yayıldığı durumları modeller: bir olayın olasılığı yalnızca büyüklüğüne bağlıdır, konumuna değil. Tek bir noktanın uzunluğu, alanı ve hacmi 0 olduğundan $P(\{\omega\}) = 0$ 'dır; aralıkların uçlarını, bölgelerin sınırlarını olaya katıp katmamak olasılığı değiştirmez. $k = 1$ ve $\Omega = [0, 1]$ için elde edilen ölçü, **Örnek 4.4**'ndeki ölçüdür.

⚠ Rastgele sözcüğü tek başına bir model belirlemez

“Rastgele bir nokta seçiliyor” demek, hangi Ω ’nın alınacağını ve düzgünlüğün hangi ölçüye göre olacağını söylemeden anlamsızdır. Farklı seçimler farklı yanıtlar verebilir; bu bir çelişki değil, modelin eksik belirlenmiş olmasıdır (klasik örnek Bertrand paradoksudur). Aşağıdaki örnek bunu somut olarak gösterir.

Örnek 4.6 (Para ve Silindir Tabanı). Yarıçapı 1 birim olan madeni bir para, taban yarıçapı 3 birim olan bir silindirin içine atılıyor ve tabana yatay olarak düşüyor. Paranın tabanın merkezini örtmesi olasılığı nedir? Deneyi (a) uzunluk ölçüsüyle ve (b) alan ölçüsüyle modelleyiniz; hangi modelin deneyi doğru betimlediğini açıklayınız.

Çözüm.

Kurulum. Tabanın merkezi O , paranın merkezi M ve $d = |OM|$ olsun. Para silindirin içinde kaldığından tabanın sınırını aşamaz: paranın her noktası M ’ye en çok 1 birim uzaklıkta olduğundan $d + 1 \leq 3$, yani $d \leq 2$ olmalıdır. Öyleyse $d \in [0, 2]$ ’dir. Para O noktasını $d \leq 1$ olduğunda örter ($d = 1$ sınır durumunun olasılığı 0’dır; eşitliği katmak ya da katmamak sonucu değiştirmez).

a) Uzunluk modeli. Sonuç olarak yalnızca d uzaklığını alalım:

$$\Omega_1 = \{d : 0 \leq d \leq 2\} = [0, 2], \quad \mathcal{U}_1 = \{A \in \mathcal{B}(\mathbb{R}) : A \subset \Omega_1\}, \quad P_1(A) = \frac{\text{uzunluk}(A)}{2}.$$

Merkezin örtülmesi olayı $A_1 = \{d : 0 \leq d \leq 1\} = [0, 1]$ ve

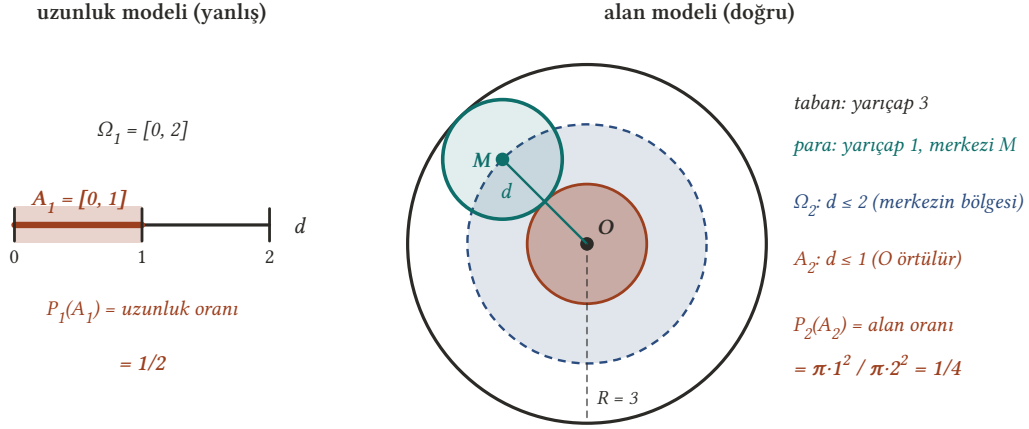
$$P_1(A_1) = \frac{1}{2}.$$

b) Alan modeli. Tabanda başlangıç noktası O olan bir dik koordinat sistemi alalım ve paranın merkezini (x, y) ile gösterelim. $d \leq 2$ koşulu, merkezin yarıçapı 2 olan diskte kalması demektir:

$$\Omega_2 = \{(x, y) : x^2 + y^2 \leq 4\}, \quad \mathcal{U}_2 = \{A \in \mathcal{B}(\mathbb{R}^2) : A \subset \Omega_2\}, \quad P_2(A) = \frac{\text{alan}(A)}{\text{alan}(\Omega_2)} = \frac{\text{alan}(A)}{4\pi}.$$

Merkezin örtülmesi olayı $A_2 = \{(x, y) : x^2 + y^2 \leq 1\}$ birim diskte ve

$$P_2(A_2) = \frac{\pi \cdot 1^2}{\pi \cdot 2^2} = \frac{1}{4}.$$



Şekil 4.3: Yarıçapı 1 olan para, taban yarıçapı 3 olan silindirin tabanına düşüyor; paranın merkezi M taban merkezine $d \leq 2$ uzaklıkta kalır ve merkez ancak $d \leq 1$ iken örtülür. Solda yalnızca d uzaklığını kaydeden uzunluk modeli, $[0, 1]$ aralığının $[0, 2]$ içindeki payı olarak 1/2 verir. Sağda paranın merkezinin düşebileceği bölge (kesikli disk) ve merkezin örtüldüğü bölge (turuncu disk) görülür; şansın alanla orantılı olduğu doğru modelde olasılık alan oranı 1/4'tür. Fark, merkezden uzak konumların yakın konumlardan daha çok yer kaplamasından gelir.

Hangi model doğru? Fiziksel deneyde para gelişigüzel atılır; merkezinin izin verilen bölgenin (yarıçapı 2 olan diskin) herhangi bir parçasına düşme şansı o parçanın **alanıyla** orantılıdır. Bu tam olarak (b) modelidir. (a) modeli ise d uzaklığının $[0, 2]$ üzerinde düzgün dağıldığını varsayar; bu varsayım fiziksel tablodan çıkmaz, hatta onunla çelişir. Gerçekten, alan modelinde $0 \leq r \leq 2$ için

$$P_2(d \leq r) = \frac{\pi r^2}{4\pi} = \frac{r^2}{4}$$

iken uzunluk modeli $P_1(d \leq r) = \frac{r}{2}$ verir. Fark şuradan gelir: merkezden r ile $r + \Delta$ uzaklıktaki halkanın alanı yaklaşık $2\pi r \Delta$ 'dır ve r ile büyür; merkezden uzak konumlar yakın konumlardan “daha çoktur”. Uzunluk modeli her uzaklığa eşit ağırlık vererek merkeze yakın konumları olduğundan fazla sayar. Sonuç olarak deneyin doğru modeli alan modelidir ve aranan olasılık $\frac{1}{4}$ 'tür. Uzunluk modeli başka bir deneyi modeller: tabanın merkezinden başlayan 2 birim uzunluğundaki bir doğru parçası (paranın merkezinin düşebileceği uzaklıklar aralığı) üzerinde rastgele bir nokta seçilmesini.

Geometrik olasılık problemlerinde izlenecek yol hep aynıdır: sonucu bir nokta olarak kodla, Ω bölgesini yaz, olayı Ω içinde bir bölge olarak betimle, büyüklükleri oranla. Kalan üç örnek bu yolu izler.

Örnek 4.7 (Çarpımı Yarımdan Küçük İki Sayı). $(0, 1)$ aralığındaki gerçel sayılardan rastgele iki sayı seçiliyor. Çarpımlarının $\frac{1}{2}$ 'den küçük olması olasılığı nedir?

Çözüm.

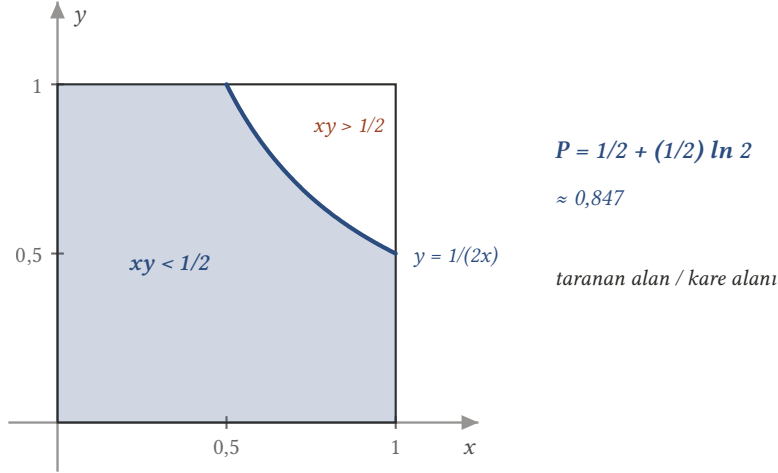
Seçilen sayılar x ve y olsun; sonuç (x, y) noktasıdır. “Rastgele” sözünü, (x, y) noktasının birim kare üzerinde düzgün dağılması olarak anlıyoruz: her iki sayı $(0, 1)$ üzerinde düzgün seçiliyor ve birinin seçimi ötekini etkilemiyor. Buna göre

$$\Omega = (0, 1) \times (0, 1), \quad \mathcal{U} = \{A \in \mathcal{B}(\mathbb{R}^2) : A \subset \Omega\}, \quad P(A) = \frac{\text{alan}(A)}{\text{alan}(\Omega)} = \text{alan}(A)$$

alınır; birim karenin alanı 1'dir. Sorulan olay

$$A = \left\{ (x, y) \in \Omega : xy < \frac{1}{2} \right\}.$$

Bölgeyi x 'e göre inceleyelim. $0 < x \leq \frac{1}{2}$ ise $y < 1$ olduğundan $xy < x \leq \frac{1}{2}$ kendiliğinden sağlanır: bu şeritte bütün y değerleri A 'dadır. $\frac{1}{2} < x < 1$ ise $xy < \frac{1}{2}$ koşulu $y < \frac{1}{2x}$ ile eşdeğerdir ve $\frac{1}{2x} \in (\frac{1}{2}, 1)$ olduğundan bu gerçek bir kısıttır. Demek ki A , $(0, \frac{1}{2}] \times (0, 1)$ dikdörtgeni ile $x \in (\frac{1}{2}, 1)$ üzerinde $y = \frac{1}{2x}$ hiperbolünün altında kalan bölgenin birleşimidir.



Şekil 4.4: $(0, 1)$ aralığından rastgele seçilen iki sayının çarpımının $1/2$ 'den küçük olması olayı. Örnek uzay birim karedir; olay, $y = 1/(2x)$ hiperbolünün altında kalan bölgedir. Olasılık, alanların oranıdır: $1/2 + (1/2) \ln 2 \approx 0,847$.

Alanı hesaplayalım:

$$\text{alan } (A) = \frac{1}{2} \cdot 1 + \int_{1/2}^1 \frac{1}{2x} dx = \frac{1}{2} + \frac{1}{2} [\ln x]_{1/2}^1 = \frac{1}{2} + \frac{1}{2} \left(0 - \ln \frac{1}{2}\right) = \frac{1}{2} + \frac{1}{2} \ln 2.$$

Öyleyse

$$P(A) = \frac{1}{2} + \frac{1}{2} \ln 2 = \frac{1 + \ln 2}{2} \approx 0,8466.$$

■

Örnek 4.8 (Doğru Parçası Üzerinde Bölme Oranı). k uzunluğunda bir AB doğru parçası üzerinde rastgele bir C noktası işaretleniyor. $a > 0$ verilmiş bir sayı olmak üzere, $\frac{|AC|}{|BC|}$ oranının a 'dan küçük olması olasılığını bulunuz.

Çözüm.

A noktasını başlangıç alıp C 'nin konumunu $x = |AC| \in [0, k]$ ile kodlayalım. “Rastgele” işaretleme, x 'in $[0, k]$ üzerinde düzgün dağılması demektir:

$$\Omega = [0, k], \quad \mathcal{U} = \{D \in \mathcal{B}(\mathbb{R}) : D \subset \Omega\}, \quad P(D) = \frac{\text{uzunluk } (D)}{k}.$$

$x = k$ (yani $C = B$) durumunda oran tanımsızdır; ama tek bir noktanın olasılığı 0 olduğundan bu durumu göz ardı edebiliriz ve $|BC| = k - x > 0$ alabiliriz. Sorulan olay

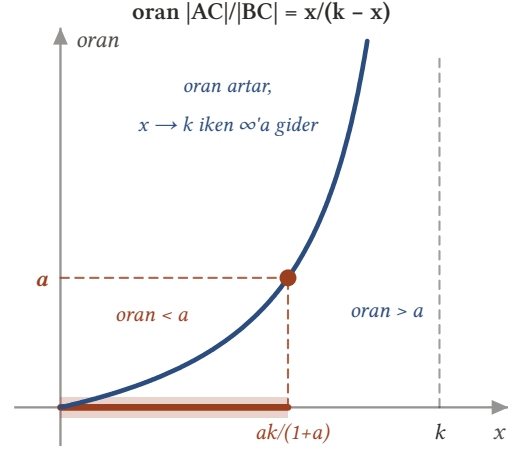
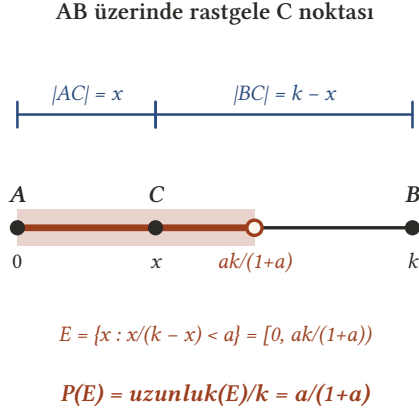
$$E = \left\{x \in [0, k) : \frac{x}{k-x} < a\right\}.$$

$k - x > 0$ olduğundan eşitsizlik paydayla çarpılabilir:

$$\frac{x}{k-x} < a \Leftrightarrow x < a(k-x) \Leftrightarrow x(1+a) < ak \Leftrightarrow x < \frac{ak}{1+a}.$$

$a > 0$ için $\frac{ak}{1+a} < k$ olduğundan $E = [0, \frac{ak}{1+a})$ bir alt aralıktır ve

$$P(E) = \frac{1}{k} \cdot \frac{ak}{1+a} = \frac{a}{1+a}.$$



Şekil 4.5: k uzunluğundaki AB doğru parçası üzerinde rastgele bir C noktası: konumu $x = |AC|$ ile kodlanır ve $[0, k]$ üzerinde düzgün dağılır. Solda, $|AC|/|BC| < a$ olayının tam olarak $[0, ak/(1+a))$ alt aralığı olduğu görülür; olasılığı uzunluk oranı $a/(1+a)$ 'dır ve k 'ye bağlı değildir. Sağda oran fonksiyonu $x/(k-x)$ çizilidir: sıfırdan başlayıp artarak sonsuza gittiğinden a düzeyinin altında kaldığı küme bir başlangıç aralığıdır.

Sonuç k 'ye bağlı değildir. Sağlama: $a = 1$ için $P(E) = \frac{1}{2}$; bu, " C 'nin A 'ya B 'den daha yakın olması" olayıdır ve simetri gereği $\frac{1}{2}$ olmalıdır. $a \rightarrow 0^+$ iken $P(E) \rightarrow 0$, $a \rightarrow \infty$ iken $P(E) \rightarrow 1$ olması da beklendiği gibidir.

■

Örnek 4.9 (İki Rastgele Nokta ve Yakınlık). k uzunluğundaki AB doğru parçası üzerinde rastgele L ve M noktaları işaretleniyor. L noktasının M noktasına, A noktasına olduğundan daha yakın olması olasılığını uygun bir olasılık uzayı kurarak bulunuz.

Çözüm.

$\ell = |AL|$ ve $m = |AM|$ olsun; sonuç $(\ell, m) \in [0, k] \times [0, k]$ noktasıdır ve kenarı k olan kare üzerinde düzgün dağılır; olasılık ölçüsü alan bölü k^2 'dir. Hesabı sadeleştirmek için $u = \ell/k$, $v = m/k$ ölçeklemesini yapalım: (u, v) noktası birim kare $\Omega = [0, 1] \times [0, 1]$ üzerinde düzgün dağılır (ölçekleme bütün alanları aynı çarpanla çarptığından oranlar değişmez) ve $P(D) = \text{alan}(D)$ olur.

Sorulan olay $|LM| < |AL|$, yani $|\ell - m| < \ell$; ölçekli değişkenlerle $|u - v| < u$ 'dur. Mutlak değeri açalım:

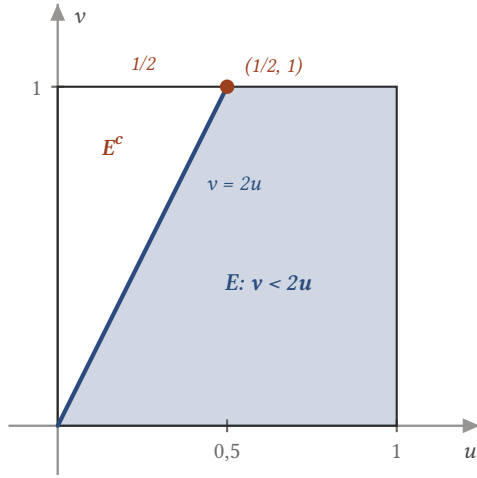
$$|u - v| < u \Leftrightarrow -u < v - u < u \Leftrightarrow 0 < v < 2u.$$

$v \geq 0$ zaten sağlandığından ve $v = 0$ doğrusunun alanı 0 olduğundan olay

$$E = \{(u, v) \in \Omega : v < 2u\}$$

bölgesidir. Tümleneni $E^c = \{(u, v) \in \Omega : v \geq 2u\} = \{(u, v) \in \Omega : u \leq v/2\}$, köşeleri $(0, 0)$, $(0, 1)$ ve $(\frac{1}{2}, 1)$ olan üçgendir; alanı

$$\text{alan}(E^c) = \frac{1}{2} \cdot \frac{1}{2} \cdot 1 = \frac{1}{4}.$$



$$\text{alan}(E^c) = 1/2 \cdot 1/2 \cdot 1 = 1/4$$

$$P(E) = 1 - 1/4 = 3/4$$

olasılık = taranan alan
(kare alanı 1)

Şekil 4.6: Bir doğru parçası üzerinde rastgele seçilen iki nokta için ölçekli sonuç uzayı: birim kare, olasılık ölçüsü alan. L 'nin M 'ye A 'dan daha yakın olması olayı $|u - v| < u$, yani $v < 2u$ bölgesidir (taralı). Tümleyeni, köşeleri $(0, 0)$, $(0, 1)$ ve $(1/2, 1)$ olan üçgendir; alanı $1/4$ olduğundan aranan olasılık $3/4$ 'tür ve parçanın uzunluğundan bağımsızdır.

Dolayısıyla

$$P(E) = 1 - \frac{1}{4} = \frac{3}{4}.$$

Sonuç k 'den bağımsızdır.

■

4.5 Alıştırmalar

Alıştırma 4.1 (Olasılık Uzayı Örnekleri). a) Hileli bir zarda i yüzünün gelme ağırlığı $p_i = \frac{i}{21}$, $i = 1, \dots, 6$ olsun. Bunun bir olasılık uzayı tanımladığını gösteriniz; çift sayı gelme, 5 ya da 6 gelme ve asal sayı gelme olasılıklarını bulunuz.

b) $\Omega = \{\omega_1, \omega_2, \dots\}$ ve $p_i = c\left(\frac{2}{3}\right)^i$, $i = 1, 2, \dots$ olsun. c sabitini belirleyiniz ve çift indisli sonuçlar kümesinin olasılığını bulunuz. Bu uzay hangi deneyi modeller?

c) $F(x) = 0$ ($x < 0$), $F(x) = \frac{x^2}{4}$ ($0 \leq x < 2$), $F(x) = 1$ ($x \geq 2$) fonksiyonunun bir dağılım fonksiyonu olduğunu gösteriniz; belirlediği P için $P((1, 2])$, $P(\{2\})$, $P([\frac{1}{2}, 1))$ ve $P((3, \infty))$ değerlerini hesaplayınız.

d) $F(x) = 0$ ($x < 0$), $F(x) = \frac{1}{4}$ ($0 \leq x < 1$), $F(x) = \frac{3}{4}$ ($1 \leq x < 2$), $F(x) = 1$ ($x \geq 2$) olsun. $P(\{0\})$, $P(\{1\})$, $P(\{2\})$, $P((0, 2))$, $P([1, \infty))$ ve $P(\mathbb{R} \setminus \{0, 1, 2\})$ değerlerini bulunuz. Bu dağılım hangi deneyi modeller?

e) Birim kareden rastgele bir (x, y) noktası seçiliyor. $x + y \leq 1$, $|x - y| < \frac{1}{2}$ ve $y > x^2$ olaylarının olasılıklarını bulunuz.

f) k uzunluğundaki bir çubuk rastgele bir noktadan kırılıyor. Uzun parçanın kısa parçanın en az iki katı olması olasılığı nedir?

Çözüm.

a) Ağırlıklar pozitifdir ve $\sum_{i=1}^6 \frac{i}{21} = \frac{21}{21} = 1$ olduğundan **Önerme 4.1** uygulanır. Çift sayı: $\frac{2+4+6}{21} = \frac{12}{21} = \frac{4}{7}$. 5 ya da 6: $\frac{5+6}{21} = \frac{11}{21}$. Asal sayı $(2, 3, 5)$: $\frac{2+3+5}{21} = \frac{10}{21}$.

b) Geometrik seri ile $\sum_{i=1}^{\infty} \left(\frac{2}{3}\right)^i = \frac{2/3}{1-2/3} = 2$; toplamın 1 olması için $c = \frac{1}{2}$ olmalıdır. Böylece $p_i = \frac{1}{2} \left(\frac{2}{3}\right)^i = \frac{1}{3} \left(\frac{2}{3}\right)^{i-1}$. Çift indisler:

$$\sum_{j=1}^{\infty} \frac{1}{2} \left(\frac{2}{3} \right)^{2j} = \frac{1}{2} \sum_{j=1}^{\infty} \left(\frac{4}{9} \right)^j = \frac{1}{2} \cdot \frac{4/9}{1-4/9} = \frac{1}{2} \cdot \frac{4}{5} = \frac{2}{5}.$$

Model: düzgün bir zar, ilk kez 5 ya da 6 gelinceye kadar atılıyor; ω_i , “ilk 5 ya da 6’nın i -inci atışta gelmesi”. i atışlık 6^i eşit olasılıklı diziden uygun olanlar $4^{i-1} \cdot 2$ tanedir; Laplace formülü $\frac{2 \cdot 4^{i-1}}{6^i} = \frac{1}{3} \left(\frac{2}{3} \right)^{i-1}$ verir.

c) $\frac{x^2}{4}$, $[0, 2]$ üzerinde artandır ve 0’dan 1’e gider; parça uçlarında $F(0) = 0$ ve $\lim_{x \rightarrow 2^-} F(x) = 1 = F(2)$ olduğundan F sürekli, dolayısıyla sağdan sürekli ve azalmayandır; $x < 0$ için 0, $x \geq 2$ için 1 olduğundan limit koşulları sağlanır. Hesaplar: $P((1, 2]) = F(2) - F(1) = 1 - \frac{1}{4} = \frac{3}{4}$; F sürekli olduğundan $P(\{2\}) = 0$; $P([\frac{1}{2}, 1]) = F(1) - F(\frac{1}{2}) = \frac{1}{4} - \frac{1}{16} = \frac{3}{16}$; $P((3, \infty)) = 1 - F(3) = 0$.

d) Sıçramalar: $P(\{0\}) = F(0) - F(0^-) = \frac{1}{4}$, $P(\{1\}) = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}$, $P(\{2\}) = 1 - \frac{3}{4} = \frac{1}{4}$. $P((0, 2)) = F(2^-) - F(0) = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}$, ki bu $P(\{1\})$ ’e eşittir. $P([1, \infty)) = 1 - F(1^-) = 1 - \frac{1}{4} = \frac{3}{4}$. Üç noktanın olasılıkları 1’e toplandığından $P(\mathbb{R} \setminus \{0, 1, 2\}) = 0$. Model: iki düzgün para atılıp gelen tura sayısı gözleniyor; $\{YY, YT, TY, TT\}$ Laplace uzayında 0, 1, 2 tura sayılarının olasılıkları $\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$ ’tür.

e) Ω birim kare, P alan. $x + y \leq 1$: köşeleri $(0, 0), (1, 0), (0, 1)$ olan üçgen, alanı $\frac{1}{2}$. $|x - y| < \frac{1}{2}$: tümleyeni $y - x \geq \frac{1}{2}$ ve $x - y \geq \frac{1}{2}$ bölgeleridir; her biri dik kenarları $\frac{1}{2}$ olan bir üçgendir, alanları $\frac{1}{8}$; olasılık $1 - \frac{2}{8} = \frac{3}{4}$. $y > x^2$: parabolün üstünde kalan bölge, alanı $\int_0^1 (1 - x^2) dx = 1 - \frac{1}{3} = \frac{2}{3}$.

f) Kırılma noktası $x \in [0, k]$ düzgün dağılsın; parçalar x ve $k - x$, kısa parça $s = \min(x, k - x)$, uzun parça $k - s$. Koşul $k - s \geq 2s$, yani $s \leq \frac{k}{3}$; bu da $x \leq \frac{k}{3}$ ya da $x \geq \frac{2k}{3}$ demektir. Uygun kümenin uzunluğu $\frac{k}{3} + \frac{k}{3} = \frac{2k}{3}$, olasılık $\frac{2}{3}$.

■

Laplace formülü olasılık hesabını $n(A)$ ve $n(\Omega)$ sayılarını bulmaya indirger; örnek uzay büyüdüğünde bu sayma için sistematik yöntemler gerekir: [Sayma Teknikleri](#).

5. Sayma Teknikleri

Olasılık Uzayı Örnekleri ve Geometrik Olasılık bölümünde Laplace olasılık uzayını tanımladık (Tanım 4.1): örnek uzay sonlu ve her sonuç eşit olasılıklı olduğunda bir A olayının olasılığı, uygun durumların sayısının bütün durumların sayısına oranıdır (Önerme 4.2):

$$P(A) = \frac{|A|}{|\Omega|}.$$

Burada $|A|$, önceki bölümde $n(A)$ ile de gösterdiğimiz eleman sayısıdır; bu bölümde yalnızca $|A|$ yazacağız. Bu formül olasılık hesabını iki sayma işine indirger: Ω 'nın eleman sayısını ve A 'nın eleman sayısını bulmak. Bir zar, iki para gibi küçük deneylerde bu sayılar elle bulunur. Ama örnek uzay binlerce ya da milyarlarca noktadan oluşuyorsa tek tek saymak olanaksızdır. Bir sınıftaki 23 kişinin doğum günleri için 365^{23} olası liste vardır; bunların içinde en az iki kişinin doğum gününün çakıştığı listeleri saymak, sistemli bir yöntem olmadan yapılamaz.

Bu bölümde bütün sayma yöntemlerinin dayandığı iki temel kuralı, **toplama** ve **çarpma** kurallarını ispatlayacağız; ardından bunlardan **permütasyon**, **kombinasyon**, **tekrarlı permütasyon** (çok terimli katsayı) ve **tekrarlı kombinasyon** (yıldız-çubuk) formüllerini türeteceğiz. Yol üstünde **binom teoremi**-ni tümevarımla ispatlayacak, **genelleştirilmiş binom katsayısını** tanımlayacağız. Her formülü somut sayma problemleriyle sınavacak, sonunda da Laplace uzayında olasılık hesaplarına döneceğiz. Burada kurulan araçlar yalnızca bu bölümün değil, ileride binom, hipergeometrik ve negatif binom dağılımlarının da temelidir.

5.1 Toplama ve Çarpma Kuralları

Her sayma yöntemi iki basit gözleme dayanır. Birincisi, birbirini dışlayan seçenekleri sayarken sayıların **toplandığını** söyler; ikincisi, art arda yapılan seçimleri sayarken sayıların **çarpıldığını** söyler. İkisini de ispatlarıyla yazıyoruz, çünkü sonraki her formül bu ikisine indirgenecektir.

Önerme 5.1 (Toplama Kuralı). *Bir iş iki farklı yoldan yapılabilir: birinci yol n_1 farklı biçimde, ikinci yol n_2 farklı biçimde gerçekleştirilebilir ve iki yolun ortak hiçbir biçimi olmasın. O zaman iş $n_1 + n_2$ farklı biçimde yapılabilir.*

Daha genel olarak, birbirini dışlayan k yol sırasıyla $n_1, n_2, \dots, n_k$ biçimde gerçekleştirilebiliyorsa iş $n_1 + n_2 + \dots + n_k$ farklı biçimde yapılabilir.

İspat.

Birinci yolun biçimlerinin kümesi A_1 , ikincinin A_2 olsun; $|A_1| = n_1$, $|A_2| = n_2$ ve varsayım gereği $A_1 \cap A_2 = \emptyset$. İş yapmanın bütün biçimleri $A_1 \cup A_2$ kümesini oluşturur. Ayırık iki sonlu kümenin birleşiminin eleman sayısı, eleman sayılarının toplamıdır: A_1 'in elemanlarını $1, 2, \dots, n_1$ sayılarıyla, A_2 'nin elemanlarını $n_1 + 1, \dots, n_1 + n_2$ sayılarıyla numaralayalım. Kümeler ayırık olduğundan hiçbir eleman iki numara almaz ve her numara tam bir elemana gider; böylece $A_1 \cup A_2$ ile $\{1, 2, \dots, n_1 + n_2\}$ arasında birebir ve örten bir eşleme kurulmuş olur. Öyleyse $|A_1 \cup A_2| = n_1 + n_2$.

k yol için ifade k üzerinde tümevarımla çıkar: $A_1 \cup \dots \cup A_k = (A_1 \cup \dots \cup A_{k-1}) \cup A_k$ yazılır. Parantez içindeki küme A_k ile ayırıktır; iki küme için kanıtlanan sonuç ve tümevarım hipotezi birlikte $|A_1 \cup \dots \cup A_k| = (n_1 + \dots + n_{k-1}) + n_k$ verir.

■

Toplama kuralı “ya biri ya öteki” biçimindeki seçimleri sayar. Bir seçimin ardından bir başka seçim geliyorsa ihtiyaç duyulan kural farklıdır.

Önerme 5.2 (Çarpma Kuralı). *Bir iş art arda iki adımda yapılsın. Birinci adım n_1 farklı biçimde yapılabilir ve birinci adım hangi biçimde yapılmış olursa olsun ikinci adım n_2 farklı biçimde yapılabilir. O zaman iş $n_1 n_2$ farklı biçimde yapılabilir.*

Daha genel olarak, k adımlı bir işte i -inci adım, önceki adımlar nasıl yapılmış olursa olsun, n_i farklı biçimde yapılabiliyorsa iş $n_1 n_2 \cdots n_k$ farklı biçimde yapılabilir.

İspat.

Birinci adımın biçimlerinin kümesi $A = \{a_1, a_2, \dots, a_{n_1}\}$ olsun. Birinci adım a_i biçiminde yapıldığında ikinci adımın biçimlerinin kümesi B_i olsun; varsayım gereği her i için $|B_i| = n_2$. İşin yapmanın bir biçimi, $b \in B_i$ olmak üzere bir (a_i, b) çiftiyle belirlenir. Bu çiftlerin kümesi

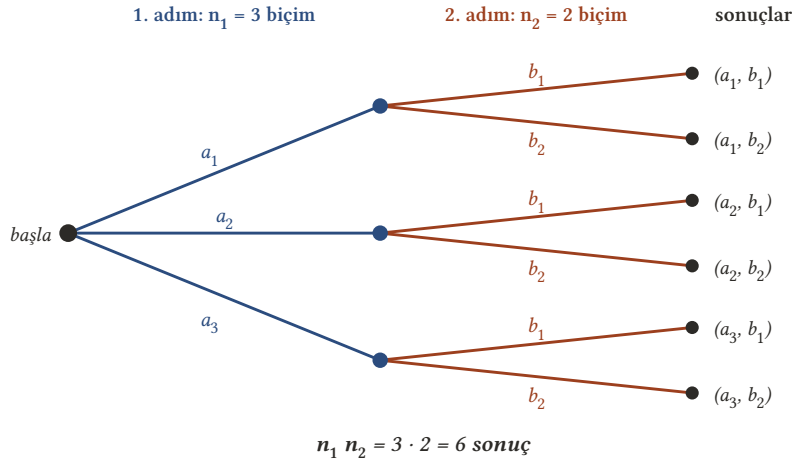
$$\{(a_i, b) : i = 1, \dots, n_1, b \in B_i\} = \bigcup_{i=1}^{n_1} (\{a_i\} \times B_i)$$

biçiminde yazılır. Sağdaki n_1 küme, ilk bileşenleri farklı olduğundan ikiye ikiye ayrılır ve her birinin n_2 elemanı vardır. Toplama kuralına göre (Önerme 5.1) birleşimin eleman sayısı

$$\underbrace{n_2 + n_2 + \cdots + n_2}_{n_1 \text{ terim}} = n_1 n_2$$

olur.

k adım için yine tümevarım: ilk $k - 1$ adım tek bir “birinci adım” gibi görülür. Tümevarım hipotezine göre bu $n_1 n_2 \cdots n_{k-1}$ biçimde yapılır; ardından k -üncü adım, öncekilerden bağımsız olarak, n_k biçimde yapılır. İki adım için kanıtlanan sonuç $n_1 n_2 \cdots n_{k-1} \cdot n_k$ verir.



Şekil 5.1: Çarpma kuralı için ağaç diyagramı. Birinci adım a_1, a_2, a_3 biçimlerinden biriyle yapılır ($n_1 = 3$ dal); hangisi seçilmiş olursa olsun ikinci adım b_1 ya da b_2 biçimiyle yapılır ($n_2 = 2$ dal). Her yaprak bir (a_i, b_j) çiftine, yani işin bir yapılış biçimine karşılık gelir; yaprak sayısı dal sayılarının çarpımı $3 \cdot 2 = 6$ 'dır.

💡 Toplama mı, çarpma mı?

Sorudaki bağlaç yol gösterir. “Matematik kitaplarından **ya da** fizik kitaplarından bir kitap seç” (3 matematik, 4 fizik kitabı): seçenekler birbirini dışlar, toplama kuralı, $3 + 4 = 7$. “Bir matematik **ve** bir fizik kitabı seç”: seçimler art arda yapılır ve ikincinin seçenek sayısı birincinin seçiminden etkilenmez, çarpma kuralı, $3 \cdot 4 = 12$.

Çarpma kuralındaki koşula dikkat edilmelidir: ikinci adımın **seçenek sayısı** birinci adımdan bağımsız olmalıdır; seçeneklerin kendileri değişebilir. Bir torbadan iade etmeden iki top çekerken ikinci çekilişte hangi topların kaldığı ilk çekilişe bağlıdır, ama kalan top sayısı hep bir eksiktir; kural uygulanır.

5.2 Permütasyon

En sık karşılaşılan sayma sorusu şudur: n farklı nesneden r tanesi seçilip bir sıraya dizilecekse kaç diziliş vardır? Sıranın önemli olduğu bu düzenlemelere ad verelim.

Tanım 5.1 (Permütasyon). n farklı nesneden oluşan bir kümenin tamamının ya da bir kısmının belirli bir sıraya dizilmesine bir **permütasyon** denir. n nesneden r tanesinin ($0 \leq r \leq n$) seçilip sıraya dizilmesiyle elde edilen düzenlemeye n nesnenin bir **r -li permütasyonu** denir; seçilen bir nesne yeniden kullanılmaz (yinelemesiz permütasyon). Bu permütasyonların sayısı $P(n, r)$ ile gösterilir.

Permütasyonda sıra önemlidir: a, b, c nesnelerinin 2-li permütasyonları ab, ba, ac, ca, bc, cb olmak üzere altı tanedir; ab ile ba farklı sayılır. Bu sayıyı genel olarak veren formül, çarpma kuralının ilk doğrudan uygulamasıdır.

Teorem 5.1 (Permütasyon Sayısı). n farklı nesnenin r -li permütasyonlarının sayısı

$$P(n, r) = n(n-1)(n-2)\cdots(n-r+1) = \frac{n!}{(n-r)!}$$

dir. Özel olarak n nesnenin tümünün sıralanışlarının sayısı $P(n, n) = n!$ ve $P(n, 0) = 1$ 'dir. Burada $0! = 1$ kabul edilir.

İspat.

Bir r -li permütasyon r adımda kurulur: birinci yere konacak nesne seçilir, sonra ikinci yere konacak nesne, ..., sonunda r -inci yere konacak nesne. Birinci yer için n seçenek vardır. Birinci yer nasıl doldurulmuş olursa olsun, kullanılan nesne bir daha kullanılmayacağından ikinci yer için $n-1$ seçenek kalır. Genel olarak i -inci yere gelindiğinde $i-1$ nesne kullanılmıştır ve $n-i+1$ seçenek vardır; r -inci yer için $n-r+1$ seçenek. Her adımdaki seçenek sayısı önceki adımların nasıl yapıldığına bağlı olmadığından çarpma kuralı (Önerme 5.2) uygulanır:

$$P(n, r) = n(n-1)\cdots(n-r+1).$$

Sağ tarafı $(n-r)!$ ile çarpıp bölelim:

$$n(n-1)\cdots(n-r+1) = \frac{n(n-1)\cdots(n-r+1)(n-r)(n-r-1)\cdots 2 \cdot 1}{(n-r)!} = \frac{n!}{(n-r)!}.$$

$r = n$ için $(n-n)! = 0! = 1$ olduğundan $P(n, n) = n!$ çıkar. $r = 0$ için hiçbir yer doldurulmaz; tek bir (boş) düzenleme vardır ve formül $n!/n! = 1$ verir.

■

Örnek 5.1 (Kitaplıktaki Üç Boş Yer). 7 farklı kitap arasından seçilen 3 kitap, bir kitaplıktaki 3 boş yere kaç farklı biçimde yerleştirilebilir?

Çözüm.

Üç boş yer birbirinden farklıdır: soldaki, ortadaki ve sağdaki. Dolayısıyla hangi kitabın hangi yere konduğu önemlidir, yani sıra önemlidir. Aranılan sayı 7 nesnenin 3-lü permütasyonlarının sayısıdır (Teorem 5.1):

$$P(7, 3) = 7 \cdot 6 \cdot 5 = \frac{7!}{4!} = 210.$$

Çarpma kuralıyla doğrudan da görülür: birinci boş yer için 7, ikincisi için 6, üçüncüsü için 5 seçenek vardır.

■

5.3 Kombinasyon

Çoğu zaman sıra önemli değildir. 10 kişiden 3 kişilik bir komite seçerken komitenin kimlerden oluştuğu önemlidir, kimin önce seçildiği değil. Bu durumda saydığımız şey diziliş değil, alt kümedir.

Tanım 5.2 (Kombinasyon). n farklı nesneden r tanesinin ($0 \leq r \leq n$) sıraya bakılmaksızın seçilmesine, yani n elemanlı bir kümenin r elemanlı bir alt kümesine, n nesnenin bir **r -li kombinasyonu** denir. Bu kombinasyonların sayısı $C(n, r)$ ya da $\binom{n}{r}$ ile gösterilir ve **binom katsayısı** diye adlandırılır.

a, b, c nesnelerinin 2-li kombinasyonları $\{a, b\}, \{a, c\}, \{b, c\}$ olmak üzere üç tanedir; 2-li permütasyonların altı tanesi, her kombinasyonun iki biçimde sıralanmasıyla oluşur. Bu gözlem genel formülün ispatının ta kendisidir.

Teorem 5.2 (Kombinasyon Sayısı). $0 \leq r \leq n$ için

$$C(n, r) = \binom{n}{r} = \frac{n!}{r!(n-r)!} = \frac{P(n, r)}{r!}$$

dir. Ayrıca $\binom{n}{r} = \binom{n}{n-r}$ ve $\binom{n}{0} = \binom{n}{n} = 1$ ’dir.

İspat.

n nesnenin her r -li permütasyonu iki adımda kurulabilir: önce hangi r nesnenin kullanılacağı seçilir (bir r -li kombinasyon; $C(n, r)$ biçimde), sonra seçilen r nesne sıraya dizilir ($P(r, r) = r!$ biçimde). Her permütasyon bu yolla tam bir kez elde edilir, çünkü bir permütasyon hem hangi nesnelerin kullanıldığını hem de sıralarını belirler. İkinci adımdaki seçenek sayısı birinci adımdaki seçimden bağımsız olduğundan çarpma kuralı (Önerme 5.2)

$$P(n, r) = r! \cdot C(n, r)$$

verir. Buradan Teorem 5.1 ile

$$C(n, r) = \frac{P(n, r)}{r!} = \frac{n!}{r!(n-r)!}$$

bulunur.

Simetri özelliği: r elemanlı bir alt küme seçmek, dışarıda bırakılacak $n - r$ elemanı seçmekle aynı şeydir; alt kümeyi tümleyenine gönderen eşleme birebir ve örtendir. Öyleyse $\binom{n}{r} = \binom{n}{n-r}$. Aynı sonuç formülden de okunur: r ile $n - r$ yer değiştirince payda değişmez. $r = 0$ için tek alt küme $\emptyset$, $r = n$ için tek alt küme kümenin kendisidir; formül de $0! = 1$ ile 1 verir.

■

Binom katsayılarının en önemli özdeşliği, binom teoreminin ispatında gereksinim duyacağımız Pascal özdeşliğidir. Bu özdeşliğin bir sayma yorumu vardır; ispatı da toplama kuralına dayanır.

Lemma 5.1 (Pascal Özdeşliği). $1 \leq r \leq n$ için

$$\binom{n+1}{r} = \binom{n}{r-1} + \binom{n}{r}.$$

İspat.

$n + 1$ elemanlı bir S kümesi ve bunun belirli bir x elemanı alalım. S ’nin r elemanlı alt kümeleri iki ayrık sınıfa ayrılır: x ’i içerenler ve içermeyenler. x ’i içeren bir alt küme, $S \setminus \{x\}$ kümesinden seçilen $r - 1$ elemanla belirlenir; $S \setminus \{x\}$ ’in n elemanı olduğundan bunlar $\binom{n}{r-1}$ tanedir. x ’i içermeyen bir alt küme, $S \setminus \{x\}$ ’in r elemanlı bir alt kümesidir; $\binom{n}{r}$ tane. Toplama kuralı (Önerme 5.1) iki sayıyı toplar ve S ’nin r elemanlı alt kümelerinin sayısı olan $\binom{n+1}{r}$ ’yi verir.

Cebirsel doğrulama da kısıdır:

$$\binom{n}{r-1} + \binom{n}{r} = \frac{n!}{(r-1)!(n-r+1)!} + \frac{n!}{r!(n-r+1)!} = \frac{n!(n+1)}{r!(n+1-r)!} = \binom{n+1}{r}.$$

satır		satır toplamı
$n = 0$	1	$2^0 = 1$
$n = 1$	1 1	$2^1 = 2$
$n = 2$	1 2 1	$2^2 = 4$
$n = 3$	1 3 3 1	$2^3 = 8$
$n = 4$	1 4 6 4 1	$2^4 = 16$
$n = 5$	1 5 10 10 5 1	$2^5 = 32$
$n = 6$	1 6 15 20 15 6 1	$2^6 = 64$

$C(n+1, r) = C(n, r-1) + C(n, r)$
 $C(5, 2) = C(4, 1) + C(4, 2) = 4 + 6 = 10$

Şekil 5.2: Pascal üçgeninin 0'dan 6'ya kadar olan satırları; n -inci satırda $C(n, 0), \dots, C(n, n)$ binom katsayıları durur. Pascal özdeşliği her sayının, bir üst satırda sol ve sağ üstündeki iki sayının toplamı olduğunu söyler: vurgulu örnekte $C(5, 2) = C(4, 1) + C(4, 2) = 4 + 6 = 10$. Sağdaki toplamalar, binom teoreminin $a = x = 1$ hâlinde çıkan $\sum_i C(n, r) = 2^n$ eşitliğini gösterir.

■

5.4 Tekrarlı Permütasyon ve Çok Terimli Katsayı

Permütasyon formülü nesnelerin birbirinden farklı olduğunu varsayar. Oysa GAZİANTEP sözcüğünün harflerini sıralarken iki A harfi birbirinden ayırt edilemez; A'ların yerini değiştirmek yeni bir sözcük vermez. Bu tür sayımlar için formülü düzeltmek gerekir.

Teorem 5.3 (Tekrarlı Permütasyon (Çok Terimli Katsayı)). n nesnenin r_1 tanesi birinci çeşitten, r_2 tanesi ikinci çeşitten, ..., r_k tanesi k -ıncı çeşitten olsun; $r_1 + r_2 + \dots + r_k = n$. Aynı çeşitten nesneler birbirinden ayırt edilmesin. Bu n nesnenin tümünün bir sıraya farklı dizilişlerinin sayısı

$$\binom{n}{r_1, r_2, \dots, r_k} = \frac{n!}{r_1! r_2! \dots r_k!}$$

dir. Bu sayıya **çok terimli katsayı** denir. $k = 2$ için $\binom{n}{r, n-r} = \binom{n}{r}$ 'dir.

İspat.

Aranan sayı N olsun. Aynı çeşitten nesneleri geçici olarak etiketleyip birbirinden ayırt edilir kılalım: birinci çeşitten nesnelere $1, \dots, r_1$, ikinci çeşittekilere $1, \dots, r_2$ etiketleri verilsin ve böyle devam edilsin. Artık n farklı nesne vardır ve Teorem 5.1 gereği bunların $n!$ sıralanışı bulunur.

Etiketli sıralanışları şöyle de sayabiliriz: önce etiketsiz bir sıralanış seçilir (N biçimde); sonra bu sıralanışta birinci çeşidin kapladığı r_1 yere etiketler dağıtılır ($r_1!$ biçimde), ikinci çeşidin kapladığı r_2 yere etiketler dağıtılır ($r_2!$ biçimde), ..., k -ıncı çeşidin r_k yerine etiketler dağıtılır ($r_k!$ biçimde). Her etiketli sıralanış bu yolla tam bir kez elde edilir: etiketler silinince hangi etiketsiz sıralanıştan geldiği, silinmeden önce de her çeşit içinde etiketlerin nasıl dağıtıldığı bellidir. Çarpma kuralı (Önerme 5.2)

$$n! = N \cdot r_1! r_2! \dots r_k!$$

verir; N çekilince istenen formül çıkar.

İkinci bir yol formülü kombinasyonlarla kurar. n yerden birinci çeşide ayrılacak r_1 yer $\binom{n}{r_1}$ biçimde seçilir; kalan $n - r_1$ yerden ikinci çeşide ayrılacak r_2 yer $\binom{n-r_1}{r_2}$ biçimde seçilir; bu böyle sürer ve sonunda kalan r_k yer k -inci çeşide verilir. Çarpma kuralı ve Teorem 5.2 ile

$$\binom{n}{r_1} \binom{n-r_1}{r_2} \dots \binom{n-r_1-r_2-\dots-r_{k-1}}{r_k} = \frac{n!}{r_1! (n-r_1)!} \cdot \frac{(n-r_1)!}{r_2! (n-r_1-r_2)!} \dots \frac{r_{k-1}!}{r_k! 0!} = \frac{n!}{r_1! r_2! \dots r_k!};$$

ara faktöriyeller ikiye ikiye sadeleşir.

■

Örnek 5.2 (GAZİANTEP Sözcüğünün Harfleri). (a) GAZİANTEP sözcüğündeki dokuz harfin tümü kullanılarak, anlamlı olup olmadığına bakılmaksızın, kaç farklı sözcük yazılabilir?

(b) Bu harflerden yedi tanesi kullanılarak kaç farklı yedi harfli sözcük yazılabilir?

Çözüm.

GAZİANTEP sözcüğünde dokuz harf vardır: G, A, Z, İ, A, N, T, E, P. A harfi iki kez geçer; öteki yedi harf (G, Z, İ, N, T, E, P) birer kez geçer.

(a) Dokuz nesnenin ikisi aynı çeşittendir (A), kalan yedisi birer çeşittir. Teorem 5.3 ile

$$\frac{9!}{2! 1! 1! 1! 1! 1! 1!} = \frac{9!}{2!} = \frac{362880}{2} = 181440.$$

(b) Yedi harf seçilip sıralanacaktır. Harflerin tümü farklı olsaydı yanıt $P(9, 7)$ olurdu; ama iki A birbirinden ayırt edilemediğinden, seçilen yedi harf arasında kaç A bulunduğuna göre durumları ayırmak gerekir. Üç durum birbirini dışlar; toplama kuralıyla (Önerme 5.1) toplanır.

İki A da kullanılır. Kalan beş harf, A dışındaki yedi farklı harften seçilir: $\binom{7}{5} = 21$ biçimde. Seçilen yedi harfin ikisi özdeş olduğundan sıralanış sayısı $\frac{7!}{2!} = 2520$ 'dir. Bu durumdaki sözcük sayısı $21 \cdot 2520 = 52920$.

Tam bir A kullanılır. A dışındaki yedi harften altısı seçilir: $\binom{7}{6} = 7$ biçimde. Yedi harfin hepsi farklıdır: $7! = 5040$ sıralanış. Bu durumdaki sözcük sayısı $7 \cdot 5040 = 35280$.

Hiç A kullanılmaz. A dışındaki yedi harfin tümü alınır: $\binom{7}{7} = 1$ biçimde; $7! = 5040$ sıralanış. Bu durumdaki sözcük sayısı 5040.

Yedi harfli sözcüklerin sayısı

$$52920 + 35280 + 5040 = 93240$$

olur.

■

Çok terimli katsayının ikinci yorumu, farklı nesneleri belirli büyüklükte gruplara ayırmaktır: n nesneyi, birincisinde r_1 , ikincisinde r_2 , ... nesne olan k etiketli gruba bölmek yollarının sayısı da $\binom{n}{r_1, \dots, r_k}$ 'dir. Nesneleri sıraya dizip her birine grubunun etiketini yazınca bu yorum tekrarlı permütasyona dönüşür.

Örnek 5.3 (Sekiz Öğretmen, Dört Okul). (a) 8 öğretmen 4 okula kaç farklı biçimde yerleştirilebilir? Bir okula istenildiği kadar öğretmen verilebilir; bazı okullar boş kalabilir.

(b) Her okula tam 2 öğretmen düşecek biçimde kaç farklı dağıtım yapılabilir?

Çözüm.

Öğretmenler de okullar da birbirinden farklıdır; bir dağıtım, her öğretmenin hangi okula gittiğini söyleyen bir listedir.

(a) Birinci öğretmen için 4 okul seçeneği vardır; ikinci öğretmen için de, birincinin seçiminden bağımsız olarak, 4 seçenek; bu böyle sürer ve sekizinci öğretmen için de 4 seçenek vardır. Çarpma kuralı (Önerme 5.2):

$$4 \cdot 4 \cdots 4 = 4^8 = 65536.$$

(b) Öğretmenleri 1, ..., 8 diye numaralayalım. Bir dağıtım, i -inci yerine i -inci öğretmenin okulunun numarası yazılmış sekiz uzunluğunda bir dizidir; her okula tam iki öğretmen düşeceğinden dizide

1, 2, 3, 4 etiketlerinin her biri tam iki kez geçer. Böyle dizilerin sayısı, dördü ikiyeşerli dört çeşitten oluşan sekiz nesnenin tekrarlı permütasyon sayısıdır (Teorem 5.3):

$$\binom{8}{2, 2, 2, 2} = \frac{8!}{2! 2! 2! 2!} = \frac{40320}{16} = 2520.$$

Aynı sonuç adım adım kombinasyonla da bulunur: birinci okulun iki öğretmeni $\binom{8}{2} = 28$, ikincinin kalan altı kişiden $\binom{6}{2} = 15$, üçüncünün $\binom{4}{2} = 6$, dördüncünün $\binom{2}{2} = 1$ biçimde seçilir; $28 \cdot 15 \cdot 6 \cdot 1 = 2520$.

■

5.5 Binom Teoremi

$\binom{n}{r}$ sayılarına binom katsayısı denmesinin nedeni, $(a + x)^n$ açılımının katsayıları olmalarıdır. Bu teorem sayma dünyasıyla cebir dünyasını birbirine bağlar ve ileride binom dağılımının olasılıklarının toplamının 1 olduğunu göstermek için gerekecektir.

Teorem 5.4 (Binom Teoremi). $n = 0, 1, 2, \dots$ ve $a, x \in \mathbb{R}$ (ya da $\mathbb{C}$) için

$$(a + x)^n = \sum_{r=0}^n \binom{n}{r} a^{n-r} x^r.$$

İspat.

n üzerinde tümevarım yapalım. $n = 0$ için iki taraf da 1'dir; $n = 1$ için sağ taraf $\binom{1}{0}a + \binom{1}{1}x = a + x$ olur.

Eşitlik bir $n \geq 1$ için doğru olsun. O zaman

$$(a + x)^{n+1} = (a + x)(a + x)^n = \sum_{r=0}^n \binom{n}{r} a^{n+1-r} x^r + \sum_{r=0}^n \binom{n}{r} a^{n-r} x^{r+1}.$$

İkinci toplamda $s = r + 1$ yazalım:

$$\sum_{r=0}^n \binom{n}{r} a^{n-r} x^{r+1} = \sum_{s=1}^{n+1} \binom{n}{s-1} a^{n+1-s} x^s.$$

Birinci toplamın $r = 0$ terimi a^{n+1} , ikinci toplamın $s = n + 1$ terimi x^{n+1} 'dir; geri kalan terimler $1 \leq r \leq n$ için birleştirilir:

$$(a + x)^{n+1} = a^{n+1} + \sum_{r=1}^n \left[\binom{n}{r} + \binom{n}{r-1} \right] a^{n+1-r} x^r + x^{n+1}.$$

Pascal özdeşliği (Lemma 5.1) köşeli parantezi $\binom{n+1}{r}$ yapar. $\binom{n+1}{0} = \binom{n+1}{n+1} = 1$ olduğundan uç terimler de aynı biçime girer:

$$(a + x)^{n+1} = \sum_{r=0}^{n+1} \binom{n+1}{r} a^{n+1-r} x^r.$$

Bu, $n + 1$ için istenen eşitliktir.

Aynı sonuç bir sayma yorumuyla da görülür. $(a + x)^n = (a + x)(a + x) \cdots (a + x)$ çarpımı açıldığında her terim, n çarpanın her birinden a ya da x seçilerek oluşur. x 'in tam r çarpandan seçildiği terimlerin her biri $a^{n-r} x^r$ 'ye eşittir ve hangi r çarpandan seçileceği $\binom{n}{r}$ biçimde belirlenir; katsayının $\binom{n}{r}$ olmasının nedeni budur.

■

i Binom teoreminin iki sonucu

$a = x = 1$ alınırsa $\sum_{r=0}^n \binom{n}{r} = 2^n$ çıkar: n elemanlı bir kümenin bütün alt kümelerinin sayısı 2^n 'dir; bu, $|\mathcal{P}(\Omega)| = 2^{|\Omega|}$ eşitliğinin (Önerme 1.2) sayma ispatıdır. $a = 1, x = -1$ alınırsa $\sum_{r=0}^n (-1)^r \binom{n}{r} = 0$ çıkar ($n \geq 1$): eleman sayısı çift olan alt kümelerin sayısı, eleman sayısı tek olan alt kümelerin sayısına eşittir; her ikisi de 2^{n-1} 'dir.

Binom katsayısının tanımındaki n yerine herhangi bir gerçel sayı konabilir; formüldeki çarpım yine anlamlıdır.

Tanım 5.3 (Genelleştirilmiş Binom Katsayısı). $\alpha \in \mathbb{R}$ ve $k = 1, 2, 3, \dots$ için

$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-k+1)}{k!}, \quad \binom{\alpha}{0} = 1$$

sayısına **genelleştirilmiş binom katsayısı** denir. $\alpha = n$ negatif olmayan bir tam sayı olduğunda $k \leq n$ için bu sayı $\binom{n}{k}$ ile çakışır; $k > n$ için çarpanlardan biri ($\alpha - n = 0$) sıfır olduğundan $\binom{n}{k} = 0$ olur.

i Genelleştirilmiş binom serisi

$\alpha \in \mathbb{R}$ ve $|z| < 1$ olan her $z \in \mathbb{C}$ için

$$(1+z)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} z^k$$

dir. Bu, $(1+z)^\alpha$ fonksiyonunun Taylor açılımıdır; ispatı analizin konusudur ve burada yalnızca ifadesi kullanılacaktır. $|z| = 1$ üzerinde serinin yakınsaması α 'ya bağlıdır; $|z| < 1$ için her zaman yakınsar. $\alpha = n$ negatif olmayan bir tam sayı olduğunda $k > n$ terimleri sıfırdır ve seri binom teoremine indirgenir. $\alpha = -1$ için $\binom{-1}{k} = (-1)^k$ olur ve seri geometrik seriye, $\frac{1}{1+z} = \sum_{k=0}^{\infty} (-1)^k z^k$ eşitliğine döner. Genel olarak $n \in \mathbb{N}$ için

$$\binom{-n}{k} = (-1)^k \binom{n+k-1}{k}$$

dir (alıştırmalarda ispatlanacaktır); bu özdeşlik negatif binom dağılımının adının kaynağıdır.

5.6 Tekrarlı Kombinasyon: Yıldız–Çubuk Yöntemi

Kombinasyonda her nesne en çok bir kez seçilir. Ama bir pastaneden üç çeşit kurabiyeden yedi tane alırken aynı çeşitten birkaç tane alabiliriz; sıra önemsizdir, tekrar serbesttir. Aynı soru başka kılıklarda da karşımıza çıkar: $x + y + z = 7$ denkleminin negatif olmayan tam sayı çözümlerini saymak, ya da 7 özdeş nesneyi 3 farklı kutuya dağıtmak. Üç sorunun yanıtı aynıdır ve zarif bir kodlamayla bulunur.

Teorem 5.5 (Tekrarlı Kombinasyon (Yıldız–Çubuk)). $n \geq 1$ ve $r \geq 0$ tam sayılar olsun.

(a) $x_1 + x_2 + \cdots + x_n = r$ denkleminin negatif olmayan tam sayılardaki ($x_i \in \{0, 1, 2, \dots\}$) çözümlerinin sayısı

$$\binom{n+r-1}{r} = \binom{n+r-1}{n-1}$$

dir. Bu sayı aynı zamanda n çeşit nesneden tekrara izin vererek r nesne seçmenin ve r özdeş nesneyi n farklı kutuya dağıtmanın yollarının sayısıdır.

(b) $r \geq n$ ise aynı denklemin pozitif tam sayılardaki ($x_i \geq 1$) çözümlerinin sayısı $\binom{r-1}{n-1}$ 'dir.

İspat.

(a) Bir $(x_1, \dots, x_n)$ çözümünü şöyle kodlayalım: x_1 tane yıldız, bir çubuk, x_2 tane yıldız, bir çubuk, ..., bir çubuk, x_n tane yıldız. Örneğin $n = 3, r = 7$ için $(2, 4, 1)$ çözümü

★★ | ★★★★★ | ★

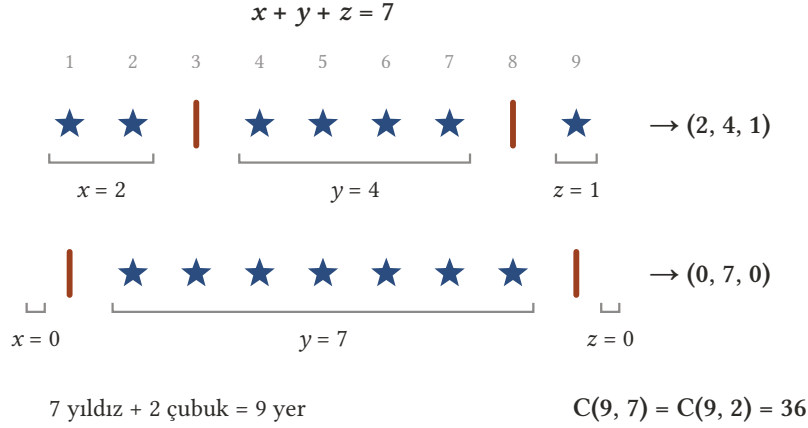
dizisine, $(0, 7, 0)$ çözümü | ★★★★★★★ | dizisine karşılık gelir. Dizide toplam r yıldız ve $n - 1$ çubuk, yani $n + r - 1$ simge vardır. Tersine, r yıldız ile $n - 1$ çubuktan oluşan her dizi tam bir çözüm belirler: çubuklar diziyi n parçaya böler ve i -inci parçadaki yıldız sayısı x_i 'dir; boş parça $x_i = 0$ demektir. Öyleyse çözümler ile bu diziler arasında birebir ve örten bir eşleme vardır. Böyle bir dizi, $n + r - 1$ yerden yıldızların konacağı r yeri seçmekle belirlenir; Teorem 5.2 ile $\binom{n+r-1}{r}$ biçimde. Çubukların yerlerini seçmek de aynı şeydir: $\binom{n+r-1}{n-1}$.

Öteki iki yorum doğrudan eşleşir: n çeşit nesneden tekrarlı r nesne seçmek, her çeşitten kaç tane alındığını (x_i) söylemektir; r özdeş nesneyi n kutuya dağıtmak, her kutuya kaç nesne düştüğünü söylemektir. İkisi de $x_1 + \dots + x_n = r$ denkleminin negatif olmayan çözümleriyle birebir eşleşir.

(b) $x_i \geq 1$ koşulunu $y_i = x_i - 1 \geq 0$ dönüşümüyle kaldıralım. $(x_1, \dots, x_n)$ pozitif bir çözümse $(y_1, \dots, y_n)$, $y_1 + \dots + y_n = r - n$ denkleminin negatif olmayan bir çözümüdür; tersine her böyle (y_i) dizisi $x_i = y_i + 1$ ile pozitif bir çözüm verir. Eşleme birebir ve örtendir. (a) şikkıyla sayı

$$\binom{n + (r - n) - 1}{n - 1} = \binom{r - 1}{n - 1}$$

olur.



Şekil 5.3: Yıldız-çubuk kodlaması. $x + y + z = 7$ denkleminin her çözümü, 7 yıldız ile 2 çubuktan oluşan 9 simgelik bir diziye karşılık gelir: çubuklar diziyi üç parçaya böler, her parçadaki yıldız sayısı bir bilinmeyen değeridir; boş parça 0 demektir. Dizi, 9 yerden yıldızların (ya da çubukların) yerini seçmekle belirlenir, dolayısıyla çözüm sayısı $C(9, 7) = C(9, 2) = 36$ 'dır.

■

Örnek 5.4 (Bir Denklemin Çözüm Sayısı). $x, y, z \in \{0, 1, 2, \dots\}$ olmak üzere $x + y + z = 7$ denkleminin kaç farklı çözümü vardır? Çözümler pozitif tam sayılarla sınırlanırsa sayı ne olur?

Çözüm.

$n = 3$ bilinmeyen, toplam $r = 7$. Teorem 5.5 (a) ile

$$\binom{3 + 7 - 1}{7} = \binom{9}{7} = \binom{9}{2} = \frac{9 \cdot 8}{2} = 36.$$

Doğrudan sayarak doğrulayalım: $x = k$ sabitlenince $y + z = 7 - k$ denkleminin $8 - k$ çözümü vardır ($y = 0, 1, \dots, 7 - k$ ve z buna göre belirlenir). $k = 0, 1, \dots, 7$ üzerinden toplam $8 + 7 + \dots + 1 = 36$.

Pozitif çözümler için (b) şikkı: $\binom{7-1}{3-1} = \binom{6}{2} = 15$.

■

Örnek 5.5 (Sekiz Özdeş Beyaz Tahta). (a) 8 özdeş beyaz tahta 4 okul arasında kaç farklı biçimde paylaştırılabilir? Bir okul hiç tahta almayabilir.
(b) Her okul en az bir tahta alacaksa kaç paylaştırma vardır?

Çözüm.

Tahtalar özdeş olduğundan bir paylaştırma yalnızca her okula kaç tahta düştüğüyle, yani $x_1 + x_2 + x_3 + x_4 = 8$ denkleminin bir çözümüyle belirlenir. Bu, öğretmen örneğinden (Örnek 5.3) temel bir farktır: orada öğretmenler birbirinden farklıydı ve kimin nereye gittiği önemliydi.

(a) Negatif olmayan çözümler (Teorem 5.5):

$$\binom{4+8-1}{8} = \binom{11}{8} = \binom{11}{3} = \frac{11 \cdot 10 \cdot 9}{6} = 165.$$

(b) Pozitif çözümler: $\binom{8-1}{4-1} = \binom{7}{3} = 35$. Aynı sonuç şöyle de görülür: önce her okula birer tahta verilir, kalan 4 özdeş tahta 4 okula serbestçe dağıtılır: $\binom{4+4-1}{4} = \binom{7}{4} = 35$.

Hangi formül?

Bir sayma sorusunda iki şey sorulur: **sıra önemli mi** ve **tekrar var mı**. n farklı nesneden r tanesi için yanıtlar şöyle özetlenir:

	sıra önemli	sıra önemsiz
tekrarsız	$P(n, r) = \frac{n!}{(n-r)!}$	$\binom{n}{r}$
tekrarlı	n^r	$\binom{n+r-1}{r}$

Sol alt hücre çarpma kuralının doğrudan sonucudur: r yerin her biri için n seçenek vardır (öğretmen örneğindeki 4^8 gibi). Nesnelerin bir kısmı özdeşse ve tümü sıralanıyorsa tekrarlı permütasyon (Teorem 5.3) devreye girer. Üçüncü bir soru, **nesnelerin ayırt edilip edilmediği**; aşağıdaki iki örnek, bu soruya verilen yanıtın sonucu nasıl değiştirdiğini gösterir.

	n farklı nesneden r tanesini seçmek	
	tekrarlı (iadeli)	tekrarsız (iadesiz)
sıra önemli (sıralı)	n^r <i>çarpma kuralı</i> $n = 3, r = 2$: 9 seçim <i>aa ab ac</i> <i>ba bb bc</i> <i>ca cb cc</i>	$P(n, r) = n!/(n-r)!$ <i>permütasyon</i> $n = 3, r = 2$: 6 seçim <i>ab ac ba</i> <i>bc ca cb</i>
sıra önemsiz (sırasız)	$C(n+r-1, r)$ <i>tekrarlı kombinasyon</i> $n = 3, r = 2$: 6 seçim <i>aa ab ac</i> <i>bb bc cc</i>	$C(n, r) = n!/(r!(n-r)!)$ <i>kombinasyon</i> $n = 3, r = 2$: 3 seçim <i>ab ac bc</i>

Şekil 5.4: n farklı nesneden r tanesini seçmenin dört yolu. Sütunlar tekrara (iadeye) izin verilip verilmediğini, satırlar sıranın önemli olup olmadığını ayırır; her hücrede sayma formülü ve a, b, c nesnelerinden ikisinin seçilmesi örneği ($n = 3, r = 2$) vardır. Sıralı seçimler 9 ve 6, sırasız seçimler 6 ve 3 tanedir; sıra önemsizleşince ab ile ba aynı seçim sayılır, tekrar yasaklanınca aa gibi seçimler düşer.

5.7 Karma Sayma Problemleri

Gerçek problemlerde kurallar tek başına değil, birlikte kullanılır: durumlar toplama kuralıyla ayrılır, her durum çarpma kuralı ve kombinasyon formülleriyle sayılır. Ayrıca sorunun hangi düzenlemeleri “farklı” saydığı, yani varsayımlar, sayımdan önce açıkça belirlenmelidir.

Örnek 5.6 (Yedi Siyah ve Yedi Beyaz Top). 7 siyah ve 7 beyaz top, her kutuda ikişer top olacak biçimde 7 kutuya kaç farklı biçimde dağıtılabilir? Kutular birbirinden farklı (numaralı), aynı renkli toplar özdeş kabul edilecektir.

Çözüm.

Varsayımlar sayımı belirler: kutular numaralı olduğundan hangi kutuda ne olduğu önemlidir; aynı renkli toplar özdeş olduğundan bir kutunun içeriği yalnızca renk bileşimiyle belirlenir. Her kutuda iki top olacağından bir kutunun içeriği üç türden biridir: iki siyah (SS), bir siyah bir beyaz (SB), iki beyaz (BB).

Bir dağıtım, i -inci kutudaki siyah top sayısı $s_i \in \{0, 1, 2\}$ olmak üzere $(s_1, \dots, s_7)$ dizisiyle belirlenir; kutudaki beyaz sayısı $2 - s_i$ olarak kendiliğinden bulunur. Toplam siyah sayısı 7 olmalıdır: $s_1 + \dots + s_7 = 7$. Bu durumda toplam beyaz sayısı $14 - 7 = 7$ olur; ikinci koşul kendiliğinden sağlanır.

SS kutularının sayısı a , SB kutularının sayısı b , BB kutularının sayısı c olsun. O zaman

$$a + b + c = 7, \quad 2a + b = 7.$$

İkinci denklemden $b = 7 - 2a$, birinciden $c = 7 - a - b = a$ çıkar. $b \geq 0$ olması için $a \leq 3$ gerekir; olası (a, b, c) üçlüleri

$$(0, 7, 0), \quad (1, 5, 1), \quad (2, 3, 2), \quad (3, 1, 3)$$

dir. Her üçlü için hangi kutuların SS, hangilerinin SB, hangilerinin BB olacağını seçmek gerekir; bu, 7 kutunun a, b, c tane olacak biçimde üç çeşide ayrılmasıdır ve Teorem 5.3 ile $\frac{7!}{a!b!c!}$ biçimde yapılır:

(a, b, c)	dağıtım sayısı
$(0, 7, 0)$	$\frac{7!}{0!7!0!} = 1$
$(1, 5, 1)$	$\frac{7!}{1!5!1!} = 42$
$(2, 3, 2)$	$\frac{7!}{2!3!2!} = 210$
$(3, 1, 3)$	$\frac{7!}{3!1!3!} = 140$

Dört durum birbirini dışlar; toplama kuralıyla

$$1 + 42 + 210 + 140 = 393$$

farklı dağıtım vardır.

Varsayım değişirse sonuç değişir. 14 topun tümü birbirinden farklı (örneğin numaralı) sayılsaydı, 14 topu 7 numaralı kutuya ikişer ikişer dağıtmanın sayısı Teorem 5.3 ile $\frac{14!}{(2!)^7} = 681080400$ olurdu.

■

Örnek 5.7 (Safari Turu). 9 kişi, kapasiteleri 2, 4 ve 5 yolcu olan 3 araçla safari turuna çıkacaktır ve bütün araçlar kullanılacaktır. 9 kişi araçlara kaç farklı biçimde dağılabilir? Şu varsayımlarla çözünüz: kişiler ve araçlar birbirinden farklıdır; bir araç içindeki koltuklar ayırt edilmez, yalnızca kimin hangi araçta olduğu önemlidir; “bütün araçlar kullanılacak” demek her araçta en az bir kişi bulunacak demektir.

Çözüm.

Toplam kapasite $2 + 4 + 5 = 11$ olduğundan iki koltuk boş kalacaktır. Araçlardaki kişi sayıları (n_1, n_2, n_3) olsun;

$$n_1 + n_2 + n_3 = 9, \quad 1 \leq n_1 \leq 2, \quad 1 \leq n_2 \leq 4, \quad 1 \leq n_3 \leq 5.$$

$n_1 = 2$ ise $n_2 + n_3 = 7$ olmalıdır: $(n_2, n_3) \in \{(2, 5), (3, 4), (4, 3)\}$. $n_1 = 1$ ise $n_2 + n_3 = 8$: $(n_2, n_3) \in \{(3, 5), (4, 4)\}$. $n_1 = 0$ birinci aracı boş bırakacağından dışlanır. Beş dağılım biçimi vardır.

Belirli bir (n_1, n_2, n_3) için 9 kişiyi araçlara ayırmak, 9 kişiyi n_1, n_2, n_3 kişilik üç etiketli gruba bölmektir; **Teorem 5.3** ile $\frac{9!}{n_1! n_2! n_3!}$ biçimde yapılır:

(n_1, n_2, n_3)	yol sayısı
(2, 2, 5)	$\frac{9!}{2! 2! 5!} = 756$
(2, 3, 4)	$\frac{9!}{2! 3! 4!} = 1260$
(2, 4, 3)	$\frac{9!}{2! 4! 3!} = 1260$
(1, 3, 5)	$\frac{9!}{1! 3! 5!} = 504$
(1, 4, 4)	$\frac{9!}{1! 4! 4!} = 630$

Durumlar birbirini dışladığından toplama kuralıyla

$$756 + 1260 + 1260 + 504 + 630 = 4410$$

farklı yol vardır.

Varsayımların etkisi: “Her araçta en az bir kişi” koşulu kaldırılıyorsa $(0, 4, 5)$ durumu da eklenir ($\frac{9!}{4! 5!} = 126$) ve toplam 4536 olurdu. Koltuklar da ayırt edilseydi 11 farklı koltuğa 9 kişi $P(11, 9) = \frac{11!}{2!} = 19958400$ biçimde oturur; bunlardan birinci aracı boş bırakanlar (iki boş koltuğun ikisi de birinci araçta ve 9 kişi kalan 9 koltuğa $9! = 362880$ biçimde oturur) çıkarılınca 19595520 bulunurdu. Hangi sayının doğru olduğu, sorunun hangi düzenlemeleri farklı saydığına bağlıdır; bu yüzden varsayımlar baştan söylenmelidir.

■

5.8 Laplace Uzayında Olasılık Hesapları

Sayma araçları hazır olduğuna göre bölümün başındaki amaca dönebiliriz. Laplace uzayında (**Tanım 4.1**) bir olayın olasılığı, olayın örnek uzayın eleman sayılarının oranıdır (**Önerme 4.2**); iki sayı da artık bu bölümün formülleriyle bulunur. Dikkat edilecek tek nokta, örnek uzayı **bütün sonuçları eşit olasılıklı kılacak biçimde** seçmektir: iki zarın toplamını sayarken $\{2, \dots, 12\}$ kümesi eşit olasılıklı değildir, ama sıralı ikililerin kümesi eşit olasılıklıdır.

Örnek 5.8 (Sayma ile Olasılık Hesapları). (a) İki zar atılıyor. Üste gelen sayıların toplamının 7 olma olasılığı nedir?

(b) 10 kişilik bir topluluktan rastgele 5 kişilik bir komite seçiliyor. Topluluktaki belirli iki kişinin ikisinin de komitede olma olasılığı nedir? Yalnızca birinin komitede olma olasılığı? Hiçbirinin olmama olasılığı?

(c) (Doğum günü problemi) Bir odada n kişi vardır. Yılın 365 günü olduğunu ve n kişinin doğum günlerinden oluşan 365^n sıralı listenin her birinin eşit olasılıklı olduğunu varsayalım. En az iki kişinin doğum gününün aynı olma olasılığını bulunuz; $n = 23$ için hesaplayınız.

Çözüm.

Üç deneyde de sonuçlar sonlu ve eşit olasılıklıdır; Laplace formülü (**Önerme 4.2**) $P(A) = \frac{|A|}{|\Omega|}$ kullanılır.

(a) $\Omega = \{(i, j) : i, j \in \{1, \dots, 6\}\}$; çarpma kuralıyla $|\Omega| = 6 \cdot 6 = 36$ ve zarlar hilesizse her sıralı ikili eşit olasılıklıdır. Toplamı 7 yapan ikililer

$$(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)$$

olmak üzere altı tanedir. Öyleyse

$$P(\text{toplam } 7) = \frac{6}{36} = \frac{1}{6}.$$

(b) Ω , 10 kişinin 5 elemanlı alt kümelerinin kümesidir: $|\Omega| = \binom{10}{5} = 252$; rastgele seçim her alt kümeyi eşit olasılıklı kılar. Belirli iki kişiye A ve B diyelim.

İkisi de komitede. A ve B komitede olduğuna göre kalan 3 üye öteki 8 kişiden seçilir: $\binom{8}{3} = 56$ komite. Olasılık $\frac{56}{252} = \frac{2}{9}$.

Yalnızca biri komitede. A'nın olup B'nin olmadığı komitelerde A alınır, kalan 4 üye B dışındaki 8 kişiden seçilir: $\binom{8}{4} = 70$. B'nin olup A'nın olmadığı komiteler de 70 tanedir. İki durum ayrık olduğundan toplama kuralıyla 140 komite; olasılık $\frac{140}{252} = \frac{5}{9}$.

Hiçbiri komitede değil. 5 üyenin tümü öteki 8 kişiden seçilir: $\binom{8}{5} = 56$; olasılık $\frac{56}{252} = \frac{2}{9}$.

Sağlama: $\frac{2}{9} + \frac{5}{9} + \frac{2}{9} = 1$; üç olay Ω 'nın bir parçasıdır.

(c) Ω , n kişinin doğum günlerinin sıralı listelerinin kümesidir: $\Omega = \{1, \dots, 365\}^n$ ve çarpma kuralıyla $|\Omega| = 365^n$; varsayım gereği her liste eşit olasılıklıdır. A , en az iki kişinin doğum gününün aynı olması olayı olsun. Doğrudan saymak zordur; tümleyeni saymak kolaydır. A^c , bütün doğum günlerinin farklı olması olayıdır ve eleman sayısı 365 günden n tanesinin sıralı seçimidir: $P(365, n) = 365 \cdot 364 \cdots (365 - n + 1)$. **Teorem 3.1** ile

$$P(A) = 1 - P(A^c) = 1 - \frac{365 \cdot 364 \cdots (365 - n + 1)}{365^n} = 1 - \prod_{i=0}^{n-1} \left(1 - \frac{i}{365}\right).$$

$n > 365$ için $365 \cdot 364 \cdots (365 - n + 1)$ çarpımında $365 - 365 = 0$ çarpanı yer aldığından pay sıfırdır ve $P(A) = 1$ olur: 366 kişi arasında çakışma kesindir. $n = 23$ için çarpım 0,4927 verir; dolayısıyla

$$P(A) \approx 1 - 0,4927 = 0,5073.$$

Yani yalnızca 23 kişilik bir odada en az iki kişinin doğum gününün çakışması, çakışmamasından daha olasıdır. Birkaç değer daha: $n = 10$ için $P(A) \approx 0,117$; $n = 30$ için 0,706; $n = 50$ için 0,970.

■

▲ Eşit olasılıklı örnek uzay seçilmelidir

Laplace formülü ancak Ω 'nın her noktası eşit olasılıklıysa geçerlidir. (a) şıkında “toplam 7 olur ya da olmaz, iki sonuç, olasılık $\frac{1}{2}$ ” demek yanlıştır; iki sonuç eşit olasılıklı değildir. Aynı biçimde (b) şıkında komiteleri sıralı 5-liler olarak da sayabilirdik ($|\Omega| = P(10, 5) = 30240$); pay da sıralı sayıldığı sürece $(P(5, 2) \cdot P(8, 3))$ türünden) oran değişmez. Önemli olan, pay ile paydanın **aynı türden** nesneleri saymasıdır.

5.9 Alıştırmalar

Alıştırma 5.1 (Sayma Teknikleri). (a) 5 kişi bir sıraya kaç farklı biçimde oturabilir? Belirli iki kişinin yan yana oturduğu diziliş sayısı kaçtır?

(b) MATEMATİK sözcüğünün harflerinin tümü kullanılarak kaç farklı sözcük yazılabilir?

(c) $x_1 + x_2 + x_3 + x_4 = 10$ denkleminin negatif olmayan tam sayı çözümlerinin sayısını bulunuz. Her $x_i \geq 2$ koşulu altında çözüm sayısı kaçtır?

(d) $(2x - y)^5$ açılımında x^2y^3 teriminin katsayısını bulunuz.

(e) Bir torbada 5 kırmızı, 4 mavi top vardır; rastgele 3 top çekiliyor. Üçünün de kırmızı olma olasılığı ile tam ikisinin kırmızı olma olasılığını bulunuz.

(f) Üç zar atılıyor. Üç zarın da farklı sayı gösterme olasılığı nedir?

(g) $\binom{-2}{3}$ ve $\binom{1/2}{2}$ genelleştirilmiş binom katsayılarını hesaplayınız; $n \in \mathbb{N}$ için $\binom{-n}{k} = (-1)^k \binom{n+k-1}{k}$ eşitliğini gösteriniz.

Çözüm.

(a) $5! = 120$ diziliş vardır (**Teorem 5.1**). Yan yana oturacak iki kişiyi tek bir blok sayalım: 4 nesne (3 kişi ve blok) $4! = 24$ biçimde dizilir, blok içinde iki kişi $2! = 2$ biçimde sıralanır; çarpma kuralıyla $24 \cdot 2 = 48$.

(b) MATEMATİK dokuz harflidir: M, A, T ikişer kez; E, İ, K birer kez. **Teorem 5.3** ile $\frac{9!}{2!2!2!} = \frac{362880}{8} = 45360$.

(c) **Teorem 5.5** (a) ile $\binom{4+10-1}{10} = \binom{13}{3} = 286$. $x_i \geq 2$ koşulu için $y_i = x_i - 2 \geq 0$ konursa $y_1 + y_2 + y_3 + y_4 = 2$ denklemini elde edilir: $\binom{4+2-1}{2} = \binom{5}{2} = 10$.

(d) Binom teoremi (**Teorem 5.4**) ile $(2x - y)^5 = \sum_{r=0}^5 \binom{5}{r} (2x)^{5-r} (-y)^r$. x^2y^3 terimi $r = 3$ 'e karşılık gelir: $\binom{5}{3} \cdot 2^2 \cdot (-1)^3 = 10 \cdot 4 \cdot (-1) = -40$.

- (e) $|\Omega| = \binom{9}{3} = 84$. Üçü kırmızı: $\binom{5}{3} = 10$ çekiliş, olasılık $\frac{10}{84} = \frac{5}{42}$. Tam ikisi kırmızı: iki kırmızı $\binom{5}{2} = 10$, bir mavi $\binom{4}{1} = 4$ biçimde seçilir; çarpma kuralıyla 40 çekiliş, olasılık $\frac{40}{84} = \frac{10}{21}$.
- (f) $|\Omega| = 6^3 = 216$. Üçü farklı: birinci zar için 6, ikincisi için 5, üçüncüsü için 4 değer, $P(6, 3) = 120$; olasılık $\frac{120}{216} = \frac{5}{9}$.
- (g) $\binom{-2}{3} = \frac{(-2)(-3)(-4)}{3!} = \frac{-24}{6} = -4$ ve $\binom{1/2}{2} = \frac{\frac{1}{2}(-\frac{1}{2})}{2!} = -\frac{1}{8}$. Genel olarak paydaki k çarpanın her birinden -1 dışarı alınır:

$$\binom{-n}{k} = \frac{(-n)(-n-1)\cdots(-n-k+1)}{k!} = (-1)^k \frac{n(n+1)\cdots(n+k-1)}{k!} = (-1)^k \frac{(n+k-1)!}{k!(n-1)!} = (-1)^k \binom{n+k-1}{k}.$$

■

Sayma araçları elimizde olduğuna göre, bir olayın gerçekleştiği bilgisinin öteki olayların olasılığını nasıl değiştirdiğine geçebiliriz: [Koşullu Olasılık](#), [Toplam Olasılık](#) ve [Bayes Teoremi](#).

6. Koşullu Olasılık, Toplam Olasılık ve Bayes Teoremi

Sayma Teknikleri bölümünde Laplace uzayında olasılık hesabını uygun durumların sayılmasına indirge-dik. Bu bölümde yeni bir soru soruyoruz: **deney hakkında kısmi bir bilgi edinildiğinde olasılıklar nasıl değişir?** İki zar atıldığını ve toplamın 4'ten büyük olduğunu öğrendiğimizi düşünelim. Bu bilgi, toplamın çift olma olasılığını değiştirir mi? Hesabın yeniden yapılması gerekir: toplamı 4 ya da daha küçük olan sonuçlar artık olanaksızdır; olasılık, geriye kalan sonuçlar arasında hesaplanmalıdır.

Bu “yeniden hesap” fikrini kesinleştiren kavram **koşullu olasılıktır**. Önce tanımı vereceğiz ve B olayı sabit tutulduğunda $A \mapsto P(A | B)$ eşlemesinin kendi başına bir olasılık ölçüsü olduğunu göstereceğiz; böylece olasılık ölçüleri için bildiğimiz her özellik koşullu olasılıklar için de geçerli olacak. Ardından tanımı tersine çevirip kesişimin olasılığını koşullu olasılıkların çarpımı olarak yazan **çarpma teoremini** ispatlayacağız.

Bölümün ikinci yarısında örnek uzayı parçalara ayırıp bir olayın olasılığını parçalar üzerinden toplayan **toplam olasılık formülünü** ve “sonuç gözlemlendiğinde nedenin olasılığı”nı veren **Bayes teoremini** ele alacağız. Bayes teoremi, olasılık teorisinin gündelik sezgiyle en çok çatıştığı yerlerden biridir; bunu bölümün sonundaki tıbbi test örneğinde açıkça göreceğiz.

6.1 Koşullu Olasılık

Önce fikri Laplace uzayında (**Tanım 4.1**) yakalayalım. Ω sonlu ve bütün sonuçlar eşit olasılıklı olsun. B olayının gerçekleştiğini öğrendiğimizde B 'nin dışındaki sonuçlar elenir; artık olabilecek sonuçların kümesi B 'dir. Bu bilgi altında A olayı, ancak $A \cap B$ 'deki bir sonuç gerçekleşmişse gerçekleşir. Öyleyse “ B biliniyorken A 'nın olasılığı” için doğal aday, uygun durumların yeni örnek uzaydaki payıdır:

$$\frac{n(A \cap B)}{n(B)} = \frac{n(A \cap B)/n(\Omega)}{n(B)/n(\Omega)} = \frac{P(A \cap B)}{P(B)}.$$

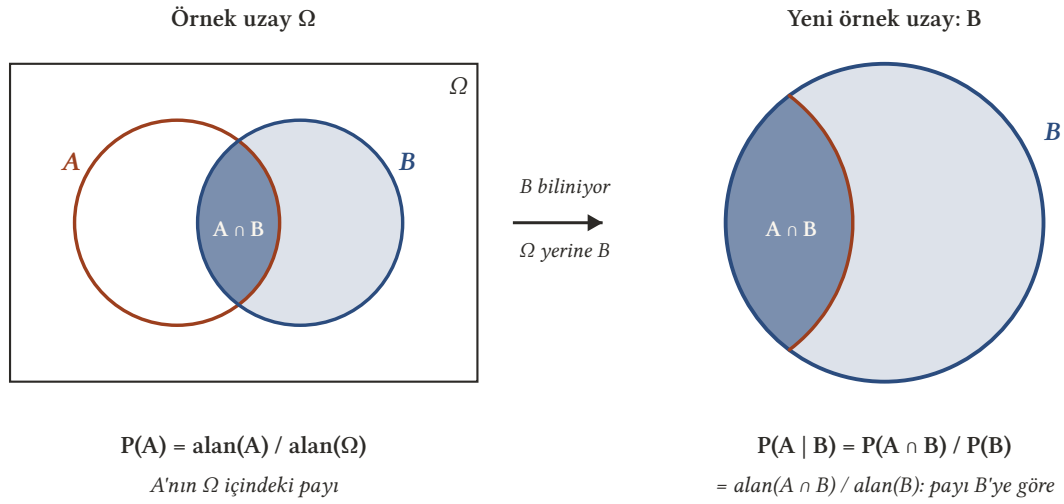
Sağdaki ifade yalnızca olasılıkları içerir; sayma yapılamayan uzaylarda da anlamlıdır. Genel tanımı bu ifadeyle veriyoruz.

Tanım 6.1 (Koşullu Olasılık). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı, $B \in \mathcal{U}$ ve $P(B) > 0$ olsun. $A \in \mathcal{U}$ olayı için

$$P(A | B) = \frac{P(A \cap B)}{P(B)}$$

sayısına A olayının B olayına göre koşullu olasılığı ya da B olayı biliniyorken A olayının gerçekleşme olasılığı denir.

Tanımda $P(B) > 0$ koşulu zorunludur; $P(B) = 0$ olan bir olaya göre koşullu olasılık tanımlanmaz. Gösterim olarak her zaman $P(A | B)$ yazacağız; eğik çizgili $P(A/B)$ yazımı bölme izlenimi verdiği için kullanılmaz. Dik çizginin ters bölüyle karıştırılmaması da ayrıca önemlidir.



Şekil 6.1: Koşullu olasılık, örnek uzayın daraltılmasıdır. Solda A ve B olayları Ω içinde durur; B açık, $A \cap B$ koyu renkle taranmıştır ve $P(A)$, A 'nın Ω içindeki payıdır. B 'nin gerçekleştiği öğrenilince Ω 'nın B dışındaki kısmı elenir ve sağda B yeni örnek uzay olur; A artık ancak $A \cap B$ üzerinden gerçekleşebilir. $P(A | B) = P(A \cap B) / P(B)$ bölmesi, olasılıkları B 'nin toplamı 1 olacak biçimde yeniden ölçekler.

⚠ Koşullu olasılık ile fark kümesinin olasılığı aynı şey değildir

$P(A | B)$ ifadesi **koşullu olasılığı**, $P(A \setminus B)$ ifadesi ise **fark kümesinin olasılığını** gösterir. Bunlar birbirinden bütünüyle farklı sayılardır:

$$P(A | B) = \frac{P(A \cap B)}{P(B)}, \quad P(A \setminus B) = P(A \cap B^c) = P(A) - P(A \cap B).$$

Düzgün bir zar atılsın; $A = \{1, 2, 3\}$ ve $B = \{2, 3, 4\}$ olsun. $A \cap B = \{2, 3\}$ olduğundan

$$P(A | B) = \frac{2/6}{3/6} = \frac{2}{3}, \quad P(A \setminus B) = P(\{1\}) = \frac{1}{6}$$

bulunur. İlki “zar 2, 3 ya da 4 geldiyse 3 ya da daha küçük olma olasılığı”dır; ikincisi “zarın A 'da olup B 'de olmama olasılığı”dır. Dik çizgi $|$ ile ters bölü $\setminus$ birbirine karıştırılmamalıdır.

Tanımdan hemen çıkan birkaç gözlem:

- Tanım, $P(A \cap B) = P(B) P(A | B)$ biçiminde de yazılabilir. Bu yazım, kesişimin olasılığını hesaplamının en sık kullanılan yoludur ve birazdan çarpma teoremine genelleşecektir.
- $A \cap B = \emptyset$ ise $P(A | B) = 0$ 'dır: B gerçekleştiyse B ile ayrık bir olay gerçekleşemez.
- $B \subseteq A$ ise $A \cap B = B$ olduğundan $P(A | B) = 1$ 'dir; özel olarak $P(\Omega | B) = P(B | B) = 1$.
- Laplace uzayında **Önerme 4.2** gereği $P(A | B) = \frac{n(A \cap B)}{n(B)}$ 'dir; yani koşullu olasılık, örnek uzayı B 'ye daraltıp yeniden saymaktan ibarettir.

6.1.1 Koşullu Olasılık Bir Olasılık Ölçüsüdür

B sabit tutulduğunda $A \mapsto P(A | B)$ eşlemesi, $\mathcal{U}$ üzerinde tanımlı yeni bir fonksiyondur. Bu fonksiyonun bir olasılık ölçüsü olduğunu göstermek istiyoruz. Bunu neden istiyoruz? Çünkü olasılık ölçüleri için ispatladığımız her sonuç (tümleyen kuralı, toplama kuralı, Boole eşitsizliği, süreklilik) o zaman koşullu olasılıklar için de yeniden ispat gerektirmeden geçerli olur.

Teorem 6.1 (Koşullu Olasılık Bir Olasılık Ölçüsüdür). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı, $B \in \mathcal{U}$ ve $P(B) > 0$ olsun.

$$P(\cdot | B) : \mathcal{U} \rightarrow \mathbb{R}, \quad A \mapsto P(A | B) = \frac{P(A \cap B)}{P(B)}$$

fonksiyonu $\mathcal{U}$ σ -cebiri üzerinde bir olasılık ölçüsüdür.

İspat.

Tanım 3.1'ndaki üç aksiyomu doğrulayacağız. Önce fonksiyonun iyi tanımlı olduğuna dikkat edelim: $A, B \in \mathcal{U}$ ise $\mathcal{U}$ kesişime kapalı olduğundan (**Teorem 2.1**) $A \cap B \in \mathcal{U}$ 'dur; dolayısıyla $P(A \cap B)$ tanımlıdır ve $P(B) > 0$ olduğundan bölme yapılabilir.

(A1) Negatif olmama. Her $A \in \mathcal{U}$ için $P(A \cap B) \geq 0$ ve $P(B) > 0$ olduğundan

$$P(A | B) = \frac{P(A \cap B)}{P(B)} \geq 0$$

olur.

(A2) Normlama. $\Omega \cap B = B$ olduğundan

$$P(\Omega | B) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$$

bulunur.

(A3) Sayılabilir toplamsallık. $A_1, A_2, \dots \in \mathcal{U}$ ikiye ikiye ayrık olaylar olsun. Her n için $A_n \cap B \in \mathcal{U}$ 'dur ve bu kümeler de ikiye ikiye ayrıktır: $n \neq m$ için

$$(A_n \cap B) \cap (A_m \cap B) \subseteq A_n \cap A_m = \emptyset.$$

Kesişimin birleşim üzerine dağılma özelliğinden

$$\left(\bigcup_{n=1}^{\infty} A_n \right) \cap B = \bigcup_{n=1}^{\infty} (A_n \cap B)$$

yazılır. Şimdi P için (A3) uygulanabilir:

$$\begin{aligned} P\left(\bigcup_{n=1}^{\infty} A_n \mid B\right) &= \frac{P\left(\left(\bigcup_{n=1}^{\infty} A_n\right) \cap B\right)}{P(B)} = \frac{P\left(\bigcup_{n=1}^{\infty} (A_n \cap B)\right)}{P(B)} \\ &= \frac{\sum_{n=1}^{\infty} P(A_n \cap B)}{P(B)} = \sum_{n=1}^{\infty} \frac{P(A_n \cap B)}{P(B)} = \sum_{n=1}^{\infty} P(A_n | B). \end{aligned}$$

Son adımda yakınsak bir seriyi sabit bir sayıya bölmenin terim terim yapılabilirdiğini kullandık. Üç aksiyom da sağlandığından $P(\cdot | B)$ bir olasılık ölçüsüdür.

■

i Olasılık ölçüsünün her özelliği koşullu olasılığa aktarılır

Teorem 6.1 sayesinde **Teorem 3.1**, **Teorem 3.2**, **Teorem 3.4** ve **Teorem 3.5**, P yerine $P(\cdot | B)$ konularak aynen kullanılabilir. Örneğin $P(B) > 0$ olan her B ve her $A, C \in \mathcal{U}$ için

$$P(A^c | B) = 1 - P(A | B), \quad P(A \cup C | B) = P(A | B) + P(C | B) - P(A \cap C | B)$$

ve $A \subseteq C$ ise $P(A | B) \leq P(C | B)$ olur. Bunların hiçbirini yeniden ispatlamak gerekmez.

6.1.2 Koşullu Olasılık Uzayı

Koşullu olasılığa ikinci bir gözle de bakılabilir: örnek uzayı Ω 'dan B 'ye daraltmak. Bunun için B 'nin alt kümelerinden oluşan bir σ -cebiri ve bu σ -cebir üzerinde bir olasılık ölçüsü gerekir. Doğal adaylar

$$\mathcal{U}_B = \{A \cap B : A \in \mathcal{U}\}$$

sınıfı ile $C \in \mathcal{U}_B$ için $P_B(C) = P(C)/P(B)$ ölçüsüdür. $\mathcal{U}_B$, $\mathcal{U}$ 'daki olayların B 'ye "izleri"nden oluşur; $\mathcal{U}$ 'da olup B 'nin içinde kalan kümelerin sınıfına eşittir. Gerçekten $C \in \mathcal{U}$ ve $C \subseteq B$ ise $C = C \cap B \in \mathcal{U}_B$; tersine $A \cap B$ hem $\mathcal{U}$ 'dadır hem de B 'nin alt kümesidir.

Teorem 6.2 (Koşullu Olasılık Uzayı). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı, $B \in \mathcal{U}$ ve $P(B) > 0$ olsun. $\mathcal{U}_B = \{A \cap B : A \in \mathcal{U}\}$ ve

$$P_B : \mathcal{U}_B \rightarrow \mathbb{R}, \quad C \mapsto P_B(C) = \frac{P(C)}{P(B)}$$

olmak üzere $(B, \mathcal{U}_B, P_B)$ bir olasılık uzayıdır. Ayrıca her $A \in \mathcal{U}$ için $P_B(A \cap B) = P(A | B)$ 'dir.

İspat.

Tanım 3.2'na göre üç şey gösterilmelidir: $B \neq \emptyset$, $\mathcal{U}_B$ 'nin B üzerinde bir σ -cebiri olduğu ve P_B 'nin $\mathcal{U}_B$ üzerinde bir olasılık ölçüsü olduğu. Son adımda da teoremin ikinci iddiasını doğrulayacağız.

Adım 1: $B \neq \emptyset$. **Teorem 3.1** gereği $P(\emptyset) = 0$ 'dır; oysa $P(B) > 0$. Öyleyse $B \neq \emptyset$.

Adım 2: $\mathcal{U}_B$ bir σ -cebirdir. $\mathcal{U}_B$, $\mathcal{U}$ 'nun B üzerindeki iz σ -cebidir; bunun bir σ -cebir olduğu **Önerme 2.4'**nde gösterilmişti. Kısaca hatırlatalım. $B = \Omega \cap B$ ve $\Omega \in \mathcal{U}$ olduğundan $B \in \mathcal{U}_B$ 'dir. $C = A \cap B \in \mathcal{U}_B$ ise C 'nin B 'ye göre tümleyeni

$$B \setminus C = B \cap (A \cap B)^c = B \cap (A^c \cup B^c) = A^c \cap B$$

olur ve $A^c \in \mathcal{U}$ olduğundan $B \setminus C \in \mathcal{U}_B$ 'dir. Son olarak $C_n = A_n \cap B \in \mathcal{U}_B$ ($n = 1, 2, \dots$) ise

$$\bigcup_{n=1}^{\infty} C_n = \left(\bigcup_{n=1}^{\infty} A_n \right) \cap B$$

ve $\bigcup_{n=1}^{\infty} A_n \in \mathcal{U}$ olduğundan $\bigcup_{n=1}^{\infty} C_n \in \mathcal{U}_B$ 'dir. Böylece **Tanım 2.2'**in üç koşulu sağlanır.

Adım 3: P_B bir olasılık ölçüsüdür. Önce P_B iyi tanımlıdır: $C = A \cap B \in \mathcal{U}_B$ ise $C \in \mathcal{U}$ 'dur, dolayısıyla $P(C)$ tanımlıdır. Şimdi aksiyomlar:

(i) Her $C \in \mathcal{U}_B$ için $P(C) \geq 0$ ve $P(B) > 0$ olduğundan $P_B(C) = P(C)/P(B) \geq 0$.

(ii) B , bu uzayın örnek uzayıdır ve $P_B(B) = P(B)/P(B) = 1$.

(iii) $C_1, C_2, \dots \in \mathcal{U}_B$ ikiye ikiye ayrık olsun. Adım 2 gereği $\bigcup_{n=1}^{\infty} C_n \in \mathcal{U}_B$ 'dir. Her C_n aynı zamanda $\mathcal{U}$ 'da olduğundan P için (A3) uygulanır:

$$P_B\left(\bigcup_{n=1}^{\infty} C_n\right) = \frac{P\left(\bigcup_{n=1}^{\infty} C_n\right)}{P(B)} = \frac{\sum_{n=1}^{\infty} P(C_n)}{P(B)} = \sum_{n=1}^{\infty} \frac{P(C_n)}{P(B)} = \sum_{n=1}^{\infty} P_B(C_n).$$

Demek ki $(B, \mathcal{U}_B, P_B)$ bir olasılık uzayıdır.

Adım 4: Koşullu olasılıkla bağ. $A \in \mathcal{U}$ için $A \cap B \in \mathcal{U}_B$ ve

$$P_B(A \cap B) = \frac{P(A \cap B)}{P(B)} = P(A | B)$$

olur.

■

İki bakış açısı aynı bilgiyi taşır. $P(\cdot | B)$ ölçüsü Ω üzerinde tanımlıdır ve B 'nin dışına sıfır olasılık verir; P_B ise doğrudan B üzerinde tanımlıdır. " B olayı gerçekleşti" bilgisi, olasılık uzayını B 'ye daraltmak ve olasılıkları $P(B)$ ile bölerek yeniden 1'e ölçeklemek demektir.

6.1.3 Örnekler

Örnek 6.1 (Toplamı Dörtten Büyük Olan İki Zar). İki düzgün zar atılıyor. Üste gelen sayıların toplamının 4'ten büyük olduğu biliniyorsa, toplamın çift sayı olması olasılığı nedir?

Çözüm.

Örnek uzay $\Omega = \{(x, y) : x, y \in \{1, 2, 3, 4, 5, 6\}\}$ ve $n(\Omega) = 36$ 'dır; zarlar düzgün olduğundan bu bir Laplace uzayıdır. A toplamın 4'ten büyük olması, B toplamın çift olması olayı olsun. Aranan olasılık $P(B | A)$ 'dir.

Önce A 'yı sayalım. Toplamı en çok 4 olan sonuçlar

$$A^c = \{(1, 1), (1, 2), (2, 1), (1, 3), (2, 2), (3, 1)\}$$

olup $n(A^c) = 6$; dolayısıyla $n(A) = 30$ ve $P(A) = \frac{30}{36} = \frac{5}{6}$ 'dır.

Şimdi $A \cap B$ 'yi, yani toplamı 4'ten büyük ve çift olan sonuçları sayalım. Olası toplamalar 6, 8, 10, 12'dir:

toplam	sonuçlar	sayı
6	(1, 5), (2, 4), (3, 3), (4, 2), (5, 1)	5
8	(2, 6), (3, 5), (4, 4), (5, 3), (6, 2)	5
10	(4, 6), (5, 5), (6, 4)	3
12	(6, 6)	1

Toplam $n(A \cap B) = 14$ ve $P(A \cap B) = \frac{14}{36}$. **Tanım 6.1** gereği

$$P(B | A) = \frac{P(A \cap B)}{P(A)} = \frac{14/36}{30/36} = \frac{14}{30} = \frac{7}{15}.$$

Aynı sonuç, örnek uzayı A 'ya daraltarak da görülür: A 'daki 30 sonucun 14'ü çift toplamlıdır, $\frac{14}{30} = \frac{7}{15}$. Sonucu koşulsuz olasılıkla karşılaştıralım: 36 sonucun 18'i çift toplamlı olduğundan $P(B) = \frac{1}{2}$ 'dir. Bilgi, olasılığı $\frac{1}{2}$ 'den $\frac{7}{15} \approx 0,467$ 'ye **düşürmüştür**. Nedeni şudur: elenen 6 sonucun 4'ü ((1, 1), (1, 3), (2, 2), (3, 1)) çift toplamlıydı; eleme, çift toplamı tek toplamıdan daha çok azaltmıştır. ■

Örnek 6.2 (Bölünebilme). 1'den 100'e kadar olan tam sayılardan biri rastgele seçiliyor. Seçilen sayının 2 ile bölünebildiği biliniyorsa, 3 veya 5 ile bölünebilmesi olasılığını bulunuz.

Çözüm.

$\Omega = \{1, 2, \dots, 100\}$ bir Laplace uzayıdır. A_2 , A_3 ve A_5 sırasıyla seçilen sayının 2, 3 ve 5 ile bölünebilmesi olayları olsun. Aranan olasılık $P(A_3 \cup A_5 | A_2)$ 'dir.

$A_2 = \{2, 4, \dots, 100\}$ olduğundan $n(A_2) = 50$ ve $P(A_2) = \frac{50}{100}$ 'dür. Kesişimin birleşim üzerine dağılmasıyla

$$(A_3 \cup A_5) \cap A_2 = (A_2 \cap A_3) \cup (A_2 \cap A_5)$$

yazılır; toplama kuralı (**Teorem 3.2**) ile

$$P((A_3 \cup A_5) \cap A_2) = P(A_2 \cap A_3) + P(A_2 \cap A_5) - P(A_2 \cap A_3 \cap A_5).$$

Buradaki olaylar sırasıyla 6 ile, 10 ile ve 30 ile bölünebilme olaylarıdır. 1 ile 100 arasında 6'nın $\lfloor 100/6 \rfloor = 16$, 10'un 10 ve 30'un 3 katı vardır. Öyleyse

$$P(A_3 \cup A_5 | A_2) = \frac{P((A_3 \cup A_5) \cap A_2)}{P(A_2)} = \frac{\frac{16}{100} + \frac{10}{100} - \frac{3}{100}}{\frac{50}{100}} = \frac{23}{50}.$$

Aynı hesabı **Teorem 6.1**'un gözüyle de yapabiliriz: $P(\cdot | A_2)$ bir olasılık ölçüsü olduğundan toplama kuralı ona doğrudan uygulanır,

$$P(A_3 \cup A_5 | A_2) = P(A_3 | A_2) + P(A_5 | A_2) - P(A_3 \cap A_5 | A_2) = \frac{16}{50} + \frac{10}{50} - \frac{3}{50} = \frac{23}{50}.$$

Burada $P(A_3 | A_2) = \frac{n(A_2 \cap A_3)}{n(A_2)} = \frac{16}{50}$ ve ötekiler de aynı biçimde hesaplanmıştır. ■

Örnek 6.3 (İki Dersten Başarısızlık). Bir sınıftaki öğrencilerin %45'i İstatistik, %35'i Bilgisayar ve %25'i hem İstatistik hem de Bilgisayar dersinde başarısızdır. Rastgele seçilen bir öğrencinin

- a) Bilgisayardan başarısız olduğu biliniyorsa İstatistikten de başarısız olma,
- b) İstatistikten başarısız olduğu biliniyorsa Bilgisayardan da başarısız olma,
- c) bu iki dersten en az birinden başarısız olma olasılığını bulunuz.

Çözüm.

I İstatistikten, B Bilgisayardan başarısız olma olayı olsun. Verilenler

$$P(I) = 0,45, \quad P(B) = 0,35, \quad P(I \cap B) = 0,25.$$

a) $P(I | B) = \frac{P(I \cap B)}{P(B)} = \frac{0,25}{0,35} = \frac{25}{35} = \frac{5}{7}.$

b) $P(B | I) = \frac{P(I \cap B)}{P(I)} = \frac{0,25}{0,45} = \frac{25}{45} = \frac{5}{9}.$

c) Toplama kuralı (Teorem 3.2) ile

$$P(I \cup B) = P(I) + P(B) - P(I \cap B) = 0,45 + 0,35 - 0,25 = 0,55.$$

(a) ile (b) şıklarının farklı çıktığına dikkat edelim: $P(I | B) \neq P(B | I)$. Koşullu olasılık, koşulun yönüne bağlıdır; pay her ikisinde de $P(I \cap B)$ 'dir ama payda koşul olayına göre değişir.

■

6.2 Çarpma Teoremi

Koşullu olasılığın tanımı $P(A \cap B) = P(B) P(A | B)$ biçiminde yazıldığında bir hesap aracına dönüşür. Art arda yapılan deneylerde koşullu olasılıklar çoğu zaman doğrudan bilinir: bir torbadan ilk top siyah çekildiyse ikincinin siyah çıkma olasılığı, torbanın yeni içeriğinden hemen okunur. Kesişimin olasılığına bu yoldan ulaşmak, sayma yapmaktan daha kolaydır. Üç olay için bu fikir

$$P(A_1 \cap A_2 \cap A_3) = P(A_1) P(A_2 | A_1) P(A_3 | A_1 \cap A_2)$$

eşitliğini verir; genel hâli aşağıdaki teoremdir.

Teorem 6.3 (Çarpma Teoremi). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı ve $A_1, A_2, \dots, A_n \in \mathcal{U}$ ($n \geq 2$) olayları

$$P(A_1 \cap A_2 \cap \dots \cap A_{n-1}) > 0$$

koşulunu sağlasın. O zaman

$$P\left(\bigcap_{i=1}^n A_i\right) = P(A_1) P(A_2 | A_1) P(A_3 | A_1 \cap A_2) \dots P(A_n | A_1 \cap A_2 \cap \dots \cap A_{n-1})$$

olur.

İspat.

Önce sağ taraftaki koşullu olasılıkların hepsinin tanımlı olduğunu görelim. $1 \leq k \leq n-1$ için

$$A_1 \cap \dots \cap A_k \supseteq A_1 \cap \dots \cap A_{n-1}$$

olduğundan, olasılığın monotonluğu (Teorem 3.1) gereği

$$P(A_1 \cap \dots \cap A_k) \geq P(A_1 \cap \dots \cap A_{n-1}) > 0$$

olur. Yani her koşul olayının olasılığı pozitiftir.

Şimdi n üzerinde tümevarım yapalım.

Temel adım ($n = 2$). $P(A_1) > 0$ olduğundan Tanım 6.1 gereği

$$P(A_2 | A_1) = \frac{P(A_1 \cap A_2)}{P(A_1)}, \quad \text{yani} \quad P(A_1 \cap A_2) = P(A_1) P(A_2 | A_1).$$

Tümevarım adımı. $n \geq 3$ olsun ve teoremin $n-1$ olay için doğru olduğunu varsayalım: $P(B_1 \cap \dots \cap B_{n-2}) > 0$ koşulunu sağlayan her $B_1, \dots, B_{n-1}$ için

$$P(B_1 \cap \dots \cap B_{n-1}) = P(B_1) P(B_2 | B_1) \dots P(B_{n-1} | B_1 \cap \dots \cap B_{n-2}).$$

$A_1, \dots, A_n$ olayları $P(A_1 \cap \dots \cap A_{n-1}) > 0$ koşulunu sağlasın. $C = A_1 \cap \dots \cap A_{n-1}$ diyelim; $P(C) > 0$ 'dır. Temel adımı C ve A_n olaylarına uygularsak

$$P\left(\bigcap_{i=1}^n A_i\right) = P(C \cap A_n) = P(C) P(A_n | C).$$

Öte yandan $A_1, \dots, A_{n-1}$ olayları $P(A_1 \cap \dots \cap A_{n-2}) \geq P(C) > 0$ koşulunu sağladığından tümevarım varsayımı bunlara uygulanır:

$$P(C) = P(A_1 \cap \dots \cap A_{n-1}) = P(A_1) P(A_2 | A_1) \dots P(A_{n-1} | A_1 \cap \dots \cap A_{n-2}).$$

Bu ifadeyi bir önceki eşitlikte yerine koyunca

$$P\left(\bigcap_{i=1}^n A_i\right) = P(A_1) P(A_2 | A_1) \dots P(A_{n-1} | A_1 \cap \dots \cap A_{n-2}) P(A_n | A_1 \cap \dots \cap A_{n-1})$$

elde edilir. Tümevarım ilkesi gereği eşitlik her $n \geq 2$ için doğrudur.

■

Teoremden olayların yazılış sırası önemli değildir: ilk $n - 1$ olayın kesişiminin olasılığı pozitif kaldığı sürece teorem $A_1, \dots, A_n$ 'nin herhangi bir sıralanışına uygulanabilir. Uygulamada bu sıra deneyin zaman sırasıdır: önce ilk çekiliş, ardından ilkin sonucu biliniyorken ikinci çekiliş, sonra üçüncüsü.

Örnek 6.4 (İadesiz Üç Siyah Top). Bir torbada 3 siyah ve 5 beyaz top vardır. İadesiz olarak art arda 3 top çekiliyor. Çekilen üç topun da siyah olması olasılığı nedir?

Çözüm.

S_i , i -inci çekilen topun siyah olması olayı olsun ($i = 1, 2, 3$). Aranan olasılık $P(S_1 \cap S_2 \cap S_3)$ 'tür. Koşullu olasılıklar torbanın içeriğinden okunur:

- İlk çekilişte 8 topun 3'ü siyahtır: $P(S_1) = \frac{3}{8}$.
- İlk top siyah çekildiyse torbada 2 siyah, 5 beyaz olmak üzere 7 top kalır: $P(S_2 | S_1) = \frac{2}{7}$.
- İlk iki top siyah çekildiyse torbada 1 siyah, 5 beyaz olmak üzere 6 top kalır: $P(S_3 | S_1 \cap S_2) = \frac{1}{6}$.

$P(S_1 \cap S_2) = \frac{3}{8} \cdot \frac{2}{7} = \frac{3}{28} > 0$ olduğundan Teorem 6.3 uygulanabilir:

$$P(S_1 \cap S_2 \cap S_3) = P(S_1) P(S_2 | S_1) P(S_3 | S_1 \cap S_2) = \frac{3}{8} \cdot \frac{2}{7} \cdot \frac{1}{6} = \frac{6}{336} = \frac{1}{56}.$$

Sonucu sayarak doğrulayalım. İadesiz üç çekilişte $8 \cdot 7 \cdot 6$ sıralı sonuç eşit olasılıklıdır ve her sırasız üçlü bunlardan $3! = 6$ tanesine karşılık gelir; öyleyse sırasız üçlüler de eşit olasılıklıdır. Sıra gözetilmeden 8 toptan 3'ünün seçilmesi $\binom{8}{3} = 56$ yolla olur (Teorem 5.2); üçünün de siyah olması için 3 siyah topun tamamı seçilmelidir, bu $\binom{3}{3} = 1$ yolla olur. Laplace formülü (Önerme 4.2) yine $\frac{1}{56}$ verir. Çarpma teoremi, sayma yapmadan aynı sonuca daha kısa yoldan ulaştırmıştır.

■

6.3 Sonlu Parçalanma ve Toplam Olasılık Formülü

Bir olayın olasılığını doğrudan hesaplamak çoğu zaman güçtür; ama deneyin hangi "durumda" olduğu bilindiğinde hesap kolaylaşır. Hangi torbanın seçildiği bilinirse beyaz top çekme olasılığı hemen yazılır; hangi paranın çekildiği bilinirse dört kez yazı gelme olasılığı hemen yazılır. Toplam olasılık formülü, bu koşullu olasılıkları durumların olasılıklarıyla ağırlıklandırıp toplar. Önce "durumlar"ın ne olduğunu kesinleştirelim.

Tanım 6.2 (Örnek Uzayın Sonlu Parçalanması). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı ve $A_1, A_2, \dots, A_n \in \mathcal{U}$ olsun. Eğer

- her $i = 1, 2, \dots, n$ için $P(A_i) > 0$,
- $1 \leq i < j \leq n$ için $A_i \cap A_j = \emptyset$ (olaylar ikiye ikiye ayrık),

(c) $A_1 \cup A_2 \cup \dots \cup A_n = \Omega$

ise $A_1, A_2, \dots, A_n$ olaylarına Ω örnek uzayının bir **sonlu parçalanması** denir.

Parçalanmanın her $\omega \in \Omega$ sonucunu tam olarak bir A_i 'ye koyduğuna dikkat edelim: (c) gereği en az birine, (b) gereği en çok birine düşer. Ayrıca sonlu toplamsallık (Teorem 3.1) gereği

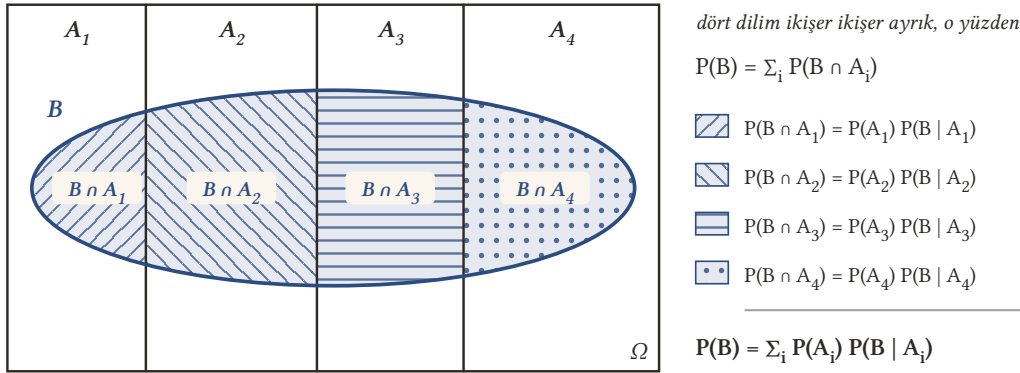
$$\sum_{i=1}^n P(A_i) = P\left(\bigcup_{i=1}^n A_i\right) = P(\Omega) = 1$$

olur. En basit parçalanma, $0 < P(B) < 1$ olan bir B olayı için $\{B, B^c\}$ ikilisidir. Aynı tanım, sayılabilir sonsuz bir $A_1, A_2, \dots$ ailesi için de kelimesi kelimesine kurulur ve aşağıdaki iki teorem o durumda da geçerlidir; ispatlarda sonlu toplamsallık yerine (A3) kullanılır.

Teorem 6.4 (Toplam Olasılık Formülü). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı ve $A_1, A_2, \dots, A_n$ olayları Ω 'nın bir sonlu parçalanması olsun. Her $B \in \mathcal{U}$ için

$$P(B) = \sum_{i=1}^n P(A_i) P(B | A_i)$$

olur.



Şekil 6.2: Toplam olasılık formülünün Venn şeması. Düşey çizgiler örnek uzay Ω 'yı ikiye ikiye ayrık A_1, A_2, A_3, A_4 parçalarına böler; B olayı da bu çizgilerle dört ayrık dilime $B \cap A_i$ ayrılır. Ayrık dilimlerin olasılıkları toplanarak $P(B)$ bulunur; her dilimin olasılığı ise koşullu olasılığın tanımıyla $P(A_i) P(B | A_i)$ biçiminde yazılır.

İspat.

$B = B \cap \Omega$ ve parçalanma tanımının (c) koşulu ile, kesişimin birleşim üzerine dağılmasından

$$B = B \cap \left(\bigcup_{i=1}^n A_i\right) = \bigcup_{i=1}^n (B \cap A_i) = (B \cap A_1) \cup (B \cap A_2) \cup \dots \cup (B \cap A_n)$$

yazılır. $B \cap A_i$ kümeleri $\mathcal{U}$ 'dadır ve ikiye ikiye ayrıktır: $i \neq j$ için

$$(B \cap A_i) \cap (B \cap A_j) \subseteq A_i \cap A_j = \emptyset.$$

Sonlu toplamsallık (Teorem 3.1) gereği

$$P(B) = \sum_{i=1}^n P(B \cap A_i).$$

Parçalanma tanımının (a) koşulu $P(A_i) > 0$ verir; dolayısıyla Tanım 6.1 gereği her i için $P(B \cap A_i) = P(A_i) P(B | A_i)$ 'dir. Yerine koyunca

$$P(B) = \sum_{i=1}^n P(A_i) P(B | A_i)$$

elde edilir.

■

İki parçalı özel durum en sık kullanılanıdır: $0 < P(A) < 1$ ise her B için

$$P(B) = P(A) P(B | A) + P(A^c) P(B | A^c).$$

İ Hipotez zayıflatılabilir

Teoremden $A_1 \cup \dots \cup A_n = \Omega$ koşulu yerine daha zayıf olan $\sum_{i=1}^n P(A_i) = 1$ koşulu da yeter. Gerçekten $D = \bigcup_i A_i$ dersek $P(D) = \sum_i P(A_i) = 1$, dolayısıyla $P(D^c) = 0$ 'dır. $B = (B \cap D) \cup (B \cap D^c)$ ayrık birleşiminde ikinci parçanın olasılığı monotonluk gereği $P(B \cap D^c) \leq P(D^c) = 0$ 'dır; öyleyse $P(B) = P(B \cap D)$ ve ispat aynen sürer. Benzer biçimde, $P(A_i) = 0$ olan parçalar toplamdan atılabilir; formül geri kalan parçalar üzerinden doğru kalır.

6.4 Bayes Teoremi

Toplam olasılık formülünde bilgi akışı “neden”den “sonuç”a doğrudur: parçalanmanın olasılıkları $P(A_i)$ ile her durumda sonucun olasılığı $P(B | A_i)$ biliniyorken $P(B)$ hesaplanır. Uygulamada çoğu zaman tersi sorulur: sonuç B gözlenmiştir, hangi durumdan geldiğini merak ediyoruz; yani $P(A_j | B)$ isteniyor. Koşullu olasılığın tanımı ile toplam olasılık formülü yan yana getirildiğinde yanıt hemen çıkar.

Teorem 6.5 (Bayes Teoremi). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı, $A_1, A_2, \dots, A_n$ olayları Ω 'nın bir sonlu parçalanması ve $B \in \mathcal{U}$, $P(B) > 0$ olsun. Her $j = 1, 2, \dots, n$ için

$$P(A_j | B) = \frac{P(A_j) P(B | A_j)}{P(B)} = \frac{P(A_j) P(B | A_j)}{\sum_{i=1}^n P(A_i) P(B | A_i)}$$

olur.

İspat.

$P(B) > 0$ olduğundan Tanım 6.1 gereği

$$P(A_j | B) = \frac{P(A_j \cap B)}{P(B)}.$$

Payı yeniden yazalım: parçalanma tanımı $P(A_j) > 0$ verdiğiinden, koşullu olasılığın tanımı (ya da Teorem 6.3'nin $n = 2$ hâli) ile

$$P(A_j \cap B) = P(A_j) P(B | A_j).$$

Paydayı ise Teorem 6.4 ile açalım:

$$P(B) = \sum_{i=1}^n P(A_i) P(B | A_i).$$

İkisini yerine koyunca teoremdaki eşitlik elde edilir.

■

Teoremdaki büyüklüklerin adları vardır. $P(A_j)$ sayısına, B gözlenmeden önceki bilgiyi yansıttığından A_j 'nin **önsel olasılığı** (a priori olasılık); $P(A_j | B)$ sayısına, gözlemden sonraki bilgiyi yansıttığından **sonsal olasılığı** (a posteriori olasılık) denir. Bayes teoremi, önsel olasılıkları gözlem ışığında sonsal

olasılıklara günceller. Sonsal olasılıkların toplamı 1'dir. Bu, Teorem 6.1 ile parçalanma tanımından çıkar; doğrudan da görülür: j üzerinden toplandığında paylar tam paydayı verir.

! Bayes teoremi koşulun yönünü çevirir

$P(B | A_j)$ ile $P(A_j | B)$ farklı sayılardır ve gündelik dilde sürekli birbirine karıştırılır. “Hastaların %99’unda test pozitif çıkar” cümlesi $P(\text{pozitif} | \text{hasta})$ hakkındadır; “testi pozitif çıkan kişi %99 olasılıkla hastadır” cümlesi ise $P(\text{hasta} | \text{pozitif})$ hakkındadır ve birincisinden **çıkamaz**. İkinciye ulaşmanın tek yolu Bayes teoremidir ve yanıt, önsel olasılığa güçlü biçimde bağlıdır.

Örnek 6.5 (Üç Torba). Üç torbada beyaz ve kırmızı toplar vardır: I. torbada 1 beyaz, 3 kırmızı; II. torbada 2 beyaz, 2 kırmızı; III. torbada 3 beyaz, 1 kırmızı top bulunmaktadır. Rastgele seçilen bir torbadan bir top çekiliyor. Çekilen top beyaz ise bu topun I. torbadan çekilmiş olma olasılığı nedir?

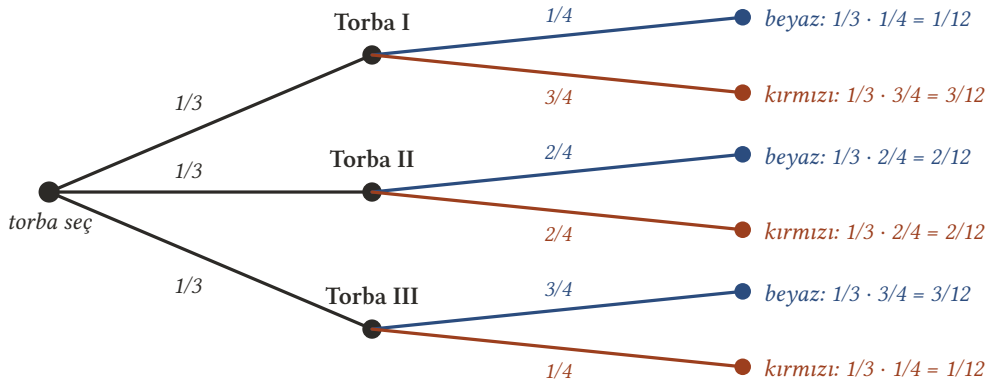
Çözüm.

A_i , seçilen torbanın i -inci torba olması olayı olsun ($i = 1, 2, 3$). Bu üç olay Ω ’nın bir sonlu parçalanmasıdır: tam olarak bir torba seçilir ve torba rastgele seçildiğinden

$$P(A_1) = P(A_2) = P(A_3) = \frac{1}{3}.$$

B çekilen topun beyaz olması olayı olsun. Hangi torbanın seçildiği bilinirse beyaz çekme olasılığı torbanın içeriğinden okunur:

$$P(B | A_1) = \frac{1}{4}, \quad P(B | A_2) = \frac{2}{4}, \quad P(B | A_3) = \frac{3}{4}.$$



$$P(B) = 1/12 + 2/12 + 3/12 = 1/2$$

$$P(I | B) = (1/12) / (1/2) = 1/6$$

Şekil 6.3: Üç torba problemi için ağaç diyagramı. Birinci dallanma torbanın seçilmesi, ikincisi topun çekilmesidir; bir yaprağın olasılığı yol üzerindeki olasılıkların çarpımıdır. Beyaz top çekilmesi olasılığı beyaz yaprakların toplamıdır (toplam olasılık formülü); çekilen top beyazken torbanın I olması olasılığı, ilgili yaprağın bu toplama oranıdır (Bayes teoremi).

Adım 1: $P(B)$. Teorem 6.4 ile

$$P(B) = \sum_{i=1}^3 P(A_i) P(B | A_i) = \frac{1}{3} \cdot \frac{1}{4} + \frac{1}{3} \cdot \frac{2}{4} + \frac{1}{3} \cdot \frac{3}{4} = \frac{1}{3} \cdot \frac{6}{4} = \frac{1}{2}.$$

Adım 2: $P(A_1 | B)$. Teorem 6.5 ile

$$P(A_1 | B) = \frac{P(A_1) P(B | A_1)}{P(B)} = \frac{\frac{1}{3} \cdot \frac{1}{4}}{\frac{1}{2}} = \frac{1/12}{1/2} = \frac{1}{6}.$$

Aynı yolla öteki sonsal olasılıklar da bulunur:

$$P(A_2 | B) = \frac{\frac{1}{3} \cdot \frac{2}{4}}{\frac{1}{2}} = \frac{1}{3}, \quad P(A_3 | B) = \frac{\frac{1}{3} \cdot \frac{3}{4}}{\frac{1}{2}} = \frac{1}{2}.$$

Toplamları $\frac{1}{6} + \frac{1}{3} + \frac{1}{2} = 1$ 'dir. Top çekilmeden önce üç torba eşit olasılıklıydı; beyaz topun gözlenmesi, beyazı en çok olan III. torbayı en olası, en az olan I. torbayı en az olası duruma getirmiştir.

■

Örnek 6.6 (İki Yüzü Yazı Olan Para). Bir çantada üç madeni para vardır: ilk ikisi düzgündür, üçüncüsünün ise iki yüzü de yazıdır. Çantadan rastgele bir para çekiliyor ve dört kez atılıyor. Dört atışın hepsinde yazı gelirse, çekilen paranın iki yüzü yazı olan para olma olasılığı nedir?

Çözüm.

A_i , i -inci paranın çekilmesi olayı olsun ($i = 1, 2, 3$). Bu olaylar Ω 'nın bir sonlu parçalanmasıdır ve para rastgele çekildiğinden

$$P(A_1) = P(A_2) = P(A_3) = \frac{1}{3}.$$

B , dört atışın hepsinde yazı gelmesi olayı olsun. Düzgün bir para dört kez atıldığında yazı-tura dizilerinin sayısı $2^4 = 16$ 'dır ve bunlar eşit olasılıklıdır; hepsinin yazı olduğu tek bir dizi vardır. Öyleyse

$$P(B | A_1) = P(B | A_2) = \frac{1}{2^4} = \frac{1}{16}.$$

Üçüncü paranın her atışı yazı verir: $P(B | A_3) = 1$.

Adım 1: $P(B)$. [Teorem 6.4](#) ile

$$P(B) = \sum_{i=1}^3 P(A_i) P(B | A_i) = \frac{1}{3} \cdot \frac{1}{16} + \frac{1}{3} \cdot \frac{1}{16} + \frac{1}{3} \cdot 1 = \frac{1}{3} \cdot \frac{18}{16} = \frac{3}{8}.$$

Adım 2: $P(A_3 | B)$. [Teorem 6.5](#) ile

$$P(A_3 | B) = \frac{P(A_3) P(B | A_3)}{P(B)} = \frac{\frac{1}{3} \cdot 1}{\frac{3}{8}} = \frac{8}{9}.$$

Önsel olasılık $\frac{1}{3}$ iken dört yazının gözlenmesi sonsal olasılığı $\frac{8}{9}$ 'a çıkarmıştır. Geriye kalan $\frac{1}{9}$, iki düzgün para arasında eşit paylaşılr: $P(A_1 | B) = P(A_2 | B) = \frac{1}{18}$.

■

Örnek 6.7 (Tıbbi Test). Bir hastalık toplumun %1'inde görülmektedir. Hastalığı saptamak için kullanılan bir test, hasta kişilerin %99'unda pozitif sonuç vermektedir (testin **duyarlılığı**); sağlıklı kişilerin ise %98'inde negatif sonuç vermektedir (testin **özgüllüğü**), yani sağlıklıların %2'sinde **yanlış pozitif** çıkmaktadır. Toplumdan rastgele seçilen bir kişiye test uygulanıyor.

a) Test pozitif çıkarsa kişinin hasta olma olasılığı nedir?

b) Test negatif çıkarsa kişinin sağlıklı olma olasılığı nedir?

Çözüm.

H kişinin hasta olması olayı olsun; $\{H, H^c\}$ örnek uzayın bir sonlu parçalanmasıdır ve

$$P(H) = 0,01, \quad P(H^c) = 0,99.$$

T testin pozitif çıkması olayı olsun. Verilenler koşullu olasılık diliyle

$$P(T | H) = 0,99, \quad P(T^c | H^c) = 0,98$$

demektir. $P(\cdot | H^c)$ ve $P(\cdot | H)$ birer olasılık ölçüsü olduğundan ([Teorem 6.1](#)) tümleyen kuralı uygulanır:

$$P(T | H^c) = 1 - 0,98 = 0,02, \quad P(T^c | H) = 1 - 0,99 = 0,01.$$

Birincisi yanlış pozitif, ikincisi yanlış negatif oranıdır; aşağıdaki hesapta yalnızca birincisi kullanılacaktır.

a) Önce Teorem 6.4 ile pozitif sonucun olasılığı:

$$P(T) = P(H) P(T | H) + P(H^c) P(T | H^c) = 0,01 \cdot 0,99 + 0,99 \cdot 0,02 = 0,0099 + 0,0198 = 0,0297.$$

Şimdi Teorem 6.5:

$$P(H | T) = \frac{P(H) P(T | H)}{P(T)} = \frac{0,0099}{0,0297} = \frac{1}{3}.$$

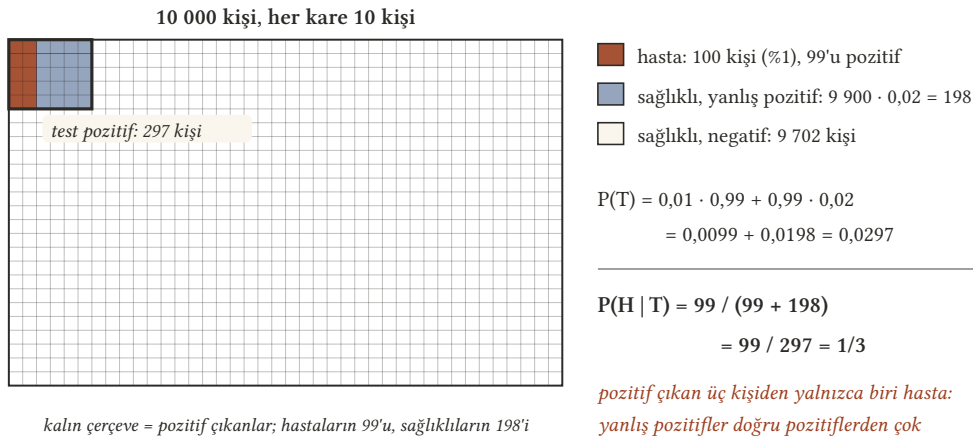
Test pozitif çıkan bir kişi ancak $\frac{1}{3}$ olasılıkla hastadır; $\frac{2}{3}$ olasılıkla sağlıklıdır.

b) $P(T^c) = 1 - P(T) = 0,9703$. Teorem 6.5 ile

$$P(H^c | T^c) = \frac{P(H^c) P(T^c | H^c)}{P(T^c)} = \frac{0,99 \cdot 0,98}{0,9703} = \frac{0,9702}{0,9703} \approx 0,9999.$$

Negatif sonuç neredeyse kesin olarak sağlıklı demektir.

(a) şıkkındaki sonuç ilk bakışta şaşırtıcıdır: duyarlılığı %99, özgüllüğü %98 olan bir test pozitif çıktığında hasta olma olasılığı yalnızca $\frac{1}{3}$ 'tür. Nedeni, hastalığın önsel olasılığının küçüklüğüdür. 10 000 kişilik bir toplulukta düşünelim: 100 kişi hastadır ve bunların 99'unda test pozitif çıkar; 9 900 kişi sağlıklıdır ve bunların %2'sinde, yani 198'inde test yanlış pozitif çıkar. Pozitif çıkan $99 + 198 = 297$ kişinin yalnızca 99'u hastadır: $\frac{99}{297} = \frac{1}{3}$. Yanlış pozitiflerin oranı küçük olsa da sağlıklı kitle hasta kitleden çok büyük olduğundan yanlış pozitiflerin sayısı doğru pozitiflerin sayısını aşar. Bu olgu, önsel olasılığı göz ardı ederek $P(T | H)$ ile $P(H | T)$ 'yi karıştırmanın ne kadar yanıltıcı olabileceğini gösterir.



Şekil 6.4: Tıbbi test örneğinin sayım resmi: 10 000 kişilik toplulukta her kare 10 kişiyi gösterir. Turuncu kareler 100 hastayı (%1), açık mavi kareler testi yanlış pozitif çıkan yaklaşık 198 sağlıklı kişiyi gösterir; kalın çerçeve testi pozitif çıkan 297 kişiyi kuşatır. Hastaların %99'unda test pozitif çıksa da hasta kitle küçük, sağlıklı kitle büyük olduğundan yanlış pozitiflerin sayısı doğru pozitifleri aşar ve $P(H | T) = 99/297 = 1/3$ kalır. Bayes teoremi bu sayımın olasılık diliyle yazılmış hâlidir.

■

6.5 Alıştırmalar

Alıştırma 6.1 (Koşullu Olasılık, Toplam Olasılık ve Bayes Teoremi Üzerine). a) Düzgün bir zar atılıyor. A üste gelen sayının çift olması, B üste gelen sayının en az 3 olması olayı olsun. $P(A | B)$, $P(B | A)$ ve $P(A \setminus B)$ sayılarını hesaplayınız ve karşılaştırınız.

b) $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı, $A, B \in \mathcal{U}$ ve $P(B) > 0$ olsun. **Teorem 6.1**'a başvurmadan, doğrudan **Tanım 6.1** ve **Teorem 3.1** ile şunları gösteriniz: (i) $A \cap B = \emptyset$ ise $P(A | B) = 0$; (ii) $B \subseteq A$ ise $P(A | B) = 1$; (iii) $A \subseteq B$ ise $P(A | B) = \frac{P(A)}{P(B)} \geq P(A)$; (iv) $P(A^c | B) = 1 - P(A | B)$.

c) İyi karıştırılmış 52 kartlık bir desteden iadesiz olarak art arda üç kart çekiliyor. Üçünün de kupa olması olasılığını çarpma teoremiyle bulunuz ve sonucu sayma yoluyla doğrulayınız. (Destede 13 kupa vardır.)

d) I. torbada 2 beyaz, 3 siyah; II. torbada 4 beyaz, 1 siyah top vardır. I. torbadan rastgele bir top çekilip bakılmadan II. torbaya atılıyor, ardından II. torbadan bir top çekiliyor. II. torbadan çekilen topun beyaz olma olasılığı nedir?

e) Bir fabrikada üretimin %50'si birinci, %30'u ikinci, %20'si üçüncü makineden çıkmaktadır. Bu makinelerin ürettiği parçaların sırasıyla %1'i, %2'si ve %5'i kusurludur. Üretimden rastgele seçilen bir parça kusurlu bulunuyor. Bu parçanın üçüncü makineden çıkmış olma olasılığı nedir? Kusurlu parçanın en olası kaynağı hangi makinedir?

f) **Örnek 6.6**'daki çantadan çekilen para n kez atılıyor ve hepsinde yazı geliyor. Çekilen paranın iki yüzü yazı olan para olma olasılığını n cinsinden bulunuz; $n = 4$ için sonucu örnekle karşılaştırınız ve $n \rightarrow \infty$ iken limiti yorumlayınız.

g) $P(A) > 0$ ve $P(B) > 0$ olsun. $P(A | B) > P(A)$ olması için gerek ve yeter koşulun $P(B | A) > P(B)$ olduğunu gösteriniz. (B bilgisi A 'yı daha olası kılıyorsa A bilgisi de B 'yi daha olası kılar.)

Çözüm.

a) $\Omega = \{1, 2, 3, 4, 5, 6\}$ Laplace uzayında $A = \{2, 4, 6\}$, $B = \{3, 4, 5, 6\}$ ve $A \cap B = \{4, 6\}$ 'dır.

$$P(A | B) = \frac{2/6}{4/6} = \frac{1}{2}, \quad P(B | A) = \frac{2/6}{3/6} = \frac{2}{3}, \quad P(A \setminus B) = P(\{2\}) = \frac{1}{6}.$$

Üç sayı da farklıdır: iki koşullu olasılık koşulun yönüne göre değişir, fark kümesinin olasılığı ise bambaşka bir olayın olasılığıdır. Öte yandan $P(A) = \frac{1}{2}$ ve $P(B) = \frac{2}{3}$ olduğundan $P(A | B) = P(A)$ ve $P(B | A) = P(B)$ 'dir: burada B bilgisi A 'nın olasılığını hiç değiştirmemiştir. Bu durumun adı bir sonraki bölümün konusudur.

b) (i) $A \cap B = \emptyset$ ise $P(A \cap B) = P(\emptyset) = 0$, dolayısıyla $P(A | B) = 0/P(B) = 0$.

(ii) $B \subseteq A$ ise $A \cap B = B$, dolayısıyla $P(A | B) = P(B)/P(B) = 1$.

(iii) $A \subseteq B$ ise $A \cap B = A$, dolayısıyla $P(A | B) = P(A)/P(B)$. $0 < P(B) \leq 1$ olduğundan $P(A)/P(B) \geq P(A)$.

(iv) $A^c \cap B = B \setminus (A \cap B)$ ve $A \cap B \subseteq B$ olduğundan **Teorem 3.1**'nin fark kuralıyla $P(A^c \cap B) = P(B) - P(A \cap B)$. Buradan

$$P(A^c | B) = \frac{P(B) - P(A \cap B)}{P(B)} = 1 - \frac{P(A \cap B)}{P(B)} = 1 - P(A | B).$$

c) K_i , i -inci kartın kupa olması olayı olsun. $P(K_1) = \frac{13}{52}$; ilk kart kupaysa 51 kartın 12'si kupadır, $P(K_2 | K_1) = \frac{12}{51}$; ilk ikisi kupaysa 50 kartın 11'i kupadır, $P(K_3 | K_1 \cap K_2) = \frac{11}{50}$. $P(K_1 \cap K_2) = \frac{13 \cdot 12}{52 \cdot 51} > 0$ olduğundan **Teorem 6.3** ile

$$P(K_1 \cap K_2 \cap K_3) = \frac{13}{52} \cdot \frac{12}{51} \cdot \frac{11}{50} = \frac{1716}{132600} = \frac{11}{850}.$$

Sayma ile doğrulama: sırasız üç kart $\binom{52}{3} = 22100$ yolla, üç kupa $\binom{13}{3} = 286$ yolla seçilir; $\frac{286}{22100} = \frac{11}{850}$.

d) W I. torbadan çekilen topun beyaz olması, B II. torbadan çekilen topun beyaz olması olayı olsun. $\{W, W^c\}$ bir parçalanmadır: $P(W) = \frac{2}{5}$, $P(W^c) = \frac{3}{5}$. Aktarılan top beyazsa II. torbada 5 beyaz, 1 siyah; siyahsa 4 beyaz, 2 siyah top olur:

$$P(B | W) = \frac{5}{6}, \quad P(B | W^c) = \frac{4}{6}.$$

Teorem 6.4 ile

$$P(B) = \frac{2}{5} \cdot \frac{5}{6} + \frac{3}{5} \cdot \frac{4}{6} = \frac{10}{30} + \frac{12}{30} = \frac{22}{30} = \frac{11}{15}.$$

e) M_1, M_2, M_3 parçanın ilgili makineden çıkması olayları olsun; bunlar bir sonlu parçalanmadır, $P(M_1) = 0,5, P(M_2) = 0,3, P(M_3) = 0,2$. D parçanın kusurlu olması olayı için $P(D | M_1) = 0,01, P(D | M_2) = 0,02, P(D | M_3) = 0,05$. [Teorem 6.4](#) ile

$$P(D) = 0,5 \cdot 0,01 + 0,3 \cdot 0,02 + 0,2 \cdot 0,05 = 0,005 + 0,006 + 0,010 = 0,021.$$

[Teorem 6.5](#) ile

$$P(M_3 | D) = \frac{0,2 \cdot 0,05}{0,021} = \frac{0,010}{0,021} = \frac{10}{21}.$$

Ötekiler $P(M_1 | D) = \frac{0,005}{0,021} = \frac{5}{21}$ ve $P(M_2 | D) = \frac{0,006}{0,021} = \frac{6}{21}$ 'dir; toplam 1'dir. Üretimin yalnızca beşte birini yapmasına karşın kusurlu parçanın en olası kaynağı üçüncü makinedir, çünkü kusur oranı ötekilerden çok daha yüksektir.

f) B_n, n atışın hepsinde yazı gelmesi olayı olsun. Düzgün para için 2^n eşit olasılıklı dizinin yalnızca biri tümüyle yazıdır: $P(B_n | A_1) = P(B_n | A_2) = 2^{-n}$; ayrıca $P(B_n | A_3) = 1$. [Teorem 6.4](#) ile

$$P(B_n) = \frac{1}{3} \cdot 2^{-n} + \frac{1}{3} \cdot 2^{-n} + \frac{1}{3} \cdot 1 = \frac{2 \cdot 2^{-n} + 1}{3} = \frac{2 + 2^n}{3 \cdot 2^n}.$$

[Teorem 6.5](#) ile

$$P(A_3 | B_n) = \frac{\frac{1}{3} \cdot 1}{\frac{2+2^n}{3 \cdot 2^n}} = \frac{2^n}{2^n + 2}.$$

$n = 4$ için $\frac{16}{18} = \frac{8}{9}$, örnekteki sonuçtur. $n \rightarrow \infty$ iken $\frac{2^n}{2^n + 2} = \frac{1}{1 + 2^{1-n}} \rightarrow 1$: art arda gelen yazılar uzakça hileli paranın sonsal olasılığı 1'e yaklaşır; düzgün bir paranın çok uzun bir yazı dizisi üretmesi giderek daha az inandırıcı olur. Buna karşın hiçbir sonlu n için olasılık tam 1 olmaz, çünkü düzgün para da ilkece n yazı üretebilir.

g) $P(B) > 0$ olduğundan $P(A | B) = \frac{P(A \cap B)}{P(B)}$; her iki taraf $P(B)$ ile çarpılırsa

$$P(A | B) > P(A) \Leftrightarrow P(A \cap B) > P(A) P(B).$$

Aynı biçimde $P(A) > 0$ olduğundan

$$P(B | A) > P(B) \Leftrightarrow P(A \cap B) > P(A) P(B).$$

İki denklik aynı koşula bağlandığından $P(A | B) > P(A) \Leftrightarrow P(B | A) > P(B)$. Aynı akıl yürütme ">" yerine "<" ve "=" için de geçerlidir; özellikle $P(A | B) = P(A)$ olması $P(A \cap B) = P(A) P(B)$ olmasına denktir. Bu son eşitlik, bir sonraki bölümün konusudur.

■

Bu bölümde B bilgisinin A 'nın olasılığını nasıl değiştirdiğini inceledik. Sıradaki bölümde bu bilginin hiçbir şey değiştirmedigi, yani $P(A | B) = P(A)$ olan durumu ele alıyoruz: [Bağımsız Olaylar](#).

7. Bağımsız Olaylar

Koşullu Olasılık, Toplam Olasılık ve Bayes Teoremi bölümünde bir olayın gerçekleştiği bilgisinin başka bir olayın olasılığını nasıl değiştirdiğini inceledik: $P(B) > 0$ iken $P(A | B) = \frac{P(A \cap B)}{P(B)}$ sayısı, B 'nin gerçekleştiği bilindiğinde A 'ya biçilen yeni olasılıktır. Bu bölümde en ilginç özel durumu ele alıyoruz: B 'nin gerçekleşmesi A 'nın olasılığını **hiç değiştirmiyorsa** ne olur? Bu durumda A ile B olaylarına **bağımsız** diyeceğiz.

Bağımsızlık, olasılık teorisinin en çok kullanılan kavramlarından biridir. Bir zarın ardışık iki atılışı, bir torbadan iadeli çekilişler, farklı makinelerin arızaları gibi “birbirini etkilemeyen” olguları modellerken hep bu kavrama başvururuz. Kavramın gücü şuradadır: bağımsız olayların kesişiminin olasılığı, olayların olasılıklarının çarpımıdır; dolayısıyla karmaşık bir olayın olasılığı basit parçaların olasılıklarından hesaplanabilir.

Bu bölümde şunları öğreneceğiz: iki olayın bağımsızlığının tanımı ve koşullu olasılıkla ilişkisi; **ayrık** ile **bağımsız** kavramlarının neden birbirine karıştırılmaması gerektiği; bir olay çiftinin bağımsızlığının tümleyenlere geçmesi; $P(A | B) = P(A | B^c)$ eşitliğinin bağımsızlıkla denk olması; ikiden çok olay için ikili, üçlü, n -li ve tam bağımsızlık kavramları ve bunların birbirini gerektirmediğini gösteren örnekler.

7.1 Bağımsızlığın Tanımı

$P(B) > 0$ olsun. B 'nin gerçekleşmesinin A 'nın olasılığını değiştirmemesi, $P(A | B) = P(A)$ demektir. Koşullu olasılığın tanımı (Tanım 6.1) gereği bu eşitlik

$$\frac{P(A \cap B)}{P(B)} = P(A), \quad \text{yani} \quad P(A \cap B) = P(A) P(B)$$

biçiminde yazılabilir. Sağdaki eşitlik $P(B) > 0$ koşulunu içermez ve A ile B 'ye göre bakışlıdır (simetrik değildir). Bu yüzden bağımsızlığı koşullu olasılıkla değil, doğrudan bu çarpım eşitliğiyle tanımlamak daha uygundur.

Tanım 7.1 (Bağımsız Olaylar). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı ve $A, B \in \mathcal{U}$ olsun. Eğer

$$P(A \cap B) = P(A) P(B)$$

ise A ile B olaylarına **bağımsız olaylar** denir. Bu eşitlik sağlanmıyorsa A ile B olaylarına **bağımlı olaylar** denir.

Tanım bakışlıdır: A, B 'den bağımsızsa B de A 'dan bağımsızdır. Bu yüzden kısaca “ A ile B bağımsızdır” deriz. Tanımda olasılıkların pozitif olması istenmediğine dikkat edin; olasılığı sıfır olan olaylar için de tanım anlamlıdır. İlk gözlemimiz bunun en uç iki örneğidir.

Önerme 7.1 (Boş Küme ve Örnek Uzay Her Olaydan Bağımsızdır). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı olsun.

(a) $\emptyset$ ve Ω olayları her $A \in \mathcal{U}$ olayından bağımsızdır.

(b) Daha genel olarak, $P(B) = 0$ ya da $P(B) = 1$ olan her $B \in \mathcal{U}$ olayı her $A \in \mathcal{U}$ olayından bağımsızdır.

İspat.

(a) $A \in \mathcal{U}$ herhangi bir olay olsun. $\emptyset \cap A = \emptyset$ ve $P(\emptyset) = 0$ olduğundan (Teorem 3.1)

$$P(\emptyset \cap A) = P(\emptyset) = 0 = 0 \cdot P(A) = P(\emptyset) P(A)$$

olur; yani $\emptyset$ ile A bağımsızdır. Benzer biçimde $\Omega \cap A = A$ ve $P(\Omega) = 1$ olduğundan

$$P(\Omega \cap A) = P(A) = 1 \cdot P(A) = P(\Omega) P(A)$$

olur; yani Ω ile A bağımsızdır.

(b) Önce $P(B) = 0$ olsun. $A \cap B \subseteq B$ olduğundan olasılığın monotonluğu (Teorem 3.1) gereği

$$0 \leq P(A \cap B) \leq P(B) = 0,$$

yani $P(A \cap B) = 0$ 'dır. Öte yandan $P(A)P(B) = P(A) \cdot 0 = 0$ 'dır. Demek ki $P(A \cap B) = P(A)P(B)$ olur.

Şimdi $P(B) = 1$ olsun. O zaman $P(B^c) = 1 - P(B) = 0$ 'dır. A olayı, ayrık $A \cap B$ ve $A \cap B^c$ olaylarının birleşimidir; toplamsallık gereği

$$P(A) = P(A \cap B) + P(A \cap B^c).$$

$A \cap B^c \subseteq B^c$ ve $P(B^c) = 0$ olduğundan, ilk paragraftaki gibi $P(A \cap B^c) = 0$ olur. Böylece

$$P(A \cap B) = P(A) = P(A) \cdot 1 = P(A)P(B)$$

elde edilir. Her iki durumda da A ile B bağımsızdır.

■

Bu önerme, bağımsızlığın “olaylar birbirine hiç dokunmaz” gibi kümesel bir kavram olmadığını gösterir: Ω her olayı kapsadığı hâlde her olaydan bağımsızdır. Bağımsızlık kümelerin değil, **olasılıkların** arasındaki bir ilişkidir.

Şimdi tanıma giden yoldaki koşullu olasılık yorumunu bir önerme olarak kaydedelim. Bu önerme, bağımsızlığı hesaplarken en sık kullanacağımız araçtır.

Önerme 7.2 (Bağımsızlık ve Koşullu Olasılık). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı, $A, B \in \mathcal{U}$ ve $P(B) > 0$ olsun. O zaman

$$A \text{ ile } B \text{ bağımsızdır} \Leftrightarrow P(A | B) = P(A).$$

Ayrıca $P(A) > 0$ da ise bu koşullar $P(B | A) = P(B)$ eşitliğine de denktir.

İspat.

$P(B) > 0$ olduğundan $P(A | B) = \frac{P(A \cap B)}{P(B)}$ tanımlıdır.

($\Rightarrow$) A ile B bağımsız olsun; yani $P(A \cap B) = P(A)P(B)$. O zaman

$$P(A | B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A).$$

($\Leftarrow$) $P(A | B) = P(A)$ olsun. Çarpma teoremi (Teorem 6.3) gereği

$$P(A \cap B) = P(B)P(A | B) = P(B)P(A) = P(A)P(B)$$

olur; yani A ile B bağımsızdır.

Son iddia için $P(A) > 0$ olsun. Yukarıdaki akıl yürütmede A ile B 'nin rollerini değiştirirsek, A ile B 'nin bağımsız olmasının $P(B | A) = P(B)$ eşitliğine denk olduğu görülür. Bağımsızlık tanımı bakışımı olduğundan üç koşul birbirine denktir.

■

Önermenin sözel anlamı şudur: A ile B bağımsızsa, B 'nin gerçekleştiğini öğrenmek A hakkındaki bilgimizi değiştirmez; koşullu olasılık koşulsuz olasılığa eşit kalır. Tersine, $P(A | B) \neq P(A)$ ise B 'nin gerçekleşmesi A 'nın olasılığını artırmış ya da azaltmıştır ve olaylar bağımlıdır.

7.2 Bağımsız ve Ayrık Olaylar

Öğrencilerin en sık düştüğü yanlışlardan biri, “ayrık” ile “bağımsız” sözcüklerini eşanlamlı sanmaktır. Oysa bu iki kavram neredeyse birbirinin karıştıdır.

⚠ Ayırık ile bağımsız aynı şey değildir

Ayrıklık kümelerle ilgili bir kavramdır: $A \cap B = \emptyset$, yani A ile B birlikte gerçekleşemez. **Bağımsızlık** ise olasılıklarla ilgili bir kavramdır: $P(A \cap B) = P(A)P(B)$.

İki olay ayrıksa, birinin gerçekleştiğini öğrenmek ötekinin **kesinlikle gerçekleşmediğini** söyler; bu, olabilecek en güçlü ters yönlü bağımlılıktır. Örneğin düzgün bir zar atılsın; $A = \{1, 2, 3\}$ ve $D = \{4, 5, 6\}$ olsun. A ile D ayrıktır; D gerçekleşmişse A gerçekleşmemiştir: $P(A | D) = 0 \neq \frac{1}{2} = P(A)$. Öte yandan $A = \{1, 2, 3\}$ ile $B = \{3, 4\}$ ayrık değildir, ama bağımsızdır:

$$P(A \cap B) = P(\{3\}) = \frac{1}{6} = \frac{1}{2} \cdot \frac{1}{3} = P(A)P(B).$$

Aşağıdaki teorem, olasılığı pozitif olan olaylar için ayrıklık ile bağımsızlığın birlikte bulunamayacağını söyler.

Teorem 7.1 (Pozitif Olasılıklı Ayırık Olaylar Bağımsız Değildir). $(\Omega, \mathcal{U}, P)$ olasılık uzayında $A, B \in \mathcal{U}$, $P(A) > 0$ ve $P(B) > 0$ olsun. A ile B ayrık ($A \cap B = \emptyset$) ise A ile B bağımsız değildir.

İspat.

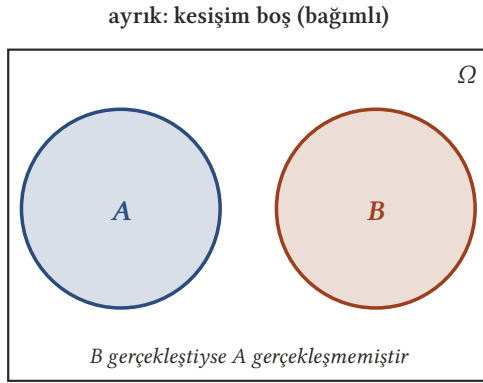
$A \cap B = \emptyset$ olduğundan

$$P(A \cap B) = P(\emptyset) = 0$$

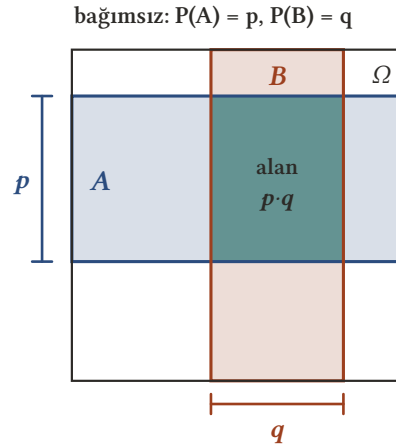
olur. Öte yandan $P(A) > 0$ ve $P(B) > 0$ olduğundan, iki pozitif sayının çarpımı olarak

$$P(A)P(B) > 0$$

dır. Demek ki $P(A \cap B) = 0 \neq P(A)P(B)$; yani bağımsızlık tanımındaki eşitlik sağlanmaz. A ile B bağımsız değildir.



$$P(A \cap B) = 0 \neq P(A)P(B) > 0$$



$$P(A \cap B) = p \cdot q = P(A)P(B) > 0$$

Şekil 7.1: Ayrıklık ile bağımsızlık farklı kavramlardır. Solda A ile B ayrıktır: birlikte gerçekleşemezler, dolayısıyla $P(A \cap B) = 0$ iken $P(A)P(B)$ pozitifdir; olaylar bağımlıdır. Sağda örnek uzay birim karedir ve olasılık alanla ölçülür: yüksekliği p olan yatay şerit A ile genişliği q olan düşey şerit B , alanı $p \cdot q = P(A)P(B)$ olan bir dikdörtgende kesişir; olaylar bağımsızdır ve kesişimleri boş değildir.

Aynı sonuca koşullu olasılıkla da ulaşılır: $P(B) > 0$ olduğundan $P(A | B) = \frac{P(A \cap B)}{P(B)} = 0 \neq P(A)$ olur ve **Önerme 7.2** gereği olaylar bağımlıdır.

■

Teoremdaki $P(A) > 0$, $P(B) > 0$ varsayımı gereklidir: $\emptyset$ ile herhangi bir A olayı hem ayrık hem de bağımsızdır (Önerme 7.1). Teoremin karşıt tersi sık kullanılan bir sonuçtur.

Sonuç 7.1 (Bağımsız Olaylar Ayrık Değildir). $(\Omega, \mathcal{U}, P)$ olasılık uzayında $A, B \in \mathcal{U}$, $P(A) > 0$ ve $P(B) > 0$ olsun. A ile B bağımsız ise A ile B ayrık değildir; dahası $P(A \cap B) > 0$ 'dır.

İspat.

A ile B bağımsız olduğundan $P(A \cap B) = P(A)P(B)$ ve bu çarpım, $P(A) > 0$ ile $P(B) > 0$ olduğundan pozitiftir. $P(A \cap B) > 0$ ise $A \cap B \neq \emptyset$ olmak zorundadır; çünkü $A \cap B = \emptyset$ olsaydı $P(A \cap B) = P(\emptyset) = 0$ olurdu. Demek ki A ile B ayrık değildir.

■

7.3 Tümlenlerin Bağımsızlığı

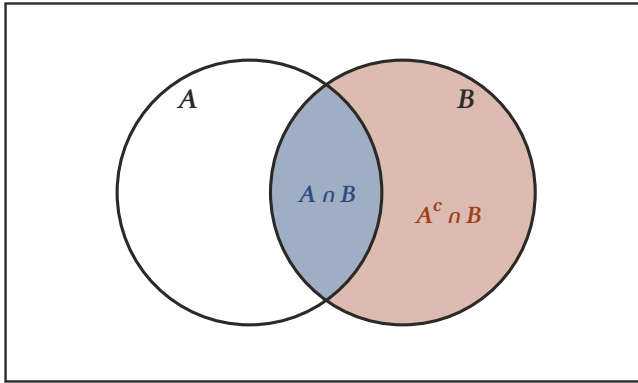
B 'nin gerçekleşmesi A 'nın olasılığını değiştirmiyorsa, B 'nin gerçekleşmemesi de değiştirmemelidir; benzer biçimde A 'nın gerçekleşmemesi de B 'den etkilenmemelidir. Bu sezgiyi doğrulayan teorem, bağımsız olay çiftlerinden tümlen alarak yeni bağımsız çiftler üretmemizi sağlar.

Teorem 7.2 (Tümlenlerin Bağımsızlığı). $(\Omega, \mathcal{U}, P)$ olasılık uzayında A ile B bağımsız olsun. O zaman

- (a) A^c ile B bağımsızdır;
- (b) A ile B^c bağımsızdır;
- (c) A^c ile B^c bağımsızdır.

İspat.

Varsayım gereği $P(A \cap B) = P(A)P(B)$ 'dir. Üç ispatta da aynı fikir kullanılır: bir olay, bir olayın gerçekleştiği ve gerçekleşmediği iki ayrık parçaya bölünür ve toplamsallık uygulanır.



B , iki ayrık parçanın birleşimi:

$A \cap B$ ve $A^c \cap B$

o yüzden

$$P(B) = P(A \cap B) + P(A^c \cap B)$$

Şekil 7.2: B olayı, A 'nın gerçekleşip gerçekleşmemesine göre iki ayrık parçaya ayrılır. A ile B bağımsızken $P(A \cap B) = P(A)P(B)$ yerine konursa, kalan parçanın olasılığı $(1 - P(A))P(B) = P(A^c)P(B)$ çıkar: tümlen de B 'den bağımsızdır.

(a) B olayı, A 'nın gerçekleştiği ve gerçekleşmediği iki parçaya ayrılır:

$$B = (A \cap B) \cup (A^c \cap B), \quad (A \cap B) \cap (A^c \cap B) = \emptyset.$$

Toplamsallık (Teorem 3.1) gereği $P(B) = P(A \cap B) + P(A^c \cap B)$ olur. Buradan

$$P(A^c \cap B) = P(B) - P(A \cap B) = P(B) - P(A)P(B) = (1 - P(A))P(B) = P(A^c)P(B)$$

elde edilir; son adımda $P(A^c) = 1 - P(A)$ kullanıldı. Demek ki A^c ile B bağımsızdır.

(b) Bu kez A olayını B 'ye göre parçalayalım:

$$A = (A \cap B) \cup (A \cap B^c), \quad (A \cap B) \cap (A \cap B^c) = \emptyset.$$

Toplamsallık gereği $P(A) = P(A \cap B) + P(A \cap B^c)$ ve dolayısıyla

$$P(A \cap B^c) = P(A) - P(A \cap B) = P(A) - P(A)P(B) = P(A)(1 - P(B)) = P(A)P(B^c)$$

olur. Demek ki A ile B^c bağımsızdır. (Bağımsızlık bakışımı olduğundan (b), (a)'da A ile B 'nin rolleri değiştirilerek de elde edilebilir.)

(c) De Morgan kuralı gereği $A^c \cap B^c = (A \cup B)^c$ olur. Toplama kuralı (Teorem 3.2) ve bağımsızlık kullanılarak

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B) - P(A)P(B)$$

bulunur. Öyleyse

$$\begin{aligned} P(A^c \cap B^c) &= P((A \cup B)^c) = 1 - P(A \cup B) \\ &= 1 - P(A) - P(B) + P(A)P(B) \\ &= (1 - P(A))(1 - P(B)) = P(A^c)P(B^c) \end{aligned}$$

olur; yani A^c ile B^c bağımsızdır.

Aynı sonuca (a) ile (b) art arda uygulanarak da ulaşılabilir: A ile B bağımsız olduğundan (a) gereği A^c ile B bağımsızdır; şimdi (b) şikkını A^c ile B çiftine uygularsak A^c ile B^c 'nin bağımsız olduğu çıkar.

■

Teoremin pratik anlamı şudur: bağımsız A ile B olayları için $\{A, A^c\}$ sınıfından bir olay ile $\{B, B^c\}$ sınıfından bir olayın kesişiminin olasılığı her zaman ilgili olasılıkların çarpımıdır. Örneğin A ile B bağımsız, $P(A) = \frac{2}{5}$ ve $P(B) = \frac{1}{2}$ ise

$$\begin{aligned} P(A \cap B) &= \frac{2}{5} \cdot \frac{1}{2} = \frac{1}{5}, & P(A^c \cap B) &= \frac{3}{5} \cdot \frac{1}{2} = \frac{3}{10}, \\ P(A \cap B^c) &= \frac{2}{5} \cdot \frac{1}{2} = \frac{1}{5}, & P(A^c \cap B^c) &= \frac{3}{5} \cdot \frac{1}{2} = \frac{3}{10} \end{aligned}$$

olur ve bu dört sayının toplamı, olması gerektiği gibi, 1'dir.

7.4 Koşullu Olasılıkların Eşitliği ve Bağımsızlık

Önerme 7.2, bağımsızlığı $P(A | B)$ ile koşulsuz $P(A)$ arasındaki eşitlikle karakterize etti. Uygulamada bazen $P(A)$ bilinmez; onun yerine B gerçekleştiğinde ve gerçekleşmediğinde A 'nın olasılıkları, yani $P(A | B)$ ile $P(A | B^c)$ bilinir. Bu iki sayının eşit olması da bağımsızlığa denktir.

Örnek 7.1 (İki Koşullu Olasılığın Eşitliğinden Bağımsızlık). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı, $A, B \in \mathcal{U}$ ve $0 < P(B) < 1$ olsun.

(a) $P(A | B) = P(A | B^c)$ ise A ile B 'nin bağımsız olduğunu gösteriniz.

(b) Tersine, A ile B bağımsız ise $P(A | B) = P(A | B^c) = P(A)$ olduğunu gösteriniz.

(c) Bir fabrikada üretilen parçaların $\frac{3}{5}$ 'i gündüz vardiyasında üretilmektedir. Gündüz üretilen parçaların $\frac{3}{10}$ 'u, gece üretilen parçaların da $\frac{3}{10}$ 'u kusurludur. Rastgele seçilen bir parçanın kusurlu olması olayı ile gündüz üretilmiş olması olayı bağımsız mıdır? Parçanın hem gündüz üretilmiş hem de kusurlu olması olasılığı nedir?

Çözüm.

$0 < P(B) < 1$ olduğundan $P(B^c) = 1 - P(B) > 0$ 'dır; dolayısıyla $P(A | B)$ ve $P(A | B^c)$ koşullu olasılıklarının ikisi de tanımlıdır.

(a) A olayı, B 'nin gerçekleştiği ve gerçekleşmediği iki ayrık parçaya ayrılır:

$$A = (A \cap B) \cup (A \cap B^c), \quad (A \cap B) \cap (A \cap B^c) = \emptyset.$$

Toplamsallık ve çarpma teoremi (Teorem 6.3) gereği

$$P(A) = P(A \cap B) + P(A \cap B^c) = P(B)P(A | B) + P(B^c)P(A | B^c).$$

Bu, $\{B, B^c\}$ parçalanmasına uygulanmış toplam olasılık formülüdür (Teorem 6.4). Varsayım gereği $P(A | B^c) = P(A | B)$ olduğundan

$$P(A) = P(B) P(A | B) + P(B^c) P(A | B) = (P(B) + P(B^c)) P(A | B) = 1 \cdot P(A | B) = P(A | B)$$

elde edilir. Böylece $P(A | B) = P(A)$ olur ve $P(B) > 0$ olduğundan Önerme 7.2 gereği A ile B bağımsızdır. Açıkça yazarsak, $P(A | B) = P(A)$ eşitliği $\frac{P(A \cap B)}{P(B)} = P(A)$, yani $P(A \cap B) = P(A) P(B)$ demektir.

(b) A ile B bağımsız olsun. $P(B) > 0$ olduğundan Önerme 7.2 gereği $P(A | B) = P(A)$ 'dır. Teorem 7.2 (b) gereği A ile B^c de bağımsızdır ve $P(B^c) > 0$ olduğundan yine Önerme 7.2 ile $P(A | B^c) = P(A)$ olur. Demek ki

$$P(A | B) = P(A) = P(A | B^c).$$

(a) ile (b) birlikte şunu söyler: $0 < P(B) < 1$ iken

$$A \text{ ile } B \text{ bağımsızdır} \Leftrightarrow P(A | B) = P(A | B^c).$$

(c) A parçanın kusurlu olması, B gündüz üretilmiş olması olayı olsun. Verilenler $P(B) = \frac{3}{5}$, $P(B^c) = \frac{2}{5}$, $P(A | B) = \frac{3}{10}$ ve $P(A | B^c) = \frac{3}{10}$ 'dur. $0 < P(B) < 1$ ve $P(A | B) = P(A | B^c)$ olduğundan (a) gereği A ile B bağımsızdır. Toplam olasılık formülüyle

$$P(A) = P(B) P(A | B) + P(B^c) P(A | B^c) = \frac{3}{5} \cdot \frac{3}{10} + \frac{2}{5} \cdot \frac{3}{10} = \frac{3}{10}$$

bulunur; beklendiği gibi $P(A) = P(A | B)$ 'dir. İstenen olasılık

$$P(A \cap B) = P(A) P(B) = \frac{3}{10} \cdot \frac{3}{5} = \frac{9}{50}$$

olur.

■

7.5 Çoklu Bağımsızlık

İkiden çok olay için bağımsızlık kavramı, ilk bakışta sanıldığından daha incedir. Üç olayın “birbirini etkilememesi” için yalnızca ikişer ikişer bağımsız olmaları yetmez; herhangi ikisinin birlikte gerçekleşmesi de üçüncüyü etkilememelidir. Bu yüzden önce alt kümeler üzerinden tanımlanan **k -li bağımsızlık** kavramını, sonra bunların hepsini birden isteyen **tam bağımsızlık** kavramını veriyoruz.

Tanım 7.2 (İkili, Üçlü, n -li ve Tam Bağımsızlık). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı, $A_1, A_2, \dots, A_n \in \mathcal{U}$ ve $2 \leq k \leq n$ olsun.

- $1 \leq i_1 < i_2 < \dots < i_k \leq n$ olan **her** indis seçimi için

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = P(A_{i_1}) P(A_{i_2}) \dots P(A_{i_k})$$

oluyorsa $A_1, A_2, \dots, A_n$ olaylarına **k -li bağımsız** denir. $k = 2$ için **ikili bağımsız**, $k = 3$ için **üçlü bağımsız** denir. $k = n$ için tek bir koşul vardır:

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1) P(A_2) \dots P(A_n);$$

bu sağlandığında olaylara **n -li bağımsız** denir.

- $A_1, A_2, \dots, A_n$ olayları her $k = 2, 3, \dots, n$ için k -li bağımsız ise bu olaylara **tam bağımsız** (karşılıklı bağımsız) denir. Başka bir deyişle, en az iki elemanlı her $I \subseteq \{1, 2, \dots, n\}$ indis kümesi için

$$P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i)$$

olmalıdır.

Sonsuz bir $A_1, A_2, \dots$ olaylar dizisi, her sonlu alt ailesi tam bağımsız olduğunda **tam bağımsız** olarak adlandırılır.

Açıkça yazılırsa ikili bağımsızlık için $\binom{n}{2}$, üçlü bağımsızlık için $\binom{n}{3}$ koşul vardır ve tam bağımsızlık, toplam $2^n - n - 1$ koşulun hepsinin sağlanmasını ister. $n = 2$ için tam bağımsızlık, iki olay için verilen bağımsızlıkla (Tanım 7.1) aynı şeydir. $n = 3$ için ise dört koşul gerekir:

$$\begin{aligned} P(A_1 \cap A_2) &= P(A_1) P(A_2), & P(A_1 \cap A_3) &= P(A_1) P(A_3), \\ P(A_2 \cap A_3) &= P(A_2) P(A_3), & P(A_1 \cap A_2 \cap A_3) &= P(A_1) P(A_2) P(A_3). \end{aligned}$$

Bu dört koşulun hiçbiri ötekilerden çıkarılamaz; aşağıdaki örnekler bunu gösterir.

! k -li bağımsızlık m -li bağımsızlığı gerektirmez

$k \neq m$ olmak üzere k -li bağımsız olaylar m -li bağımsız olmak zorunda değildir. Bunun iki yönü de vardır: ikili bağımsız olduğu hâlde üçlü bağımsız olmayan olaylar ve üçlü bağımsız olduğu hâlde ikili bağımsız olmayan olaylar vardır. Aşağıdaki iki örnek bu iki durumu gösterir. Dolayısıyla tam bağımsızlığı doğrulamak için **bütün** alt kümelerle bakmak gerekir; yalnızca ikili koşulları ya da yalnızca n -li koşulu denetlemek yetmez.

Örnek 7.2 (İkili Bağımsız Ama Tam Bağımsız Olmayan Üç Olay). Düzgün bir zar ardışık olarak iki kez atılsın. A ilk atışta 2 gelmesi, B ikinci atışta 5 gelmesi ve C gelen sonuçların toplamının 7 olması olayı olsun. A , B ve C olaylarının tam bağımsız olup olmadığını inceleyiniz.

Çözüm.

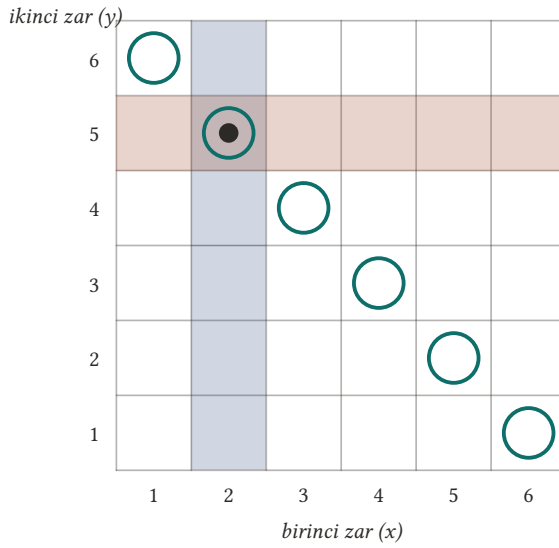
Örnek uzay $\Omega = \{(x, y) : x, y \in \{1, 2, \dots, 6\}\}$ olup $n(\Omega) = 36$ 'dır. Zar düzgün olduğundan bütün sonuçlar eşit olasılıklıdır; $(\Omega, \mathcal{P}(\Omega), P)$ bir Laplace uzayıdır (Tanım 4.1) ve her $E \subseteq \Omega$ için $P(E) = \frac{n(E)}{36}$ 'dır (Önerme 4.2).

Olayları açıkça yazalım:

$$\begin{aligned} A &= \{(2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6)\}, & n(A) &= 6, \\ B &= \{(1, 5), (2, 5), (3, 5), (4, 5), (5, 5), (6, 5)\}, & n(B) &= 6, \\ C &= \{(x, y) : x + y = 7\} = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}, & n(C) &= 6. \end{aligned}$$

Buradan

$$P(A) = P(B) = P(C) = \frac{6}{36} = \frac{1}{6}.$$



A: $x = 2$ (sütun)

B: $y = 5$ (sıra)

C: $x + y = 7$ (çember)

her biri $6/36 = 1/6$

ikili kesişimler tek nokta: $1/36 = 1/6 \cdot 1/6$

$A \cap B \cap C = \{(2, 5)\}$: $1/36 \neq 1/216$

Şekil 7.3: İki zar atışının 36 sonucu. A (ilk zar 2) bir sütun, B (ikinci zar 5) bir sıra, C (toplam 7) bir köşegendir; üçünün de olasılığı $1/6$ 'dır ve her ikili kesişim tek bir kareden oluşur, yani olaylar ikişer ikişer bağımsızdır. Ama üçlü kesişim de aynı tek kare (2, 5)'tir: $1/36 \neq (1/6)^3$, dolayısıyla üçü birlikte bağımsız değildir.

İkili kesişimler. $A \cap B$, ilk atışta 2 ve ikinci atışta 5 gelmesi olayıdır: $A \cap B = \{(2, 5)\}$. $A \cap C$ için ilk atış 2 ise toplamın 7 olması için ikinci atış 5 olmalıdır: $A \cap C = \{(2, 5)\}$. $B \cap C$ için ikinci atış 5 ise ilk atış 2 olmalıdır: $B \cap C = \{(2, 5)\}$. Demek ki

$$P(A \cap B) = P(A \cap C) = P(B \cap C) = \frac{1}{36}.$$

Öte yandan

$$P(A)P(B) = P(A)P(C) = P(B)P(C) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

olduğundan

$$P(A \cap B) = P(A)P(B), \quad P(A \cap C) = P(A)P(C), \quad P(B \cap C) = P(B)P(C)$$

eşitliklerinin üçü de sağlanır: A, B ve C **ikili bağımsızdır**.

Üçlü kesişim. $A \cap B \cap C = \{(2, 5)\}$ olduğundan $P(A \cap B \cap C) = \frac{1}{36}$ 'dır. Ancak

$$P(A)P(B)P(C) = \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{216} \neq \frac{1}{36} = P(A \cap B \cap C).$$

Demek ki A, B ve C üçlü bağımsız değildir; dolayısıyla **tam bağımsız değildir**.

Sonucun yorumu açıktır: A ile B birlikte gerçekleştiğinde, yani sonuç (2, 5) olduğunda, toplam kesinlikle 7'dir. Yani $P(C | A \cap B) = 1 \neq \frac{1}{6} = P(C)$; A ve B tek tek C hakkında bilgi vermez, ama ikisi birlikte C'yi tümüyle belirler.

■

Ters yön de mümkündür: üçlü koşul sağlanırken ikili koşulların hiçbirisi sağlanmayabilir.

Örnek 7.3 (Üçlü Bağımsız Ama İkili Bağımsız Olmayan Üç Olay). 1'den 8'e kadar numaralanmış sekiz toptan biri rastgele çekiliyor. $A = \{1, 2, 3, 4\}$, $B = \{1, 5, 6, 7\}$ ve $C = \{1, 2, 3, 8\}$ olsun. A, B ve C olaylarının üçlü bağımsız olduğunu, ama ikili bağımsız olmadığını gösteriniz.

Çözüm.

$\Omega = \{1, 2, \dots, 8\}$ ve her $E \subseteq \Omega$ için $P(E) = \frac{n(E)}{8}$ 'dir. Üç olayın da dört elemanı olduğundan

$$P(A) = P(B) = P(C) = \frac{4}{8} = \frac{1}{2}.$$

Üçlü kesişim. Üç kümenin ortak tek elemanı 1'dir: $A \cap B \cap C = \{1\}$. Öyleyse

$$P(A \cap B \cap C) = \frac{1}{8} = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = P(A) P(B) P(C)$$

olur; A, B ve C üçlü bağımsızdır.

İkili kesişimler. $A \cap B = \{1\}$, $A \cap C = \{1, 2, 3\}$ ve $B \cap C = \{1\}$ olduğundan

$$P(A \cap B) = \frac{1}{8}, \quad P(A \cap C) = \frac{3}{8}, \quad P(B \cap C) = \frac{1}{8}.$$

Oysa her çift için olasılıkların çarpımı $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$ 'tür ve

$$\frac{1}{8} \neq \frac{1}{4}, \quad \frac{3}{8} \neq \frac{1}{4}, \quad \frac{1}{8} \neq \frac{1}{4}$$

olduğundan üç çiftin hiçbirisi bağımsız değildir. Demek ki A, B ve C üçlü bağımsızdır, ama ikili bağımsız değildir; dolayısıyla tam bağımsız da değildir.

■

Son olarak, üçlü koşul ile ikili koşulların ikisi sağlanırken üçüncü ikili koşul da sağlanmayabilir. Yine $\Omega = \{1, 2, \dots, 8\}$ üzerinde eşit olasılıklı çekilişte $A = \{1, 2, 3, 4\}$, $B = \{1, 2, 3, 5\}$ ve $C = \{1, 4, 5, 6\}$ alınsın. $P(A) = P(B) = P(C) = \frac{1}{2}$, $A \cap C = \{1, 4\}$, $B \cap C = \{1, 5\}$ ve $A \cap B \cap C = \{1\}$ olduğundan

$$P(A \cap C) = P(B \cap C) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2}, \quad P(A \cap B \cap C) = \frac{1}{8} = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}$$

olur; ama $A \cap B = \{1, 2, 3\}$ ve $P(A \cap B) = \frac{3}{8} \neq \frac{1}{4} = P(A) P(B)$ 'dir. Olayların adları değiştirilerek aynı şey her ikili koşul için söylenebilir.

Birinci örnek ile son örnek birlikte, $n = 3$ için yukarıdaki dört koşulun (Tanım 7.2) hiçbirinin ötekilerden çıkarılamayacağını gösterir: her koşul, öteki üçü sağlanırken bozulabilir. İkinci örnek ise daha uç bir durumu, üçlü koşul sağlanırken üç ikili koşulun birden bozulabileceğini ortaya koyar. Tam bağımsızlık, ancak bütün koşullar tek tek doğrulanarak ya da deneyin yapısından çıkarılarak kurulur. İkinci yolun tipik örneği, iadeli çekilişlerdir.

Örnek 7.4 (İadeli Çekilişte Üç Siyah Top). Bir torbada 3 siyah ve 5 beyaz top vardır. Torbadan iadeli olarak, yani her çekilen top torbaya geri konarak, art arda 3 top çekiliyor. Çekilen 3 topun da siyah olması olasılığı nedir?

Çözüm.

Topları 1'den 8'e kadar numaralalım; 1, 2, 3 siyah, 4, 5, 6, 7, 8 beyaz olsun. Her çekilişte torbada aynı sekiz top bulunduğundan deneyin örnek uzayı

$$\Omega = \{(\omega_1, \omega_2, \omega_3) : \omega_i \in \{1, 2, \dots, 8\}\}, \quad n(\Omega) = 8^3 = 512$$

dir ve her üçlü eşit olasılıklıdır: $P(E) = \frac{n(E)}{512}$.

S_i , i -inci çekilen topun siyah olması olayı olsun ($i = 1, 2, 3$):

$$S_i = \{(\omega_1, \omega_2, \omega_3) \in \Omega : \omega_i \in \{1, 2, 3\}\}.$$

S_1 'in elemanlarını sayalım: ω_1 için 3, ω_2 ve ω_3 için 8'er seçenek vardır; $n(S_1) = 3 \cdot 8 \cdot 8 = 192$. Aynı biçimde $n(S_2) = n(S_3) = 192$ olur ve

$$P(S_1) = P(S_2) = P(S_3) = \frac{192}{512} = \frac{3}{8}.$$

Olaylar tam bağımsızdır. $S_1 \cap S_2$ için ω_1 ve ω_2 'nin 3'er, ω_3 'ün 8 seçeneği vardır: $n(S_1 \cap S_2) = 3 \cdot 3 \cdot 8 = 72$ ve

$$P(S_1 \cap S_2) = \frac{72}{512} = \frac{9}{64} = \frac{3}{8} \cdot \frac{3}{8} = P(S_1) P(S_2).$$

Aynı hesap $S_1 \cap S_3$ ve $S_2 \cap S_3$ için de geçerlidir; olaylar ikili bağımsızdır. Üçlü kesişim için her koordinatın 3 seçeneği vardır: $n(S_1 \cap S_2 \cap S_3) = 3^3 = 27$ ve

$$P(S_1 \cap S_2 \cap S_3) = \frac{27}{512} = \left(\frac{3}{8}\right)^3 = P(S_1) P(S_2) P(S_3).$$

Demek ki S_1, S_2, S_3 tam bağımsızdır. İstenen olasılık

$$P(S_1 \cap S_2 \cap S_3) = \left(\frac{3}{8}\right)^3 = \frac{27}{512}$$

dir.

Karşılaştırma. Çekilişler iadesiz olsaydı, çarpma teoremiyle (Teorem 6.3)

$$P(S_1 \cap S_2 \cap S_3) = P(S_1) P(S_2 | S_1) P(S_3 | S_1 \cap S_2) = \frac{3}{8} \cdot \frac{2}{7} \cdot \frac{1}{6} = \frac{1}{56}$$

bulunurdu. İadeli çekilişte her çekiliş torbayı başlangıç durumuna döndürdüğünden koşullu olasılıklar koşulsuz olasılıklara eşittir ve çarpma teoremi, tam bağımsızlıktaki çarpım kuralına indirgenir. Genel olarak, iadeli çekilişlerde farklı çekilişlere ilişkin olaylar tam bağımsızdır; bu, iadeli çekilişleri iadesiz çekilişlerden ayıran temel özelliktir.

■

7.6 Bağımsızlık Hesapları

Bu bölümde öğrendiklerimizi birkaç hesapta birleştirelim. Aşağıdaki örneğin ilk iki şıkkı, bağımsızlığın “olaylar birbiriyle ilgisizdir” sezgisine her zaman uymadığını da gösterir: aynı olay bir olaydan bağımsız, ona çok benzeyen bir başka olaydan bağımlı olabilir.

Örnek 7.5 (Bağımsızlık Hesapları). Düzgün bir zar iki kez atılsın; A ilk atışın çift gelmesi, B toplamın 7 olması ve C toplamın 8 olması olayı olsun.

(a) A ile B bağımsız mıdır?

(b) A ile C bağımsız mıdır?

(c) Bir olasılık uzayında E ile F olayları $P(E) = \frac{1}{2}$, $P(F) = \frac{1}{3}$ ve $P(E \cup F) = \frac{2}{3}$ eşitliklerini sağlıyorsa bağımsız mıdır?

(d) Bağımsız E ile F olayları için $P(E) = \frac{2}{5}$ ve $P(F) = \frac{1}{2}$ olsun. $P(E \cup F)$, $P(E^c \cap F^c)$, $P(E | F)$ ve $P(F | E^c)$ olasılıklarını hesaplayınız.

Çözüm.

İlk iki şıkta $\Omega = \{(x, y) : x, y \in \{1, \dots, 6\}\}$ olup $n(\Omega) = 36$ 'dır ve bir olayın olasılığı, eleman sayısının 36'ya bölümüdür. Son iki şık ise herhangi bir olasılık uzayında geçerlidir.

(a) $A = \{(x, y) : x \in \{2, 4, 6\}\}$ olup $n(A) = 3 \cdot 6 = 18$ ve $P(A) = \frac{1}{2}$ 'dir. $B = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}$ olup $P(B) = \frac{6}{36} = \frac{1}{6}$ 'dır. Kesişim, ilk bileşeni çift olan toplamı 7 çiftlerinden oluşur:

$$A \cap B = \{(2, 5), (4, 3), (6, 1)\}, \quad P(A \cap B) = \frac{3}{36} = \frac{1}{12}.$$

$P(A) P(B) = \frac{1}{2} \cdot \frac{1}{6} = \frac{1}{12} = P(A \cap B)$ olduğundan A ile B **bağımsızdır**. Sezgisel açıklama: ilk zar ne gelirse gelsin, toplamın 7 olması için ikinci zarın alması gereken tam bir değer vardır; bu yüzden $P(B | A) = \frac{1}{6} = P(B)$.

(b) $C = \{(2, 6), (3, 5), (4, 4), (5, 3), (6, 2)\}$ olup $P(C) = \frac{5}{36}$ 'dır. Kesişim

$$A \cap C = \{(2, 6), (4, 4), (6, 2)\}, \quad P(A \cap C) = \frac{3}{36} = \frac{1}{12}$$

dir. Oysa

$$P(A)P(C) = \frac{1}{2} \cdot \frac{5}{36} = \frac{5}{72} \neq \frac{6}{72} = \frac{1}{12} = P(A \cap C).$$

Demek ki A ile C **bağımsız değildir**. Gerçekten $P(C | A) = \frac{1/12}{1/2} = \frac{1}{6} > \frac{5}{36} = P(C)$: ilk zarın çift gelmesi toplamın 8 olma olasılığını artırır, çünkü toplamı 8 yapan beş çiftin üçü çift sayıyla başlar.

(c) Toplama kuralı (Teorem 3.2) gereği

$$P(E \cap F) = P(E) + P(F) - P(E \cup F) = \frac{1}{2} + \frac{1}{3} - \frac{2}{3} = \frac{1}{6}$$

olur. $P(E)P(F) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6} = P(E \cap F)$ olduğundan olaylar **bağımsızdır**.

(d) Bağımsızlık gereği $P(E \cap F) = \frac{2}{5} \cdot \frac{1}{2} = \frac{1}{5}$. Toplama kuralıyla

$$P(E \cup F) = \frac{2}{5} + \frac{1}{2} - \frac{1}{5} = \frac{7}{10}.$$

Teorem 7.2 (c) gereği E^c ile F^c bağımsızdır; $P(E^c) = \frac{3}{5}$ ve $P(F^c) = \frac{1}{2}$ olduğundan

$$P(E^c \cap F^c) = \frac{3}{5} \cdot \frac{1}{2} = \frac{3}{10}.$$

Bu sonuç $P(E^c \cap F^c) = 1 - P(E \cup F) = 1 - \frac{7}{10} = \frac{3}{10}$ ile uyumludur. Önerme 7.2 gereği

$$P(E | F) = P(E) = \frac{2}{5}.$$

Son olarak Teorem 7.2 (a) gereği E^c ile F bağımsızdır ve $P(E^c) = \frac{3}{5} > 0$ olduğundan

$$P(F | E^c) = P(F) = \frac{1}{2}.$$

■

7.7 Alıştırmalar

Alıştırma 7.1 (Bağımsız Olaylar). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı olsun.

(a) $A \in \mathcal{U}$ olayının kendisinden bağımsız olması için gerek ve yeter koşulun $P(A) = 0$ ya da $P(A) = 1$ olduğunu gösteriniz.

(b) $A \subseteq B$ ve A ile B bağımsız ise $P(A) = 0$ ya da $P(B) = 1$ olduğunu gösteriniz.

(c) A ile B bağımsız ise $P(A \cup B) = 1 - P(A^c)P(B^c)$ olduğunu gösteriniz. $P(A) = \frac{1}{3}$ ve $P(B) = \frac{1}{4}$ için $P(A \cup B)$ 'yi hesaplayınız.

(d) $A_1, A_2, \dots, A_n$ tam bağımsız ise, bu olaylardan istenilen kadarının yerine tümleyeninin konmasıyla elde edilen ailenin de tam bağımsız olduğunu gösteriniz. Buradan

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = 1 - \prod_{i=1}^n (1 - P(A_i))$$

olduğunu çıkarınız. Üç atıcı bir hedefe birbirinden bağımsız olarak birer atış yapıyor; isabet olasılıkları sırasıyla $\frac{1}{2}$, $\frac{3}{5}$ ve $\frac{7}{10}$ 'dur. Hedefin en az bir kez vurulması olasılığı nedir?

(e) Düzgün bir zar iki kez atılsın. A ilk atışın tek, B ikinci atışın tek ve C toplamın tek olması olayı olsun. A, B, C olaylarının ikili bağımsız olduğunu, ama tam bağımsız olmadığını gösteriniz.

(f) 3 siyah ve 5 beyaz top bulunan torbadan (Örnek 7.4) iadeli olarak 3 top çekiliyor. Tam olarak bir topun siyah olması ve en az bir topun siyah olması olasılıklarını hesaplayınız.

(g) A, B, C tam bağımsız ise A ile $B \cup C$ olaylarının ve A ile $B \cap C$ olaylarının bağımsız olduğunu gösteriniz.

(h) A ile B ayrık, $P(A) = \frac{1}{3}$ ve $P(B) = \frac{1}{4}$ olsun. $P(A | B)$ ve $P(A | A \cup B)$ olasılıklarını hesaplayınız; A ile B 'nin bağımsız olmadığını doğrulayınız.

Çözüm.

(a) A 'nın kendisinden bağımsız olması, $A \cap A = A$ olduğundan, $P(A) = P(A \cap A) = P(A) P(A) = P(A)^2$ demektir. $P(A)^2 = P(A)$ eşitliği $P(A) (P(A) - 1) = 0$ biçiminde yazılır; bu da ancak $P(A) = 0$ ya da $P(A) = 1$ iken sağlanır. Tersine $P(A) \in \{0, 1\}$ ise $P(A)^2 = P(A)$ olduğundan A kendisinden bağımsızdır.

(b) $A \subseteq B$ olduğundan $A \cap B = A$ 'dır. Bağımsızlık gereği $P(A) = P(A \cap B) = P(A) P(B)$, yani $P(A) (1 - P(B)) = 0$ olur. Bir çarpım sıfırda çarpanlardan biri sıfırdır: ya $P(A) = 0$ ya da $P(B) = 1$.

(c) Teorem 7.2 (c) gereği A^c ile B^c bağımsızdır. De Morgan kuralıyla $(A \cup B)^c = A^c \cap B^c$ olduğundan

$$P(A \cup B) = 1 - P(A^c \cap B^c) = 1 - P(A^c) P(B^c).$$

$$P(A) = \frac{1}{3}, P(B) = \frac{1}{4} \text{ için } P(A^c) = \frac{2}{3}, P(B^c) = \frac{3}{4} \text{ ve}$$

$$P(A \cup B) = 1 - \frac{2}{3} \cdot \frac{3}{4} = 1 - \frac{1}{2} = \frac{1}{2}.$$

Toplama kuralıyla doğrulama: $P(A \cup B) = \frac{1}{3} + \frac{1}{4} - \frac{1}{3} \cdot \frac{1}{4} = \frac{4+3-1}{12} = \frac{1}{2}$.

(d) Önce tek bir olayın, diyelim A_1 'in, yerine A_1^c konduğunda ailenin tam bağımsız kaldığını gösterelim. $I \subseteq \{1, \dots, n\}$ en az iki elemanlı bir indis kümesi olsun. $1 \notin I$ ise ilgili koşulda A_1 geçmez ve koşul varsayım gereği sağlanır. $1 \in I$ ise $J = I \setminus \{1\}$ diyelim ($J \neq \emptyset$) ve $D = \bigcap_{i \in J} A_i$ olsun. D olayı, ayrık $A_1 \cap D$ ve $A_1^c \cap D$ parçalarına ayrıldığından

$$P(A_1^c \cap D) = P(D) - P(A_1 \cap D) = \prod_{i \in J} P(A_i) - P(A_1) \prod_{i \in J} P(A_i) = (1 - P(A_1)) \prod_{i \in J} P(A_i) = P(A_1^c) \prod_{i \in J} P(A_i)$$

olur; burada $P(D)$ ve $P(A_1 \cap D)$ için tam bağımsızlık kullanıldı. Demek ki $A_1^c, A_2, \dots, A_n$ tam bağımsızdır. Indis adı önemli olmadığından aynı akıl yürütme herhangi bir A_j için geçerlidir; ayrıca sonuç her uygulamada tam bağımsız bir aile verdiğinden, işlem istenilen sayıda olay için art arda uygulanabilir. Böylece istenilen olayların tümleyeniyle değiştirilmesi tam bağımsızlığı korur.

Özel olarak $A_1^c, \dots, A_n^c$ tam bağımsızdır ve De Morgan kuralıyla

$$P(A_1 \cup \dots \cup A_n) = 1 - P(A_1^c \cap \dots \cap A_n^c) = 1 - \prod_{i=1}^n P(A_i^c) = 1 - \prod_{i=1}^n (1 - P(A_i)).$$

Atıcılar için A_i , i -inci atıcının isabet ettirmesi olayı olsun; A_1, A_2, A_3 tam bağımsızdır. Hedefin en az bir kez vurulması $A_1 \cup A_2 \cup A_3$ olayıdır ve

$$P(A_1 \cup A_2 \cup A_3) = 1 - \left(1 - \frac{1}{2}\right) \left(1 - \frac{3}{5}\right) \left(1 - \frac{7}{10}\right) = 1 - \frac{1}{2} \cdot \frac{2}{5} \cdot \frac{3}{10} = 1 - \frac{3}{50} = \frac{47}{50} = 0,94.$$

(e) $\Omega = \{(x, y) : x, y \in \{1, \dots, 6\}\}$, $n(\Omega) = 36$. A 'da x için 3, y için 6 seçenek vardır: $P(A) = \frac{18}{36} = \frac{1}{2}$; benzer biçimde $P(B) = \frac{1}{2}$. Toplamın tek olması için zarlardan biri tek, öteki çift olmalıdır: $n(C) = 3 \cdot 3 + 3 \cdot 3 = 18$ ve $P(C) = \frac{1}{2}$.

İkili kesişimler: $A \cap B$ iki zarın da tek gelmesidir, $n(A \cap B) = 9$; $A \cap C$ ilk zarın tek, toplamın tek, yani ikinci zarın çift gelmesidir, $n(A \cap C) = 9$; $B \cap C$ ikinci zarın tek, ilk zarın çift gelmesidir, $n(B \cap C) = 9$. Öyleyse

$$P(A \cap B) = P(A \cap C) = P(B \cap C) = \frac{9}{36} = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2},$$

yani olaylar ikili bağımsızdır.

Üçlü kesişim: iki zar da tek gelirse toplam çifttir; $A \cap B \cap C = \emptyset$ ve $P(A \cap B \cap C) = 0 \neq \frac{1}{8} = P(A) P(B) P(C)$. Demek ki olaylar üçlü bağımsız değildir, dolayısıyla tam bağımsız değildir.

(f) S_i , i -inci topun siyah olması olayı olsun; Örnek 7.4 gereği S_1, S_2, S_3 tam bağımsızdır ve $P(S_i) = \frac{3}{8}$, $P(S_i^c) = \frac{5}{8}$ 'dir. Tümleyenler de (d) şıkkı gereği tam bağımsız olduğundan, tam olarak bir siyah top çekilmesi olasılığı

$$P(S_1 \cap S_2^c \cap S_3^c) + P(S_1^c \cap S_2 \cap S_3^c) + P(S_1^c \cap S_2^c \cap S_3) = 3 \cdot \frac{3}{8} \cdot \frac{5}{8} \cdot \frac{5}{8} = \frac{225}{512}$$

dir (üç olay ayırıktır). En az bir siyah top çekilmesi, hiç siyah çekilmemesinin tümleyenidir:

$$P(S_1 \cup S_2 \cup S_3) = 1 - P(S_1^c \cap S_2^c \cap S_3^c) = 1 - \left(\frac{5}{8}\right)^3 = 1 - \frac{125}{512} = \frac{387}{512}.$$

(g) Tam bağımsızlık gereği $P(A \cap B) = P(A)P(B)$, $P(A \cap C) = P(A)P(C)$, $P(B \cap C) = P(B)P(C)$ ve $P(A \cap B \cap C) = P(A)P(B)P(C)$ 'dir. Dağılıma özelliği ile $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ olur ve toplama kuralı gereği

$$\begin{aligned} P(A \cap (B \cup C)) &= P(A \cap B) + P(A \cap C) - P(A \cap B \cap C) \\ &= P(A)P(B) + P(A)P(C) - P(A)P(B)P(C) \\ &= P(A)(P(B) + P(C) - P(B)P(C)) \\ &= P(A)(P(B) + P(C) - P(B \cap C)) = P(A)P(B \cup C). \end{aligned}$$

Demek ki A ile $B \cup C$ bağımsızdır. Öte yandan

$$P(A \cap (B \cap C)) = P(A \cap B \cap C) = P(A)P(B)P(C) = P(A)P(B \cap C)$$

olduğundan A ile $B \cap C$ de bağımsızdır.

(h) $A \cap B = \emptyset$ olduğundan $P(A \cap B) = 0$ ve

$$P(A | B) = \frac{P(A \cap B)}{P(B)} = \frac{0}{1/4} = 0.$$

Ayrıklık gereği $P(A \cup B) = P(A) + P(B) = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}$ ve $A \cap (A \cup B) = A$ olduğundan

$$P(A | A \cup B) = \frac{P(A)}{P(A \cup B)} = \frac{1/3}{7/12} = \frac{4}{7}.$$

$P(A \cap B) = 0 \neq \frac{1}{12} = P(A)P(B)$ olduğundan A ile B bağımsız değildir; bu, Teorem 7.1'in bir örneğidir. $P(A | B) = 0 \neq \frac{1}{3} = P(A)$ eşitsizliği de aynı sonucu verir.

■

Bu bölümle olasılık uzayları üzerine ilk kısmı tamamlamış olduk: olayları, olasılık ölçüsünü, koşullu olasılığı ve bağımsızlığı tanıdık. Bundan sonra olayları tek tek incelemek yerine, deneyin sonucuna sayısal değerler atayan fonksiyonları, yani rastgele değişkenleri ele alacağız. Kitabın ikinci kısmının ilk bölümü: [Rastgele Değişken](#).

8. Rastgele Değişken

Bağımsız Olaylar bölümüyle birlikte kitabın ilk kısmı tamamlandı. Elimizde artık bir olasılık uzayı $(\Omega, \mathcal{U}, P)$ var: rastgele deneyin bütün sonuçlarını toplayan örnek uzay Ω , olasılığı sorulabilecek olayları toplayan σ -cebiri $\mathcal{U}$ ve her olaya bir sayı karşılık getiren olasılık ölçüsü P . Bu üçlü, olasılık teorisinin bütün sorularının sorulduğu zemindir.

Ama uygulamada bir deneyin sonucunun kendisiyle nadiren ilgileniriz; ilgilendiğimiz şey çoğunlukla o sonuca bağlı bir **sayı**dır. Üç para atıldığında yazı ve turaların hangi sırayla geldiği değil, kaç tura geldiği; iki zar atıldığında hangi yüzlerin geldiği değil, üst yüzlerin toplamı; bir ampulün nasıl bozulduğu değil, kaç saat dayandığı sorulur. Bu sayıyı üreten şey, örnek uzayın her noktasına bir gerçel sayı karşılık getiren bir fonksiyondur: $X : \Omega \rightarrow \mathbb{R}$. Böyle bir fonksiyon, deneyin sonuçları hakkındaki soruları sayı doğrusundaki sorulara çevirir. “Kaç tura geldi?” sorusu “ X ’in değeri nedir?” sorusuna, “en çok bir tura gelmesi” olayı ise $\{\omega \in \Omega : X(\omega) \leq 1\}$ kümesine dönüşür.

Bu çevirinin işe yaraması için tek bir koşul gerekir: sayı doğrusunda sorulan her makul sorunun Ω ’daki karşılığı bir **olay** olmalı, yani $\mathcal{U}$ ’nun elemanı olmalıdır; aksi hâlde P o kümeye bir olasılık veremez. Bu koşulu sağlayan fonksiyonlara **rastgele değişken** denir. Bu bölümde rastgele değişkenin tanımını, olay kısaltmalarını ve değer kümesini vereceğiz; hangi fonksiyonların rastgele değişken olduğunu somut örneklerde inceleyeceğiz; tanımdaki koşulun aslında bütün Borel kümeleri için geçerli olduğunu söyleyen Borel ölçütünü ispatlayacağız; son olarak bir rastgele değişkenin $\mathbb{R}$ ’nin Borel cebri üzerinde yeni bir olasılık ölçüsü, yani bir **olasılık dağılımı** tanımladığını göreceğiz. Kitabın ikinci kısmının geri kalanı bu dağılımın incelenmesine ayrılmıştır.

8.1 Rastgele Değişkenin Tanımı

Birazdan vereceğimiz tanım yalnızca $\{\omega : X(\omega) \leq a\}$ kümelerinin olay olmasını ister; bu koşulun biçimi ilk bakışta keyfi görünebilir. Neden yalnızca bu kümeler? Neden pratiktir. Sayı doğrusunda sorulabilecek soruların en basiti “ X en çok a mıdır?” sorusudur; bu bölümün ilerleyen kısmında (**Önerme 8.2**) bu basit soruların olay olmasının, “ X bir Borel kümesine düşer mi?” biçimindeki bütün soruların olay olmasını zaten gerektirdiğini göreceğiz. Yani tanımda en az şey istenir, en çok şey elde edilir.

Tanım 8.1 (Rastgele Değişken). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı (**Tanım 3.2**) ve

$$X : \Omega \rightarrow \mathbb{R}, \quad \omega \mapsto X(\omega)$$

bir fonksiyon olsun. Her $a \in \mathbb{R}$ için

$$(X \leq a) = \{\omega \in \Omega : X(\omega) \leq a\} \in \mathcal{U}$$

ise X fonksiyonuna bir **rastgele değişken** (rastgele, tesadüfi, raslantı ya da rassal değişken de denir) adı verilir.

Tanımla ilgili birkaç gözlem:

- Rastgele değişkenler genellikle $X, Y, Z, V, W, \dots$ gibi büyük harflerle, aldıkları değerler ise $x, y, z, \dots$ gibi küçük harflerle gösterilir.
- Adındaki “rastgele” sözcüğüne karşın X **rastgele bir şey değildir**; Ω ’dan $\mathbb{R}$ ’ye tümüyle belirli, sıradan bir fonksiyondur. Rastgele olan, deney sonucunda hangi ω ’nın gerçekleşeceğini; ω belli olunca $X(\omega)$ sayısı da bellidir.
- Bir fonksiyonun rastgele değişken olup olmaması yalnızca Ω ’ya ve σ -cebiri $\mathcal{U}$ ’ya bağlıdır; olasılık ölçüsü P ’nin tanımda hiçbir rolü yoktur. $\mathcal{U} = \mathcal{P}(\Omega)$ ise her alt küme bir olay olduğundan, Ω üzerindeki **her** gerçel değerli fonksiyon bir rastgele değişkendir. Koşulun gerçekten kısıtlayıcı olduğu durumlar, $\mathcal{U}$ ’nun kuvvet kümesinden küçük olduğu durumlardır; bunları aşağıda göreceğiz.

Tanımda görülen $(X \leq a)$ yazımı, bölüm boyunca kullanacağımız bir kısaltma ailesinin ilk örneğidir. Bu kısaltmaları ve bunların küme işlemleriyle nasıl uyduğunu şimdi toplu hâlde verelim; teoremlerin ispatlarında hep bunlara başvuracağız.

Önerme 8.1 (Olay Kısaltmaları ve Ters Görüntü Özellikleri). $X : \Omega \rightarrow \mathbb{R}$ bir fonksiyon ve $A \subseteq \mathbb{R}$ olsun. A 'nın X altındaki **ters görüntüsü** kısaca

$$(X \in A) = \{\omega \in \Omega : X(\omega) \in A\} = X^{-1}(A)$$

ile gösterilir. Özel olarak $a, b \in \mathbb{R}$, $a < b$ için

$$\begin{aligned} (X = a) &= \{\omega \in \Omega : X(\omega) = a\}, & (X \leq a) &= \{\omega \in \Omega : X(\omega) \leq a\}, \\ (X < a) &= \{\omega \in \Omega : X(\omega) < a\}, & (X > a) &= \{\omega \in \Omega : X(\omega) > a\}, \\ (X \geq a) &= \{\omega \in \Omega : X(\omega) \geq a\}, & (a < X \leq b) &= \{\omega \in \Omega : a < X(\omega) \leq b\} \end{aligned}$$

yazılır; $(a \leq X \leq b)$, $(a < X < b)$, $(a \leq X < b)$ kısaltmaları da aynı biçimde anlaşılır. Bu kümeler $\mathcal{U}$ 'nun elemanı olduğunda olasılıkları $P(X \in A)$, $P(X = a)$, $P(a < X \leq b)$ biçiminde yazılır. Ters görüntü küme işlemleriyle uyur: $A, B, A_1, A_2, \dots \subseteq \mathbb{R}$ için

- (i) $(X \in A^c) = (X \in A)^c$; burada soldaki tümleyen $\mathbb{R}$ 'ye ($A^c = \mathbb{R} \setminus A$), sağdaki Ω 'ya göre alınır;
- (ii) $(X \in \bigcup_{n=1}^{\infty} A_n) = \bigcup_{n=1}^{\infty} (X \in A_n)$ ve $(X \in \bigcap_{n=1}^{\infty} A_n) = \bigcap_{n=1}^{\infty} (X \in A_n)$;
- (iii) $(X \in A \setminus B) = (X \in A) \setminus (X \in B)$ ve $A \subseteq B$ ise $(X \in A) \subseteq (X \in B)$;
- (iv) $A \cap B = \emptyset$ ise $(X \in A) \cap (X \in B) = \emptyset$.

İspat.

Bütün maddeler tanımın doğrudan sonucudur; hepsi " ω sol taraftadır $\Leftrightarrow \omega$ sağ taraftadır" biçiminde gösterilir.

(i) Tümleyen tanımını iki yanda da açalım:

$$\begin{aligned} \omega \in (X \in A^c) &\Leftrightarrow X(\omega) \in A^c \Leftrightarrow X(\omega) \notin A \\ &\Leftrightarrow \omega \notin (X \in A) \Leftrightarrow \omega \in (X \in A)^c. \end{aligned}$$

(ii) $\omega \in (X \in \bigcup_n A_n) \Leftrightarrow X(\omega) \in \bigcup_n A_n \Leftrightarrow$ en az bir n için $X(\omega) \in A_n \Leftrightarrow$ en az bir n için $\omega \in (X \in A_n) \Leftrightarrow \omega \in \bigcup_n (X \in A_n)$. Kesişim için "en az bir n " yerine "her n " yazılır.

(iii) $A \setminus B = A \cap B^c$ olduğundan (i) ve (ii)'den $(X \in A \setminus B) = (X \in A) \cap (X \in B)^c = (X \in A) \setminus (X \in B)$ bulunur. $A \subseteq B$ ise $X(\omega) \in A$ olan her ω için $X(\omega) \in B$ olur; bu da $(X \in A) \subseteq (X \in B)$ demektir.

(iv) (ii)'ye göre $(X \in A) \cap (X \in B) = (X \in A \cap B) = (X \in \emptyset)$ olur. Hiçbir ω için $X(\omega) \in \emptyset$ olamayacağından bu küme boştur.

■

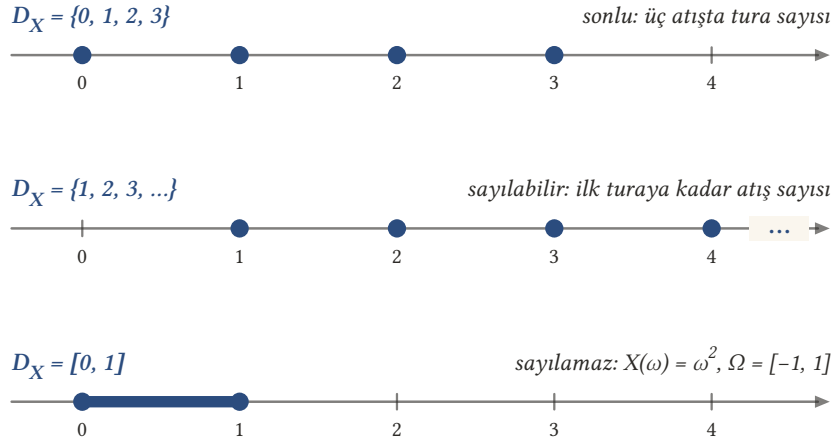
Önermenin (iv) maddesi, ayrık kümelerin ters görüntülerinin ayrık olaylar olduğunu söyler; bu, ileride σ -toplamsallığı aktarırken kilit rol oynayacak. Şimdi bir rastgele değişkenin hangi sayıları alabildiğini adlandıralım.

Tanım 8.2 (Değer Kümesi). $X : \Omega \rightarrow \mathbb{R}$ bir rastgele değişken olsun. X 'in görüntü kümesi

$$D_X = X(\Omega) = \{x \in \mathbb{R} : \text{bir } \omega \in \Omega \text{ için } X(\omega) = x\}$$

kümesine X 'in **değer kümesi** denir.

X değer kümesinin dışında hiçbir değer alamaz: $B \subseteq \mathbb{R}$ için $(X \in B) = (X \in B \cap D_X)$ olur, çünkü $X(\omega) \in B$ ise zaten $X(\omega) \in D_X$ 'tir. Bu basit gözlem hesaplarda çok iş görür; D_X sonlu olduğunda her sorunun yanıtı sonlu sayıda noktaya bakılarak verilir. Değer kümesi sonlu ya da sayılabilir olan rastgele değişkenlere ileride **kesikli** rastgele değişken diyeceğiz (Tanım 10.1); değer kümesi bir aralık olanların önemli bir sınıfı ise olasılığı bir yoğunluk fonksiyonunun integraliyle verilen **sürekli** rastgele değişkenlerdir (Tanım 11.2). Şimdilik her iki türden de örnek görelim.



Şekil 8.1: Üç rastgele değişkenin değer kümesi D_X aynı sayı doğrusunda. Üç para atışında tura sayısı sonlu bir küme, ilk turaya kadar yapılan atış sayısı sayılabilir sonsuz bir küme, $[-1, 1]$ üzerinde ω^2 ise bir aralık verir. İlk iki tür kesikli rastgele değişkenlerin, üçüncüsü sürekli rastgele değişkenlerin örneğidir; X değer kümesinin dışında hiçbir değer almaz.

8.2 İlk Örnekler

Aşağıdaki üç örnekte bir fonksiyonun rastgele değişken olduğunu tanımın istediği gibi, yani a 'nın bütün gerçel değerleri için $(X \leq a)$ kümesini açıkça bulup $\mathcal{U}$ 'da olduğunu görerek göstereceğiz. a , değer kümesinin altında mı, iki değer arasında mı, yoksa üstünde mi kalıyor; $(X \leq a)$ kümesi buna göre değişir. Bu yüzden hesap her zaman a eksen aralıklara bölünerek yapılır.

Örnek 8.1 (Bir Para Atışında Tura Sayısı). Bir para bir kez atılıyor. Örnek uzay $\Omega = \{Y, T\}$ (Y yazı, T tura), $\mathcal{U} = \mathcal{P}(\Omega)$ ve her $A \in \mathcal{U}$ için $P(A) = \frac{n(A)}{n(\Omega)}$ (Laplace uzayı, Tanım 4.1) olmak üzere $(\Omega, \mathcal{U}, P)$ olasılık uzayı alınsın.

$$X : \Omega \rightarrow \mathbb{R}, \quad X(\omega) = \begin{cases} 0, & \omega = Y \\ 1, & \omega = T \end{cases}$$

fonksiyonunun bir rastgele değişken olduğunu gösteriniz; değer kümesini ve $P(X \leq 0)$, $P(X = 1)$ olasılıklarını bulunuz.

Çözüm.

X yalnızca 0 ve 1 değerlerini aldığından $D_X = \{0, 1\}$ 'dir. $a \in \mathbb{R}$ için $(X \leq a) = \{\omega \in \Omega : X(\omega) \leq a\}$ kümesini a 'nın konumuna göre bulalım.

- $a < 0$ ise hiçbir ω için $X(\omega) \leq a$ olamaz, çünkü X 'in en küçük değeri 0'dır: $(X \leq a) = \emptyset \in \mathcal{U}$.
- $0 \leq a < 1$ ise $X(\omega) \leq a$ koşulunu yalnızca $X(\omega) = 0$ sağlar: $(X \leq a) = \{Y\} \in \mathcal{U}$.
- $a \geq 1$ ise her iki değer de a 'yı aşmaz: $(X \leq a) = \{Y, T\} = \Omega \in \mathcal{U}$.

Her $a \in \mathbb{R}$ için $(X \leq a) \in \mathcal{U}$ olduğundan X bir rastgele değişkendir. X , atışta gelen **tura sayısını** sayar. Olasılıklar Laplace formülüyle (Önerme 4.2)

$$P(X \leq 0) = P(\{Y\}) = \frac{1}{2}, \quad P(X = 1) = P(\{T\}) = \frac{1}{2}$$

olarak bulunur.

■

Aynı yöntem birden çok atış için de geçerlidir; yalnızca incelenecek aralık sayısı artar.

Örnek 8.2 (Üç Para Atışında Tura Sayısı). Bir para üç kez atılıyor. Örnek uzay

$$\Omega = \{YYY, YYT, YTY, YTT, TYY, TYT, TTY, TTT\},$$

$\mathcal{U} = \mathcal{P}(\Omega)$ ve $P(A) = \frac{n(A)}{8}$ olmak üzere $(\Omega, \mathcal{U}, P)$ olasılık uzayı alınsın. $X : \Omega \rightarrow \mathbb{R}$ fonksiyonu her sonuçtaki tura sayısı olsun:

$$X(\omega) = \begin{cases} 0, \omega = YYY \\ 1, \omega \in \{YYT, YTY, TYY\} \\ 2, \omega \in \{YTT, TYT, TTY\} \\ 3, \omega = TTT. \end{cases}$$

X 'in bir rastgele değişken olduğunu gösteriniz; $P(X \leq 0)$, $P(X \leq 1)$, $P(X = 2)$ olasılıklarını hesaplayınız ve $x \in D_X$ için $P(X = x)$ değerlerini bir tabloda veriniz.

Çözüm.

$D_X = \{0, 1, 2, 3\}$ 'tür. $a \in \mathbb{R}$ için $(X \leq a)$ kümesi, X değeri a 'yı aşmayan sonuçların kümesidir:

- $a < 0$ ise $(X \leq a) = \emptyset \in \mathcal{U}$.
- $0 \leq a < 1$ ise $(X \leq a) = (X = 0) = \{YYY\} \in \mathcal{U}$.
- $1 \leq a < 2$ ise $(X \leq a) = (X = 0) \cup (X = 1) = \{YYY, YYT, YTY, TYY\} \in \mathcal{U}$.
- $2 \leq a < 3$ ise $(X \leq a) = (X = 0) \cup (X = 1) \cup (X = 2) = \Omega \setminus \{TTT\} \in \mathcal{U}$.
- $a \geq 3$ ise $(X \leq a) = \Omega \in \mathcal{U}$.

$2 \leq a < 3$ durumundaki birleşimin elemanları açıkça şunlardır:

$$(X = 0) \cup (X = 1) \cup (X = 2) = \{YYY, YYT, YTY, TYY, YTT, TYT, TTY\}.$$

Demek ki X bir rastgele değişkendir. Şimdi istenen olasılıklar, olay kısaltmalarını açıp Laplace formülünü uygulayarak bulunur:

$$P(X \leq 0) = P(\{\omega \in \Omega : X(\omega) \leq 0\}) = P(\{YYY\}) = \frac{1}{8},$$

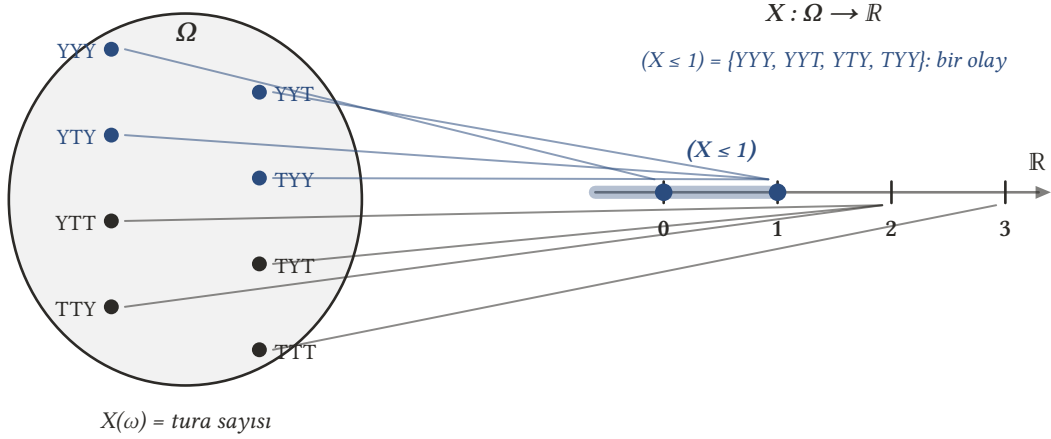
$$P(X \leq 1) = P(\{YYY, YYT, YTY, TYY\}) = \frac{4}{8} = \frac{1}{2},$$

$$P(X = 2) = P(\{YTT, TYT, TTY\}) = \frac{3}{8}.$$

Değer kümesinin her noktası için $P(X = x)$ olasılığı, x tura içeren sonuçların sayısının 8'e bölümüdür:

x	0	1	2	3
$P(X = x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Tablodaki dört sayının toplamı 1'dir; bu bir rastlantı değildir, çünkü $(X = 0)$, $(X = 1)$, $(X = 2)$, $(X = 3)$ olayları ayrıktır ve birleşimleri Ω 'dır. Böyle bir tablo, bir rastgele değişkenin **olasılık fonksiyonu** kavramının ilk örneğidir (Tanım 10.2).



Şekil 8.2: Rastgele değişken, örnek uzayın her sonucuna bir sayı eşleyen fonksiyondur; burada üç para atışında tura sayısı. " $X \leq 1$ " gibi bir koşul, sayı doğrusunda bir yarı doğruya, örnek uzayda ise oraya giden sonuçların kümesine karşılık gelir. Tanım, bu kümelerin hepsinin birer olay (σ -cebrin elemanı) olmasını ister; olasılıkları ancak o zaman anlamlıdır.

■

Örnek uzay sonsuz, hatta sayılamaz olduğunda da tanım aynı biçimde uygulanır; yalnızca $(X \leq a)$ kümeleri artık sonlu listeler değil, aralıklardır.

Örnek 8.3 (Kare Dönüşümü). $\Omega = [-1, 1]$, $\mathcal{U} = \mathcal{B}(\Omega)$ (Ω 'nın Borel alt kümelerinin σ -cebiri) ve $A \in \mathcal{U}$ için

$$P(A) = \frac{\text{uzunluk}(A)}{\text{uzunluk}(\Omega)} = \frac{\text{uzunluk}(A)}{2}$$

olmak üzere $(\Omega, \mathcal{U}, P)$ olasılık uzayı alınsın (geometrik olasılık, [Tanım 4.5](#)). $X : \Omega \rightarrow \mathbb{R}$, $X(\omega) = \omega^2$ fonksiyonunun bir rastgele değişken olduğunu gösteriniz; değer kümesini bulunuz ve $0 \leq a \leq 1$ için $P(X \leq a)$ olasılığını hesaplayınız.

Çözüm.

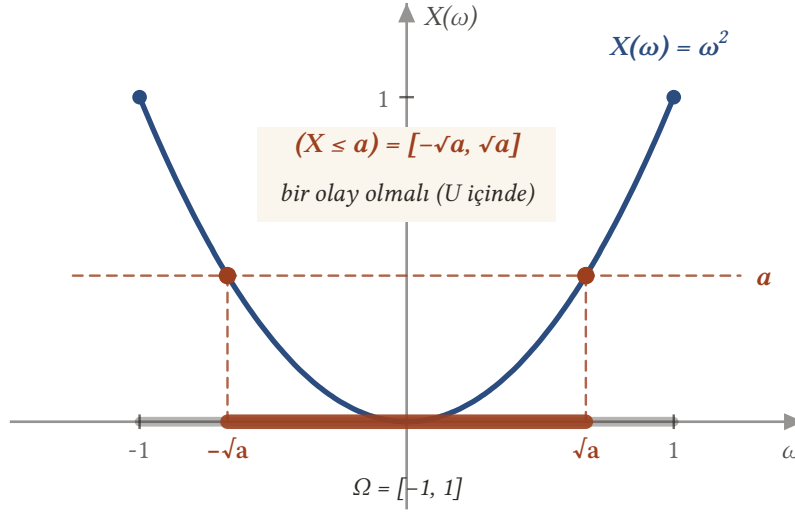
$\omega \in [-1, 1]$ iken $\omega^2 \in [0, 1]$ olur ve $[0, 1]$ 'deki her x değeri $\omega = \sqrt{x} \in \Omega$ noktasında alınır; dolayısıyla $D_X = [0, 1]$ 'dir. $a \in \mathbb{R}$ için $(X \leq a) = \{\omega \in [-1, 1] : \omega^2 \leq a\}$ kümesini bulalım.

- $a < 0$ ise $\omega^2 \leq a$ eşitsizliğinin çözümü yoktur: $(X \leq a) = \emptyset \in \mathcal{U}$.
- $0 \leq a \leq 1$ ise $\omega^2 \leq a \Leftrightarrow -\sqrt{a} \leq \omega \leq \sqrt{a}$ olur ve $\sqrt{a} \leq 1$ olduğundan bu aralık Ω 'nın içindedir:

$$(X \leq a) = \{\omega \in [-1, 1] : -\sqrt{a} \leq \omega \leq \sqrt{a}\} = [-\sqrt{a}, \sqrt{a}].$$

Kapalı bir aralık bir Borel kümesidir ([Önerme 2.5](#)), dolayısıyla $[-\sqrt{a}, \sqrt{a}] \in \mathcal{U}$.

- $a > 1$ ise Ω 'nın her noktasında $\omega^2 \leq 1 < a$ olduğundan $(X \leq a) = \Omega \in \mathcal{U}$.



Şekil 8.3: Rastgele değişken koşulunun geometrisi: $\Omega = [-1, 1]$ üzerinde $X(\omega) = \omega^2$ alınmıştır. " $X \leq a$ " sorusu, eğrinin a düzeyinin altında kaldığı ω noktalarını, yani $[-\sqrt{a}, \sqrt{a}]$ aralığını seçer. Tanım, her a için bu ters görüntünün σ -cebir $\mathcal{U}$ 'nin bir elemanı olmasını ister; burada kapalı aralık bir Borel kümesi olduğundan koşul sağlanır ve $P(X \leq a) = \sqrt{a}$ olur.

Her a için $(X \leq a) \in \mathcal{U}$ olduğundan X bir rastgele değişkendir. $0 \leq a \leq 1$ için

$$P(X \leq a) = P([-\sqrt{a}, \sqrt{a}]) = \frac{\sqrt{a} - (-\sqrt{a})}{2} = \sqrt{a}$$

bulunur. Örneğin $P(X \leq \frac{1}{4}) = \frac{1}{2}$ ve $P(\frac{1}{4} < X \leq 1) = 1 - \frac{1}{2} = \frac{1}{2}$ 'dir. Burada, önceki iki örnekten farklı olarak, her a için tek bir noktanın olasılığı sıfırdır: $0 < a \leq 1$ iken $(X = a) = \{-\sqrt{a}, \sqrt{a}\}$, $a = 0$ iken $(X = 0) = \{0\}$, öteki a değerleri için ise $(X = a) = \emptyset$ 'dir; sonlu bir kümenin uzunluğu sıfır olduğundan her durumda $P(X = a) = 0$ bulunur.

■

8.3 Hangi Fonksiyonlar Rastgele Değişkendir?

Yukarıdaki örneklerin ilk ikisinde $\mathcal{U} = \mathcal{P}(\Omega)$ olduğundan koşulun sağlanacağı baştan belliydi; üçüncüsünde $\mathcal{U}$ Borel cebriydi ve koşul yine sağlandı. Koşulun gerçekten eleyici olduğunu görmek için σ -cebirin kuvvet kümesinden **küçük** olduğu bir uzay almak gerekir. Aşağıdaki örnek bunu yapar: dört fonksiyondan ikisi rastgele değişkendir, ikisi değildir.

Örnek 8.4 (Hangileri Rastgele Değişkendir). $\Omega = \{\omega_1, \omega_2, \omega_3\}$,

$$\mathcal{U} = \{ \emptyset, \{\omega_1, \omega_2\}, \{\omega_3\}, \Omega \}$$

ve $P(\{\omega_3\}) = \frac{1}{2}$ (dolayısıyla $P(\{\omega_1, \omega_2\}) = \frac{1}{2}$) olmak üzere $(\Omega, \mathcal{U}, P)$ olasılık uzayı verilsin. Aşağıdaki fonksiyonlardan hangileri rastgele değişkendir?

- a) $X(\omega_i) = i$, $i = 1, 2, 3$;
- b) $Y(\omega_i) = 0$, $i = 1, 2, 3$;
- c) $Z(\omega_1) = Z(\omega_2) = 1$, $Z(\omega_3) = 0$;
- d) $V(\omega_1) = V(\omega_2) = 1$, $V(\omega_3) = 0$.

Çözüm.

Önce $\mathcal{U}$ 'nin gerçekten bir σ -cebir olduğunu not edelim: $\{\omega_1, \omega_2\}$ ile $\{\omega_3\}$ birbirinin tümleyenidir, $\emptyset$ ile Ω da öyledir ve bu dört kümenin her birleşimi yine bu dört kümeden biridir. $\mathcal{U}$ 'nin elemanı **olmayan** alt kümeler $\{\omega_1\}$, $\{\omega_2\}$, $\{\omega_1, \omega_3\}$ ve $\{\omega_2, \omega_3\}$ 'tür. Bir fonksiyonun rastgele değişken olmaması için

tek bir a değerinde $(X \leq a)$ 'nın bu dört kümeden birine eşit çıkması yeter. Olasılık ölçüsü P 'nin bu soruda hiçbir rolü yoktur.

a) $X(\omega_1) = 1, X(\omega_2) = 2, X(\omega_3) = 3$ olduğundan $D_X = \{1, 2, 3\}$ 'tür. $a = 1$ alalım:

$$(X \leq 1) = \{\omega \in \Omega : X(\omega) \leq 1\} = \{\omega_1\} \notin \mathcal{U}.$$

Tek bir a için koşul bozulduğundan X **rastgele değişken değildir**. (Öteki aralıklara bakmaya gerek yoktur; ama bakılırsa $a < 1$ için $\emptyset, 1 \leq a < 2$ için $\{\omega_1\}, 2 \leq a < 3$ için $\{\omega_1, \omega_2\}, a \geq 3$ için Ω bulunur; yalnızca $1 \leq a < 2$ aralığı sorun çıkarır.)

b) Y sabit sıfır fonksiyonudur, $D_Y = \{0\}$. $a \in \mathbb{R}$ için

$$(Y \leq a) = \begin{cases} \emptyset, & a < 0 \\ \Omega, & a \geq 0 \end{cases}$$

olur ve $\emptyset, \Omega \in \mathcal{U}$ 'dur. Demek ki Y **bir rastgele değişkendir**. Aynı akıl yürütme her sabit fonksiyon için geçerlidir: sabit fonksiyonlar her olasılık uzayında rastgele değişkendir.

c) $D_Z = \{0, 1\}$ 'dir ve $Z, 0$ değerini yalnızca ω_3 'te alır. $a \in \mathbb{R}$ için

$$(Z \leq a) = \begin{cases} \emptyset, & a < 0 \\ \{\omega_3\}, & 0 \leq a < 1 \\ \Omega, & a \geq 1. \end{cases}$$

Üç küme de $\mathcal{U}$ 'nun elemanıdır; Z **bir rastgele değişkendir**.

d) $D_V = \{0, 1\}$ 'dir ve $V, 0$ değerini yalnızca ω_2 'de alır. $a = 0$ için

$$(V \leq 0) = \{\omega \in \Omega : V(\omega) \leq 0\} = \{\omega_2\} \notin \mathcal{U}.$$

Dolayısıyla V **rastgele değişken değildir**. ($0 \leq a < 1$ aralığının tamamında $(V \leq a) = \{\omega_2\}$ olur.) Sonuç: yalnızca Y ve Z rastgele değişkendir. Z ile V arasındaki fark öğreticidir: ikisi de yalnızca 0 ve 1 değerlerini alır, ama Z değerlerini $\mathcal{U}$ 'daki kümeler üzerinde sabit tutarken $V, \mathcal{U}$ 'nun bölemediği $\{\omega_1, \omega_2\}$ kümesini ikiye ayırmaya çalışır. σ -cebiri bu ayrımı "göremediği" için V 'ye olasılık sorulamaz: $P(V = 0) = P(\{\omega_2\})$ tanımlı değildir.

■

i Rastgele değişken olmak yalnızca fonksiyona değil, σ -cebire de bağlıdır

Aynı $X : \Omega \rightarrow \mathbb{R}$ fonksiyonu bir σ -cebire göre rastgele değişken olabilir, bir başkasına göre olmaya-bilir. İki uç durum bunu iyi gösterir.

- $\mathcal{U} = \mathcal{P}(\Omega)$ (en büyük σ -cebiri) ise $(X \leq a)$ kümesi ne olursa olsun $\mathcal{U}$ 'dadır; Ω üzerindeki **her** fonksiyon rastgele değişkendir.
- $\mathcal{U} = \{\emptyset, \Omega\}$ (en küçük σ -cebiri) ise **yalnızca sabit** fonksiyonlar rastgele değişkendir. Gerçekten, X iki farklı $c_1 < c_2$ değeri alıyorsa $(X \leq c_1)$ kümesi c_1 'i veren noktayı içerir, c_2 'yi veren noktayı içermez; yani ne boştur ne Ω 'dır, dolayısıyla $\mathcal{U}$ 'da değildir.

Örnek 8.4 bu iki ucun arasındadır: $\mathcal{U}, \Omega$ 'yı yalnızca $\{\omega_1, \omega_2\}$ ve $\{\omega_3\}$ parçalarına ayırabildiğinden, rastgele değişkenler tam olarak bu parçaların her biri üzerinde sabit olan fonksiyonlardır. Genel kural da budur: **bir fonksiyon, ancak σ -cebirin ayırt edebildiği kadar ayrıntı taşıyorsa rastgele değişkendir**.

8.4 Borel Ölçütü

Tanım yalnızca $(X \leq a)$ kümelerinin olay olmasını ister. Oysa uygulamada $(X = a), (a < X \leq b), (X > a)$ gibi kümelerin, hatta " X rasyonel mi?" gibi soruların olasılığını da sormak isteriz. Aşağıdaki önerme bu kaygıyı ortadan kaldırır: tanımdaki koşul, $\mathbb{R}$ 'nin **bütün Borel kümeleri** için ters görüntünün olay olmasına denktir. Bu sonuca ihtiyacımız iki yönlüdür. Bir yandan olay kısaltmalarının hepsine olasılık verebilmemizi sağlar; öte yandan bu bölümün son kısmında $P_X(B) = P(X \in B)$ ifadesinin her Borel B için anlamlı olduğunu bu önerme garanti edecektir.

Önerme 8.2 (Borel Ölçütü). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı ve $X : \Omega \rightarrow \mathbb{R}$ bir fonksiyon olsun. Aşağıdaki iki koşul birbirine denktir:

(i) her $a \in \mathbb{R}$ için $(X \leq a) \in \mathcal{U}$, yani X bir rastgele değişkendir;

(ii) her Borel kümesi $B \in \mathcal{B}(\mathbb{R})$ için $(X \in B) = X^{-1}(B) \in \mathcal{U}$.

Özel olarak X bir rastgele değişken ve $a < b$ ise

$$(X < a), (X > a), (X \geq a), (X = a), (a < X \leq b), (a \leq X \leq b), (a < X < b), (a \leq X < b)$$

kümelerinin her biri bir olaydır.

İspat.

(ii) $\Rightarrow$ (i). Her $a \in \mathbb{R}$ için $(-\infty, a]$ yarı doğrusu bir Borel kümesidir (Önerme 2.5 (c)). (ii)'ye göre $(X \leq a) = (X \in (-\infty, a]) \in \mathcal{U}$ olur.

(i) $\Rightarrow$ (ii). Bu yön asıl içeriktir; üç adımda ispatlayacağız. Boyunca (i)'nin sağlandığını varsayıyoruz ve Önerme 8.1'ndeki ters görüntü özelliklerini kullanıyoruz.

Adım 1: Yarı doğrular, aralıklar ve tek noktalar. $\mathcal{U}$ bir σ -cebiri olduğundan tümleyene, sayılabilir birleşime ve kesişime kapalıdır (Teorem 2.1). $a < b$ olsun.

- $(X > a) = (X \leq a)^c \in \mathcal{U}$.
- $(X < a) = \bigcup_{n=1}^{\infty} (X \leq a - \frac{1}{n})$. Gerçekten, $X(\omega) < a$ ise $a - X(\omega) > 0$ olduğundan $\frac{1}{n} \leq a - X(\omega)$ olacak biçimde bir n vardır, yani $X(\omega) \leq a - \frac{1}{n}$; tersine $X(\omega) \leq a - \frac{1}{n}$ ise $X(\omega) < a$ 'dır. Sağ taraf $\mathcal{U}$ 'daki kümelerin sayılabilir birleşimi olduğundan $(X < a) \in \mathcal{U}$.
- $(X \geq a) = (X < a)^c \in \mathcal{U}$.
- $(X = a) = (X \leq a) \cap (X \geq a) \in \mathcal{U}$.
- $(a < X \leq b) = (X \leq b) \cap (X > a) \in \mathcal{U}$; benzer biçimde $(a \leq X \leq b) = (X \leq b) \cap (X \geq a)$, $(a < X < b) = (X < b) \cap (X > a)$ ve $(a \leq X < b) = (X < b) \cap (X \geq a)$ kümeleri de $\mathcal{U}$ 'dadır.

Böylece önermenin son cümlesi ispatlanmış oldu. Aynı zamanda şunu göstermiş olduk: $I \subseteq \mathbb{R}$ hangi türden olursa olsun bir aralıksa $((-\infty, a], (-\infty, a), (a, \infty), [a, \infty), (a, b], [a, b], (a, b), [a, b))$ ya da tek noktalı $\{a\} = [a, a]$ $(X \in I) \in \mathcal{U}$ 'dur.

Adım 2: Ters görüntüsü olay olan kümelerin sınıfı bir σ -cebirdir.

$$\mathcal{D} = \{B \subseteq \mathbb{R} : (X \in B) \in \mathcal{U}\}$$

sınıfını tanımlayalım. $\mathcal{D}$ 'nin $\mathbb{R}$ üzerinde bir σ -cebiri olduğunu (Tanım 2.2) gösterelim.

- $(X \in \mathbb{R}) = \Omega \in \mathcal{U}$ olduğundan $\mathbb{R} \in \mathcal{D}$.
- $B \in \mathcal{D}$ olsun. Önerme 8.1 (i)'ye göre $(X \in B^c) = (X \in B)^c$ ve $\mathcal{U}$ tümleyene kapalı olduğundan $(X \in B^c) \in \mathcal{U}$; yani $B^c \in \mathcal{D}$.
- $B_1, B_2, \dots \in \mathcal{D}$ olsun. Önerme 8.1 (ii)'ye göre $(X \in \bigcup_n B_n) = \bigcup_n (X \in B_n)$ ve $\mathcal{U}$ sayılabilir birleşime kapalı olduğundan bu küme $\mathcal{U}$ 'dadır; yani $\bigcup_n B_n \in \mathcal{D}$.

Demek ki $\mathcal{D}$ bir σ -cebiri. Adım 1'e göre $\mathbb{R}$ 'nin her aralığı $\mathcal{D}$ 'nin elemanıdır.

Adım 3: Borel cebri $\mathcal{D}$ 'nin içindedir. Borel cebri, açık aralıklar sınıfı $\mathcal{L}_1 = \{(a, b) : a < b\}$ 'in ürettiği σ -cebiri (Tanım 2.4): $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{L}_1)$, yani $\mathcal{L}_1$ 'i kapsayan en küçük σ -cebiri. Adım 1'e göre her açık aralık $\mathcal{D}$ 'nin elemanıdır, yani $\mathcal{L}_1 \subseteq \mathcal{D}$; Adım 2'ye göre $\mathcal{D}$ bir σ -cebiri. Teorem 2.3 (b) gereği $\mathcal{L}_1$ 'i kapsayan her σ -cebiri $\sigma(\mathcal{L}_1)$ 'i kapsar; öyleyse $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{D}$ 'dir. Bu da tam olarak (ii)'dir: her $B \in \mathcal{B}(\mathbb{R})$ için $(X \in B) \in \mathcal{U}$.

■

Önermenin iki pratik sonucu vardır. Birincisi, tanımda $(X \leq a)$ yerine $(X < a)$, $(X \geq a)$ ya da $(X > a)$ kümeleri kullanılsaydı aynı rastgele değişken kavramına ulaşıldı. Adım 1'deki eşitlikler, $(X \leq a)$ kümelerinin olay olmasından öteki üçünün olay olmasını çıkarır; ters yön için

$$(X \leq a) = (X > a)^c = \bigcap_{n=1}^{\infty} \left(X < a + \frac{1}{n} \right), \quad (X < a) = (X \geq a)^c$$

eşitlikleri yeter (ilk eşitlikteki kesişim, $X(\omega) \leq a$ ile "her n için $X(\omega) < a + \frac{1}{n}$ " koşullarının denk olmasından gelir). Böylece dört koşul birbirine denktir. İkincisi, X bir rastgele değişkense ve B herhangi bir Borel kümesiyse $P(X \in B)$ olasılığı anlamlıdır; $\mathbb{R}$ 'nin uygulamada karşılaşılan her alt kümesi (aralıklar,

sayılabilir kümeler, açık ve kapalı kümeler) Borel olduğundan, bir rastgele değişken hakkında sorulabilecek her makul sorunun bir olasılığı vardır.

💡 Sonlu değer kümesi için pratik ölçüt

D_X sonlu ya da sayılabilir ise X 'in rastgele değişken olması için gerek ve yeter koşul, her $x \in D_X$ için $(X = x) \in \mathcal{U}$ olmasıdır. Gereklik **Önerme 8.2**'ten gelir. Yeterlik için

$$(X \leq a) = \bigcup_{x \in D_X, x \leq a} (X = x)$$

yazmak yeter: sağ taraf $\mathcal{U}$ 'daki sonlu ya da sayılabilir çoklukta kümenin birleşimidir, dolayısıyla $\mathcal{U}$ 'dadır. **Örnek 8.4**'e bu gözle bakılırsa yanıt anında görülür: Z için $(Z = 0) = \{\omega_3\}$ ve $(Z = 1) = \{\omega_1, \omega_2\}$ olaydır; V için $(V = 0) = \{\omega_2\}$ olay değildir.

8.5 Rastgele Değişkenin Olasılık Dağılımı

Bir rastgele değişkenin asıl işlevi, olasılığı Ω 'dan sayı doğrusuna taşımaktır. Borel ölçütü sayesinde her Borel kümesi B için $P(X \in B)$ sayısı tanımlıdır; bu sayıya “ X 'in B 'ye düşme olasılığı” deriz. $B \mapsto P(X \in B)$ kuralı, $\mathbb{R}$ 'nin Borel kümeleri üzerinde tanımlı bir fonksiyondur. Aşağıdaki teorem bu fonksiyonun bir olasılık ölçüsü olduğunu söyler; yani X , elimizdeki $(\Omega, \mathcal{U}, P)$ uzayından $\mathbb{R}$ üzerinde yepyeni bir olasılık uzayı üretir. Bu teoreme ihtiyaç duymamızın nedeni açıktır: kitabın bundan sonraki bölümlerinde Ω 'yı hemen hemen unutup yalnızca sayı doğrusu üzerindeki bu ölçüyle çalışacağız; bunun meşru olması için o ölçünün Kolmogorov aksiyomlarını sağladığını bir kez kesin olarak görmemiz gerekir.

Teorem 8.1 (Rastgele Değişkenin Ürettiği Olasılık Ölçüsü). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı ve $X : \Omega \rightarrow \mathbb{R}$ bir rastgele değişken olsun.

$$P_X : \mathcal{B}(\mathbb{R}) \rightarrow \mathbb{R}, \quad P_X(B) = P(X \in B) = P(\{\omega \in \Omega : X(\omega) \in B\})$$

fonksiyonu $\mathcal{B}(\mathbb{R})$ üzerinde bir olasılık ölçüsüdür. Dolayısıyla $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$ bir olasılık uzayıdır.

İspat.

Önce P_X 'in iyi tanımlı olduğunu görelim: $B \in \mathcal{B}(\mathbb{R})$ için **Önerme 8.2**'e göre $(X \in B) \in \mathcal{U}$ 'dur, dolayısıyla $P(X \in B)$ sayısı tanımlıdır. Şimdi P_X 'in olasılık ölçüsü aksiyomlarını (**Tanım 3.1**) sağladığını gösterelim.

(A1) Negatif olmama. Her $B \in \mathcal{B}(\mathbb{R})$ için $(X \in B)$ bir olaydır ve P her olaya negatif olmayan bir sayı verir: $P_X(B) = P(X \in B) \geq 0$.

(A2) Normlama. X 'in her değeri $\mathbb{R}$ 'de olduğundan $(X \in \mathbb{R}) = \Omega$ 'dır; dolayısıyla $P_X(\mathbb{R}) = P(\Omega) = 1$.

(A3) Sayılabilir toplamsallık. $B_1, B_2, \dots \in \mathcal{B}(\mathbb{R})$ ikişer ikişer ayrık Borel kümeleri olsun. Borel cebri sayılabilir birleşime kapalı olduğundan $\bigcup_n B_n \in \mathcal{B}(\mathbb{R})$ 'dir ve $P_X(\bigcup_n B_n)$ tanımlıdır. **Önerme 8.1** (ii)'ye göre

$$\left(X \in \bigcup_{n=1}^{\infty} B_n \right) = \bigcup_{n=1}^{\infty} (X \in B_n).$$

$m \neq n$ için $B_m \cap B_n = \emptyset$ olduğundan **Önerme 8.1** (iv)'e göre $(X \in B_m) \cap (X \in B_n) = \emptyset$; yani $(X \in B_1), (X \in B_2), \dots$ olayları da ikişer ikişer ayrıktır. P 'nin sayılabilir toplamsallığı (aksiyom (A3)) bu ayrık olay dizisine uygulanırsa

$$P_X\left(\bigcup_{n=1}^{\infty} B_n\right) = P\left(\bigcup_{n=1}^{\infty} (X \in B_n)\right) = \sum_{n=1}^{\infty} P(X \in B_n) = \sum_{n=1}^{\infty} P_X(B_n)$$

elde edilir.

Üç aksiyom da sağlandığından $P_X, \mathcal{B}(\mathbb{R})$ üzerinde bir olasılık ölçüsüdür; $\mathcal{B}(\mathbb{R})$ de $\mathbb{R}$ üzerinde bir σ -cebir olduğundan $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$ bir olasılık uzayıdır (**Tanım 3.2**). ■

Teoremdaki ölçünün adı vardır ve bundan sonraki bölümlerin ana nesnesi odur.

Tanım 8.3 (Olasılık Dağılımı). X bir rastgele değişken olsun. Teorem 8.1’nde tanımlanan, $\mathcal{B}(\mathbb{R})$ üzerindeki

$$P_X(B) = P(X \in B), \quad B \in \mathcal{B}(\mathbb{R})$$

olasılık ölçüsüne X rastgele değişkeninin **olasılık dağılımı** (ya da X ’in olasılık ölçüsü) denir. $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$ uzayına da X ’in **ürettiği olasılık uzayı** denir.

Tanımla ilgili üç not:

- P_X bir olasılık ölçüsü olduğundan, olasılık ölçüleri için ispatlanmış her şey ona da uygulanır: $P_X(B^c) = 1 - P_X(B)$, $P_X(\emptyset) = 0$, $A \subseteq B \Rightarrow P_X(A) \leq P_X(B)$, toplama kuralı $P_X(A \cup B) = P_X(A) + P_X(B) - P_X(A \cap B)$ (Teorem 3.1, Teorem 3.2) ve artan ya da azalan Borel kümesi dizileri için süreklilik (Teorem 3.5).
- D_X bir Borel kümesiye, X değer kümesinin dışına düşemediğinden her $B \in \mathcal{B}(\mathbb{R})$ için $P_X(B) = P_X(B \cap D_X)$ ve özel olarak $P_X(D_X) = 1$ olur; yani P_X bütün kütesini D_X üzerinde taşır. Bu kitapta karşılaşılan bütün değer kümeleri (sonlu ya da sayılabilir kümeler ve aralıklar) Borel’dir.
- P_X , iki farklı örnek uzayda tanımlı rastgele değişkenleri karşılaştırmayı mümkün kılar: X ile Y farklı deneylere ait olsa bile ikisinin dağılımı da $\mathcal{B}(\mathbb{R})$ üzerinde birer ölçüdür ve doğrudan karşılaştırılabilir.

Dağılımı somut örneklerde hesaplayalım. Değer kümesi sonlu olduğunda her Borel kümesinin olasılığı, o kümeye düşen değerlerin olasılıkları toplanarak bulunur; değer kümesi bir aralık olduğunda ise uzunluk hesabına dönülür.

Örnek 8.5 (Olasılık Dağılımı Hesabı). a) Örnek 8.2’deki tura sayısı X için $P_X(\{1, 2\})$, $P_X((-\infty, 1])$, $P_X((1, 3))$, $P_X([\frac{1}{2}, \frac{5}{2}])$ ve $P_X(\mathbb{R} \setminus \{0, 1, 2, 3\})$ değerlerini hesaplayınız. Genel bir $B \in \mathcal{B}(\mathbb{R})$ için $P_X(B)$ ’yi veren formülü yazınız.

b) Örnek 8.3’deki $X(\omega) = \omega^2$ rastgele değişkeni için $P_X([0, \frac{1}{4}])$, $P_X((\frac{1}{4}, \frac{9}{16}])$ ve $P_X(\{\frac{1}{4}\})$ değerlerini hesaplayınız.

Çözüm.

a) Her durumda $P_X(B) = P(X \in B)$ tanımını açıp B ’ye düşen sonuçları sayacağız.

$$P_X(\{1, 2\}) = P(X \in \{1, 2\}) = P(\{YYT, YTY, TYY, YTT, TYT, TTY\}) = \frac{6}{8} = \frac{3}{4}.$$

$B = (-\infty, 1]$ için $X(\omega) \leq 1$ olan sonuçlar aranır:

$$P_X((-\infty, 1]) = P(X \leq 1) = P(\{YYY, YYT, YTY, TYY\}) = \frac{4}{8} = \frac{1}{2}.$$

$B = (1, 3)$ açık aralığına D_X ’ten yalnızca 2 düşer:

$$P_X((1, 3)) = P(1 < X < 3) = P(X = 2) = \frac{3}{8}.$$

$B = [\frac{1}{2}, \frac{5}{2}]$ aralığına D_X ’ten 1 ve 2 düşer; $B \cap D_X = \{1, 2\}$ olduğundan

$$P_X\left(\left[\frac{1}{2}, \frac{5}{2}\right]\right) = P_X(\{1, 2\}) = \frac{3}{4}.$$

Son olarak $B = \mathbb{R} \setminus \{0, 1, 2, 3\} = \mathbb{R} \setminus D_X$ için $B \cap D_X = \emptyset$ ’dir, dolayısıyla $(X \in B) = \emptyset$ ve $P_X(B) = 0$. Aynı şey $P_X(\mathbb{R}) = 1$ ile tümleyen kuralından da görülür: $P_X(D_X) = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = 1$. Genel formül: $B \in \mathcal{B}(\mathbb{R})$ için $(X \in B) = (X \in B \cap D_X) = \bigcup_{x \in B \cap D_X} (X = x)$ olur ve bu birleşim en çok dört ayrık olaydan oluşur. P ’nin toplamsallığıyla

$$P_X(B) = \sum_{x \in B \cap \{0, 1, 2, 3\}} P(X = x),$$

yani P_X tümüyle Örnek 8.2’deki tablo tarafından belirlenir. Bu, kesikli rastgele değişkenlerin dağılımının olasılık fonksiyonuyla belirlenmesinin ilk örneğidir.

b) $\Omega = [-1, 1]$ ve P uzunluğun yarısıdır. $B = [0, \frac{1}{4}]$ için

$$(X \in B) = \left\{ \omega \in [-1, 1] : 0 \leq \omega^2 \leq \frac{1}{4} \right\} = \left[-\frac{1}{2}, \frac{1}{2} \right], \quad P_X(B) = \frac{1}{2} \cdot 1 = \frac{1}{2}.$$

$B = (\frac{1}{4}, \frac{9}{16}]$ için $\frac{1}{4} < \omega^2 \leq \frac{9}{16} \Leftrightarrow \frac{1}{2} < |\omega| \leq \frac{3}{4}$ olduğundan

$$(X \in B) = \left[-\frac{3}{4}, -\frac{1}{2} \right) \cup \left(\frac{1}{2}, \frac{3}{4} \right], \quad P_X(B) = \frac{1}{2} \left(\frac{1}{4} + \frac{1}{4} \right) = \frac{1}{4}.$$

Aynı sonuç $P_X((\frac{1}{4}, \frac{9}{16}]) = P(X \leq \frac{9}{16}) - P(X \leq \frac{1}{4}) = \frac{3}{4} - \frac{1}{2} = \frac{1}{4}$ biçiminde, **Örnek 8.3'**de bulunan $P(X \leq a) = \sqrt{a}$ formülünden de elde edilir.

Tek nokta için $(X \in \{\frac{1}{4}\}) = (X = \frac{1}{4}) = \{-\frac{1}{2}, \frac{1}{2}\}$ ve iki noktalı kümenin uzunluğu sıfır olduğundan $P_X(\{\frac{1}{4}\}) = 0$. Burada dağılım tek noktalara değil, aralıklara kütle yükler; bu, sürekli rastgele değişkenlerin tipik davranışdır.

i Aynı dağılım, farklı rastgele değişkenler

P_X , rastgele değişkeni belirlemez. **Örnek 8.2'**ndeki uzayda $Y(\omega)$, ω sonucundaki **yazı** sayısı olsun; $Y = 3 - X$ 'tir ve Y de bir rastgele değişkendir. x yazı içeren sonuçların sayısı, x tura içeren sonuçların sayısına eşit olduğundan

$$P(Y = 0) = \frac{1}{8}, \quad P(Y = 1) = \frac{3}{8}, \quad P(Y = 2) = \frac{3}{8}, \quad P(Y = 3) = \frac{1}{8};$$

yani her Borel B için $P_Y(B) = P_X(B)$, kısaca $P_Y = P_X$. Oysa $X \neq Y$: örneğin $X(TTT) = 3$ iken $Y(TTT) = 0$ 'dır. Dağılım, rastgele değişkenin **hangi olasılıkla hangi değeri aldığını** söyler; **hangi sonuçta** aldığını söylemez. Olasılık teorisinin büyük bölümü için bu bilgi yeterlidir; işte bu yüzden Ω çoğu zaman arka planda kalır.

8.6 Alıştırmalar

Alıştırma 8.1 (Alıştırmalar). a) İki zar atılıyor: $\Omega = \{(i, j) : i, j \in \{1, 2, 3, 4, 5, 6\}\}$, $\mathcal{U} = \mathcal{P}(\Omega)$, $P(A) = \frac{n(A)}{36}$. $X(i, j) = i + j$ (üst yüzlerin toplamı) olsun. X 'in bir rastgele değişken olduğunu gösteriniz; D_X 'i bulunuz; $P_X(\{7\})$, $P_X([10, \infty))$ ve $P_X((-\infty, 3])$ değerlerini hesaplayınız.

b) $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ ve $\mathcal{U} = \{\emptyset, \{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}, \Omega\}$ olsun. Aşağıdaki fonksiyonlardan hangileri rastgele değişkendir?

$$X(\omega_1) = X(\omega_2) = -1, \quad X(\omega_3) = X(\omega_4) = 5;$$

$$Y(\omega_1) = 1, \quad Y(\omega_2) = 2, \quad Y(\omega_3) = Y(\omega_4) = 3;$$

$$Z(\omega_i) = (-1)^i.$$

c) $\Omega = [0, 1]$, $\mathcal{U} = \mathcal{B}(\Omega)$ ve $P(A) = \text{uzunluk}(A)$ olsun. $X(\omega) = 2\omega - 1$ için her $a \in \mathbb{R}$ 'de $(X \leq a)$ kümesini bulunuz, X 'in rastgele değişken olduğunu gösteriniz ve $P_X([0, \frac{1}{2}])$ 'yi hesaplayınız.

d) (c)'deki uzayda $X(\omega) = \min\{\omega, \frac{1}{2}\}$ olsun. X 'in rastgele değişken olduğunu gösteriniz; D_X 'i, $P_X(\{\frac{1}{2}\})$ ve $P_X([0, \frac{1}{4}])$ değerlerini bulunuz.

e) X bir rastgele değişken ve $c, d \in \mathbb{R}$ olsun. $Y = cX + d$ fonksiyonunun, yani $Y(\omega) = cX(\omega) + d$ ile tanımlanan $Y : \Omega \rightarrow \mathbb{R}$ 'nin bir rastgele değişken olduğunu gösteriniz.

f) X ve Y aynı olasılık uzayında tanımlı iki rastgele değişken olsun. $M(\omega) = \max\{X(\omega), Y(\omega)\}$ ve $m(\omega) = \min\{X(\omega), Y(\omega)\}$ fonksiyonlarının rastgele değişken olduğunu gösteriniz.

g) X bir rastgele değişken olsun. $|X|$ ve X^2 fonksiyonlarının rastgele değişken olduğunu gösteriniz.

Çözüm.

a) $\mathcal{U} = \mathcal{P}(\Omega)$ olduğundan Ω 'nın her alt kümesi, özel olarak her $(X \leq a)$ kümesi $\mathcal{U}$ 'dadır; X bir rastgele değişkendir. $i + j$ toplamı 2'den 12'ye kadar her tam sayı değerini alır: $D_X = \{2, 3, \dots, 12\}$. Toplamı 7 yapan çiftler $(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)$ olmak üzere altı tanedir:

$$P_X(\{7\}) = P(X = 7) = \frac{6}{36} = \frac{1}{6}.$$

$[10, \infty) \cap D_X = \{10, 11, 12\}$ 'dir; toplamı 10 yapan üç çift $((4, 6), (5, 5), (6, 4))$, 11 yapan iki çift $((5, 6), (6, 5))$, 12 yapan bir çift $((6, 6))$ vardır:

$$P_X([10, \infty)) = P(X \geq 10) = \frac{3 + 2 + 1}{36} = \frac{6}{36} = \frac{1}{6}.$$

$(-\infty, 3] \cap D_X = \{2, 3\}$ 'tür; toplamı 2 yapan bir çift $((1, 1))$, 3 yapan iki çift $((1, 2), (2, 1))$ vardır:

$$P_X((-\infty, 3]) = P(X \leq 3) = \frac{1 + 2}{36} = \frac{1}{12}.$$

b) $\mathcal{U}$ 'nun elemanı olmayan alt kümelerin her biri $\{\omega_1, \omega_2\}$ ya da $\{\omega_3, \omega_4\}$ kümelerinden en az birini böler. $D_X = \{-1, 5\}$ olmak üzere

$$(X \leq a) = \begin{cases} \emptyset, & a < -1 \\ \{\omega_1, \omega_2\}, & -1 \leq a < 5 \\ \Omega, & a \geq 5 \end{cases}$$

kümelerinin üçü de $\mathcal{U}$ 'dadır; X **rastgele değişkendir**. Y için $a = 1$ alınırsa $(Y \leq 1) = \{\omega_1\} \notin \mathcal{U}$; Y **rastgele değişken değildir**. $Z(\omega_1) = -1, Z(\omega_2) = 1, Z(\omega_3) = -1, Z(\omega_4) = 1$ olduğundan $(Z \leq -1) = \{\omega_1, \omega_3\} \notin \mathcal{U}$; Z de **rastgele değişken değildir**.

c) $X(\omega) \leq a \Leftrightarrow 2\omega - 1 \leq a \Leftrightarrow \omega \leq \frac{a+1}{2}$. $\omega \in [0, 1]$ olduğundan

$$(X \leq a) = \begin{cases} \emptyset, & a < -1 \\ [0, \frac{a+1}{2}], & -1 \leq a \leq 1 \\ \Omega, & a > 1. \end{cases}$$

Boş küme, kapalı aralık ve Ω Borel kümeleri olduğundan her a için $(X \leq a) \in \mathcal{U}$; X bir rastgele değişkendir ve $D_X = [-1, 1]$ 'dir. $0 \leq 2\omega - 1 \leq \frac{1}{2} \Leftrightarrow \frac{1}{2} \leq \omega \leq \frac{3}{4}$ olduğundan

$$P_X\left(\left[0, \frac{1}{2}\right]\right) = P\left(\left[\frac{1}{2}, \frac{3}{4}\right]\right) = \frac{3}{4} - \frac{1}{2} = \frac{1}{4}.$$

d) $\omega \leq \frac{1}{2}$ ise $X(\omega) = \omega$, $\omega > \frac{1}{2}$ ise $X(\omega) = \frac{1}{2}$ 'dir; dolayısıyla $D_X = [0, \frac{1}{2}]$. $a \in \mathbb{R}$ için

$$(X \leq a) = \begin{cases} \emptyset, & a < 0 \\ [0, a], & 0 \leq a < \frac{1}{2} \\ \Omega, & a \geq \frac{1}{2}, \end{cases}$$

çünkü $0 \leq a < \frac{1}{2}$ iken $\min\{\omega, \frac{1}{2}\} \leq a$ koşulu tam olarak $\omega \leq a$ demektir, $a \geq \frac{1}{2}$ iken ise her ω için sağlanır. Üç küme de Borel'dir; X bir rastgele değişkendir.

$$P_X\left(\left\{\frac{1}{2}\right\}\right) = P\left(X = \frac{1}{2}\right) = P\left(\left[\frac{1}{2}, 1\right]\right) = \frac{1}{2}, \quad P_X\left(\left[0, \frac{1}{4}\right]\right) = P\left(\left[0, \frac{1}{4}\right]\right) = \frac{1}{4}.$$

Örnek uzay bir aralık olduğu hâlde P_X tek bir noktaya $\frac{1}{2}$ kütle yüklüyor: $\frac{1}{2}$ 'den büyük her sonuç aynı değere sıkıştırılmıştır. Böyle bir rastgele değişken ne kesiklidir ne sürekli; değer kümesi bir aralık, ama dağılımın kütesinin yarısı tek bir noktada toplanmıştır (kalan yarısı $[0, \frac{1}{2}]$ üzerine yayılmıştır).

e) Üç durum vardır.

- $c > 0$: $cX(\omega) + d \leq a \Leftrightarrow X(\omega) \leq \frac{a-d}{c}$, dolayısıyla $(Y \leq a) = (X \leq \frac{a-d}{c}) \in \mathcal{U}$.
- $c < 0$: c ile bölünce eşitsizlik yön değiştirir, $cX(\omega) + d \leq a \Leftrightarrow X(\omega) \geq \frac{a-d}{c}$; dolayısıyla $(Y \leq a) = (X \geq \frac{a-d}{c})$ ve **Önerme 8.2'**e göre bu küme $\mathcal{U}$ 'dadır.
- $c = 0$: Y sabit d fonksiyonudur; $a < d$ için $(Y \leq a) = \emptyset$, $a \geq d$ için $(Y \leq a) = \Omega$ ve ikisi de $\mathcal{U}$ 'dadır.

Her durumda Y bir rastgele değişkendir.

f) $\max\{X(\omega), Y(\omega)\} \leq a$ olması, $X(\omega) \leq a$ ve $Y(\omega) \leq a$ olmasına denktir:

$$(M \leq a) = (X \leq a) \cap (Y \leq a) \in \mathcal{U},$$

ünkü $\mathcal{U}$ kesiřime kapalıdır. $\min\{X(\omega), Y(\omega)\} \leq a$ olması ise $X(\omega) \leq a$ **veya** $Y(\omega) \leq a$ olmasına denktir:

$$(m \leq a) = (X \leq a) \cup (Y \leq a) \in \mathcal{U}.$$

Demek ki M ve m rastgele deęiřkendir. Aynı akıl yürütme sonlu sayıda rastgele deęiřkenin en büyüęü ve en küçüęü için de geçerlidir.

g) $a < 0$ ise $|X(\omega)| \leq a$ hiçbir ω için saęlanmaz: $(|X| \leq a) = \emptyset \in \mathcal{U}$. $a \geq 0$ ise $|X(\omega)| \leq a \Leftrightarrow -a \leq X(\omega) \leq a$, dolayısıyla

$$(|X| \leq a) = (-a \leq X \leq a) = (X \leq a) \cap (X \geq -a) \in \mathcal{U}$$

(Önerme 8.2). Benzer biçimde $a < 0$ için $(X^2 \leq a) = \emptyset$; $a \geq 0$ için $X(\omega)^2 \leq a \Leftrightarrow |X(\omega)| \leq \sqrt{a}$, yani $(X^2 \leq a) = (|X| \leq \sqrt{a})$ ve bu küme az önce gösterildięi gibi $\mathcal{U}$ 'dadır. Öyleyse $|X|$ ve X^2 rastgele deęiřkendir. (Genel kural, bir rastgele deęiřkenin sürekli bir fonksiyonunun yine rastgele deęiřken olduęudur: [Teorem 12.1](#).)

■

Bir rastgele deęiřkenin daęılımı P_X , Borel kümelerinin tamamı üzerinde tanımlı bir ölçüdür; ama bu ölçüyü elde tutmak için her Borel kümesine ne verdięini bilmek gerekmez. Yalnızca $(-\infty, a]$ yarı doęrularına verdięi deęerler, yani $a \mapsto P(X \leq a)$ fonksiyonu, P_X 'i bütünüyle belirler. Bu tek deęiřkenli fonksiyon, kitabın bundan sonraki ana aracıdır: [Daęılım Fonksiyonu](#).

9. Dağılım Fonksiyonu

Rastgele Değişken bölümünde bir olasılık uzayı üzerinde tanımlı ve her $(X \leq a)$ olayını $\mathcal{U}$ içinde tutan fonksiyonları rastgele değişken diye adlandırdık (Tanım 8.1). Ardından her rastgele değişkenin, $P_X(B) = P(X \in B)$ kuralıyla, reel sayıların Borel cebri üzerinde bir olasılık ölçüsü doğurduğunu gördük (Teorem 8.1); bu ölçüye X 'in olasılık dağılımı dedik (Tanım 8.3). P_X , X ile ilgili bütün olasılık bilgisini taşır: “ X hangi Borel kümesine hangi olasılıkla düşer?” sorusunun her yanıtı P_X 'in içindedir.

Ne var ki P_X bir ölçüdür; onu bilmek demek, sayılamayan çoklukta Borel kümesinin her birinde değeri bilmek demektir. Bununla çalışmak zordur. Bu bölümde aynı bilgiyi tek bir reel fonksiyona sıkıştırıyoruz: her $x \in \mathbb{R}$ için $F_X(x) = P(X \leq x)$ sayısını veren **dağılım fonksiyonu**. Bu fonksiyon yalnızca $(-\infty, x]$ biçimindeki kümelerin olasılığını kaydeder; ama bu kadarının, Borel cebri üzerindeki bütün olasılıkları geri kazanmaya yettiğini göreceğiz.

Bölümün akışı şöyledir. Önce tanımı veriyor ve dört somut örnekte, sonlu bir örnek uzaydan başlayıp bir aralık üzerindeki uzunluk ölçüsüne kadar, F_X 'i parça parça hesaplıyoruz. Sonra her dağılım fonksiyonunun taşıdığı üç yapısal özelliği (azalmayanlık, sağdan süreklilik, uç limitler) rastgele değişken diliyle ifade edip ispatlıyoruz. Bunun ardından asıl kullanışlı sonuca geçiyoruz: $P(a < X \leq b)$, $P(X = a)$, $P(X < a)$ gibi bütün aralık ve nokta olasılıklarının yalnızca F_X 'in değerleri ve sol limitleriyle hesaplanması. Bölümü, dağılım fonksiyonunun olasılık dağılımını tek biçimde belirlediğini gösteren önermeyle ve alıştırmalarla kapatıyoruz.

9.1 Dağılım Fonksiyonunun Tanımı

X bir rastgele değişken olsun. Rastgele değişken tanımı gereği her $x \in \mathbb{R}$ için $(X \leq x) = \{\omega \in \Omega : X(\omega) \leq x\}$ kümesi $\mathcal{U}$ 'nın bir elemanıdır; dolayısıyla $P(X \leq x)$ sayısı anlamlıdır. Bu sayıyı x 'in fonksiyonu olarak görmek, bölümün bütün fikridir.

Tanım 9.1 (Rastgele Değişkenin Dağılım Fonksiyonu). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı ve $X : \Omega \rightarrow \mathbb{R}$ bir rastgele değişken olsun.

$$F_X : \mathbb{R} \rightarrow [0, 1], \quad F_X(x) = P(X \leq x) = P(\{\omega \in \Omega : X(\omega) \leq x\}) = P_X((-\infty, x])$$

fonksiyonuna X rastgele değişkeninin **dağılım fonksiyonu** (birikimli dağılım fonksiyonu) denir.

Tanımdaki üç eşitlik aynı sayının üç yazımıdır. İlki olay kısaltmasıdır; ikincisi bu olayı açık küme olarak yazar; üçüncüsü ise P_X 'in tanımını kullanır: $(X \leq x) = (X \in (-\infty, x])$ olduğundan $P(X \leq x) = P_X((-\infty, x])$ 'tir. Bu son yazım önemli bir gözlemi hemen verir: $P_X, \mathcal{B}(\mathbb{R})$ üzerinde bir Borel olasılık ölçüsüdür (Tanım 4.2, Teorem 8.1) ve F_X , tam olarak bu ölçüye karşılık gelen, Tanım 4.3 anlamındaki dağılım fonksiyonudur: $x \mapsto P_X((-\infty, x])$. Yani rastgele değişkenin dağılım fonksiyonu yeni bir kavram değildir; **Olasılık Uzayı Örnekleri ve Geometrik Olasılık** bölümünde Borel olasılık ölçüleri için tanımladığımız kavramın P_X ölçüsüne uygulanmış hâlidir. Ölçüler için ispatladığımız her şey F_X için de geçerlidir ve bu bölümde bunları rastgele değişken diline çevireceğiz.

Bağlamdan hangi rastgele değişkenin söz konusu olduğu belliyse F_X yerine kısaca F yazılır. F_X 'in tanım kümesi her zaman bütün $\mathbb{R}$ 'dir; X 'in değer kümesi D_X ne olursa olsun, $F_X(x)$ her reel x için hesaplanır. x , X 'in alabileceği bütün değerlerin altındaysa $(X \leq x)$ boş olaydır ve $F_X(x) = 0$ olur; x bütün değerlerin üstündeyse $(X \leq x) = \Omega$ ve $F_X(x) = 1$ olur. İlginç olan, aradaki bölgedir.

9.2 Dağılım Fonksiyonu Örnekleri

Dağılım fonksiyonunu hesaplamamanın tek yolu vardır: her x için $(X \leq x)$ olayını Ω 'nın bir alt kümesi olarak açıkça belirlemek ve P ile ölçmek. Olayın biçimi x 'in bulunduğu bölgeye göre değiştiğinden, F_X hemen her zaman parçalı tanımlı çıkar. Aşağıdaki dört örnek bu hesabın tipik durumlarını gösterir.

Örnek 9.1 (Üç Para Atışında Tura Sayısının Dağılım Fonksiyonu). Bir para üç kez atılıyor. Örnek uzay

$$\Omega = \{YYY, YYT, YTY, YTT, TYY, TYT, TTY, TTT\},$$

$\mathcal{U} = \mathcal{P}(\Omega)$ ve her $A \in \mathcal{U}$ için $P(A) = \frac{n(A)}{8}$ olsun (Tanım 4.1). Gelen tura sayısını sayan

$$X(\omega) = \begin{cases} 0, \omega = YYY \\ 1, \omega \in \{YYT, YTY, TYY\} \\ 2, \omega \in \{YTT, TYT, TTY\} \\ 3, \omega = TTT \end{cases}$$

rastgele değişkeninin dağılım fonksiyonunu bulunuz.

Çözüm.

X 'in değer kümesi $D_X = \{0, 1, 2, 3\}$ 'tür. Verilen bir x için $(X \leq x)$ olayı, X 'in x 'i aşmayan değerlerini aldığı sonuçlardan oluşur; bu yüzden x 'in 0, 1, 2, 3 sayılarına göre nerede durduğuna bakmak yeter.

$x < 0$ ise X 'in hiçbir değeri x 'i aşmaz; $(X \leq x) = \emptyset$ ve $F_X(x) = P(\emptyset) = 0$.

$0 \leq x < 1$ ise yalnızca 0 değeri x 'i aşmaz:

$$(X \leq x) = (X = 0) = \{YYY\}, \quad F_X(x) = \frac{1}{8}.$$

$1 \leq x < 2$ ise 0 ve 1 değerleri x 'i aşmaz:

$$(X \leq x) = (X = 0) \cup (X = 1) = \{YYY, YYT, YTY, TYY\}, \quad F_X(x) = \frac{4}{8} = \frac{1}{2}.$$

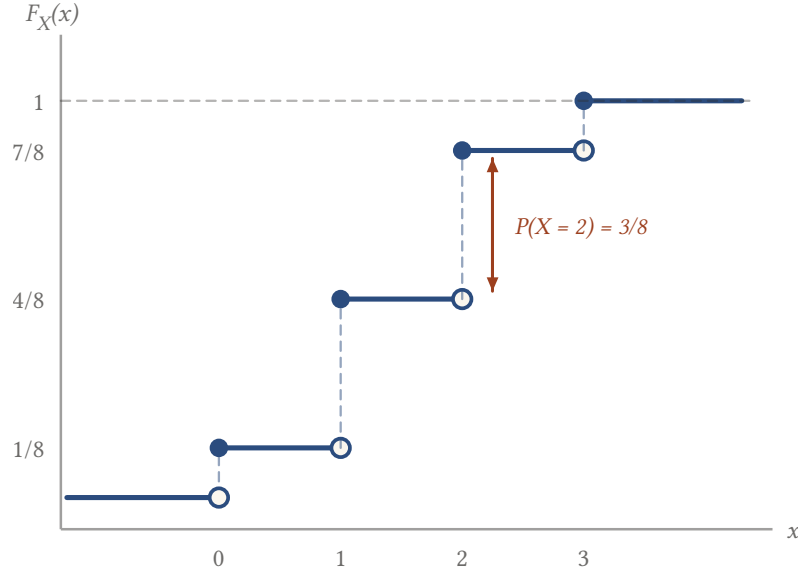
$2 \leq x < 3$ ise 3 dışındaki bütün değerler x 'i aşmaz:

$$(X \leq x) = \Omega \setminus \{TTT\}, \quad F_X(x) = \frac{7}{8}.$$

$x \geq 3$ ise her sonuç olaya girer: $(X \leq x) = \Omega$ ve $F_X(x) = 1$.

Parçaları bir araya getirirsek

$$F_X(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{8}, & 0 \leq x < 1 \\ \frac{1}{2}, & 1 \leq x < 2 \\ \frac{7}{8}, & 2 \leq x < 3 \\ 1, & x \geq 3. \end{cases}$$



Şekil 9.1: Düzgün bir paranın üç kez atılmasında tura sayısı X 'in dağılım fonksiyonu. Basamaklar 0, 1, 2, 3 noktalarındadır; her basamağın yüksekliği o noktanın olasılığıdır ($P(X=2) = 3/8$). Fonksiyon sağdan süreklidir: her basamakta sağ uç dolu, sol uç boş.

F_X bir **basamak fonksiyonudur**: X 'in almadığı değerler boyunca sabit kalır, X 'in aldığı her değerde yukarı sıçrar. Sıçramaların boyları $\frac{1}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{8}$ 'dir ve bunlar tam olarak $P(X=0)$, $P(X=1)$, $P(X=2)$, $P(X=3)$ olasılıklarıdır; bu gözlemi ileride genel bir teoreme dönüştüreceğiz. Her sıçrama noktasında F_X 'in değeri **üst** basamağa aittir: örneğin $F_X(1) = \frac{1}{2}$ 'dir, $\frac{1}{8}$ değil. Bunun nedeni tanımdaki eşitsizliğin $\leq$ olmasıdır; ($X \leq 1$) olayı $X=1$ sonuçlarını içerir.

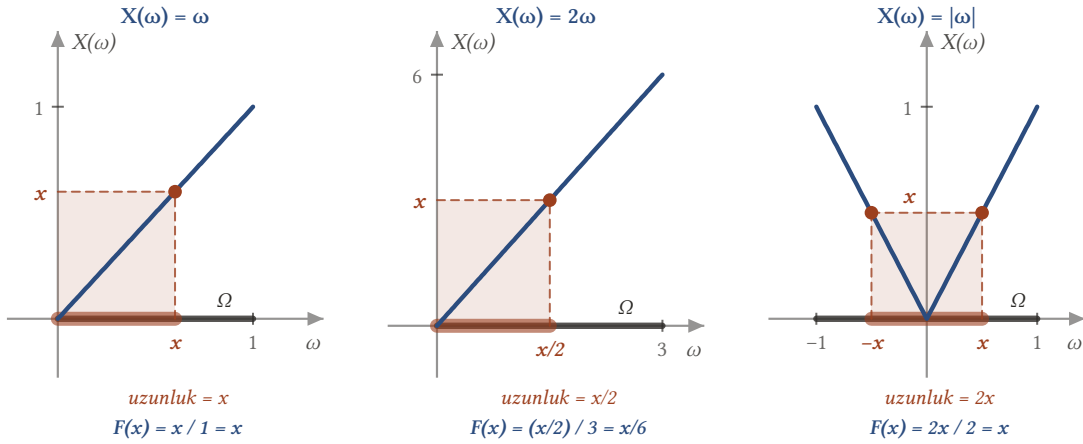
Şimdi sonsuz bir örnek uzaya geçelim. Bir aralık üzerinde uzunluk ölçüsüyle kurulan olasılık uzayında (Tanım 4.5) ($X \leq x$) olayı bir aralık (ya da birkaç aralığın birleşimi) olacak ve olasılığı uzunluğundan okunacaktır. Bu **uzunluk yöntemi** üç adımdan oluşur:

1. x verilmişken $\{\omega \in \Omega : X(\omega) \leq x\}$ kümesini, $X(\omega) \leq x$ eşitsizliğini ω için çözerek bul.
2. Bu kümenin uzunluğunu hesapla.
3. Uzunluğu Ω 'nın uzunluğuna böl; $F_X(x)$ budur.

Geometrik olarak: $X(\omega)$ 'nın grafiği çizilir, x düzeyinde yatay bir doğru çekilir, grafiğin bu doğrunun **altında** kaldığı ω 'lar ω ekseninde işaretlenir; işaretli parçanın uzunluğu, Ω 'nın uzunluğuna bölününce $F_X(x)$ 'i verir. Aşağıdaki üç kısa soru yöntemi en yalın hâlden başlayarak alıştıırır; asıl örnek (Örnek 9.4) bunların üstüne kurulur.

Örnek 9.2 (Uzunluk Yöntemine Alışma: Üç Isınma Sorusu). Aşağıdaki her şıkta $\mathcal{U} = \mathcal{B}(\Omega)$ ve $A \in \mathcal{U}$ için $P(A) = \frac{A \text{ 'nin uzunluğu}}{\Omega \text{ 'nin uzunluğu}}$ alınıyor.

- a) $\Omega = [0, 1]$ ve $X(\omega) = \omega$ olsun. F_X 'i bulunuz.
- b) $\Omega = [0, 3]$ ve $X(\omega) = 2\omega$ olsun. F_X 'i bulunuz.
- c) $\Omega = [-1, 1]$ ve $X(\omega) = |\omega|$ olsun. F_X 'i bulunuz.



Şekil 9.2: Uzunluk yöntemi: bir aralık üzerinde uzunluk ölçüsüyle kurulan olasılık uzayında $F(x)$, $(X \leq x)$ olayının uzunluğunun Ω 'nın uzunluğuna oranıdır. Her panelde x düzeyinin altında kalan ω 'lar ω ekseninde taranmıştır; bu aralığın uzunluğu x 'e bağlı olarak okunur ve Ω 'nın uzunluğuna (1, 3, 2) bölünür. Aynı x için taralı küme fonksiyona göre değişir: doğru için tek parça, mutlak değer için 0'a göre simetrik bir aralık.

Çözüm.

a) $\omega \in [0, 1]$ iken $X(\omega) = \omega \in [0, 1]$; $D_X = [0, 1]$.

Adım 1. $X(\omega) \leq x$ eşitsizliği burada doğrudan $\omega \leq x$ demektir. $x < 0$ için hiçbir ω bunu sağlamaz: $(X \leq x) = \emptyset$. $0 \leq x < 1$ için $(X \leq x) = \{\omega \in [0, 1] : \omega \leq x\} = [0, x]$. $x \geq 1$ için her ω sağlar: $(X \leq x) = \Omega$.

Adım 2-3. $[0, x]$ aralığının uzunluğu x , Ω 'nın uzunluğu 1'dir:

$$F_X(x) = \begin{cases} 0, & x < 0 \\ x, & 0 \leq x < 1 \\ 1, & x \geq 1. \end{cases}$$

Grafik, $[0, 1]$ üzerinde 45°'lik bir doğru parçasıdır: sonucun kendisi ω olan bir deneyde " $X \leq x$ olma olasılığı" tam olarak x 'tir. Bu, Örnek 4.4'ndeki düzgün dağılımdır.

b) $\omega \in [0, 3]$ iken $2\omega \in [0, 6]$; $D_X = [0, 6]$.

Adım 1. $2\omega \leq x \Leftrightarrow \omega \leq \frac{x}{2}$. Öyleyse $0 \leq x < 6$ için $(X \leq x) = [0, \frac{x}{2}]$; çünkü $\frac{x}{2} < 3$ olduğundan aralık Ω 'nın içinde kalır.

Adım 2-3. Uzunluk $\frac{x}{2}$, Ω 'nın uzunluğu 3:

$$F_X(x) = \begin{cases} 0, & x < 0 \\ \frac{x}{6}, & 0 \leq x < 6 \\ 1, & x \geq 6. \end{cases}$$

Denetim: $x = 6$ 'da $\frac{6}{6} = 1$, sürekli bir geçiş. Ölçeklemenin etkisi görülüyor: X 'in değerleri ω 'nın iki katı olduğundan $(X \leq x)$ olayı ω ekseninde yarı hızla büyür ve F_X , 6 birimlik bir aralığa yayılarak yükselir.

c) $\omega \in [-1, 1]$ iken $|\omega| \in [0, 1]$; $D_X = [0, 1]$.

Adım 1. Buradaki yenilik eşitsizliğin **iki yana** çözülmesidir: $|\omega| \leq x \Leftrightarrow -x \leq \omega \leq x$. $0 \leq x < 1$ için $[-x, x] \subseteq [-1, 1]$ olduğundan $(X \leq x) = [-x, x]$.

Adım 2-3. Uzunluk $x - (-x) = 2x$, Ω 'nın uzunluğu 2:

$$F_X(x) = \begin{cases} 0, & x < 0 \\ \frac{2x}{2} = x, & 0 \leq x < 1 \\ 1, & x \geq 1. \end{cases}$$

Grafikte $X(\omega) = |\omega|$ bir V harfidir; x düzeyinin altında kalan parça 0'a göre simetrik $[-x, x]$ aralıktır. Sonuç (a) ile aynı çıktı: ω ile $|\omega|$ birbirinden farklı fonksiyonlardır, ama aynı dağılıma sahiptirler. Dağılım fonksiyonu değişkenin kendisini değil, olasılığı nasıl dağıttığını betimler. Üç soruda da $(X \leq x)$ tek bir aralıktı ve Ω 'nın içinde kaldı. Asıl örnekte aralığın bir ucu Ω 'nın sınırına takılacak; formül bu yüzden parçalara ayrılacak.

■

Önce simetrik olmayan aralığa geçmeden, ω^2 'yi Ω 'nın tek yanlı olduğu bir aralıkta ele alalım.

Örnek 9.3 (Aralık Üzerinde Karenin Dağılım Fonksiyonu). $\Omega = [0, 2]$, $\mathcal{U} = \mathcal{B}(\Omega)$ ve $A \in \mathcal{U}$ için

$$P(A) = \frac{A \text{ 'nın uzunluğu}}{2}$$

olmak üzere $(\Omega, \mathcal{U}, P)$ olasılık uzayı alınsın. $X(\omega) = \omega^2$ ile tanımlanan $X : \Omega \rightarrow \mathbb{R}$ fonksiyonunun bir rastgele değişken olduğunu gösteriniz ve dağılım fonksiyonunu bulunuz.

Çözüm.

X bir rastgele değişkendir. $\omega \in [0, 2]$ için $\omega^2 \in [0, 4]$ olduğundan $D_X = [0, 4]$ 'tür. $a \in \mathbb{R}$ olsun.

$a < 0$ ise $(X \leq a) = \emptyset \in \mathcal{U}$.

$0 \leq a < 4$ ise, $\omega \geq 0$ olduğundan $\omega^2 \leq a$ eşitsizliği $\omega \leq \sqrt{a}$ ile aynıdır ve $\sqrt{a} < 2$ olduğundan

$$(X \leq a) = \{\omega \in [0, 2] : \omega^2 \leq a\} = [0, \sqrt{a}] \in \mathcal{U}.$$

$a \geq 4$ ise $(X \leq a) = \Omega \in \mathcal{U}$.

Her durumda $(X \leq a)$ bir Borel kümesidir; [Tanım 8.1](#) gereği X bir rastgele değişkendir.

Dağılım fonksiyonu. Yukarıdaki olayları P ile ölçmek yeter.

$x < 0$ için $F_X(x) = P(\emptyset) = 0$.

$0 \leq x < 4$ için

$$F_X(x) = P([0, \sqrt{x}]) = \frac{\sqrt{x} - 0}{2} = \frac{\sqrt{x}}{2}.$$

$x \geq 4$ için $F_X(x) = P(\Omega) = 1$.

Böylece

$$F_X(x) = \begin{cases} 0, & x < 0 \\ \frac{\sqrt{x}}{2}, & 0 \leq x < 4 \\ 1, & x \geq 4. \end{cases}$$

Parçaların birleşim noktalarına bakalım: $x = 0$ 'da $\frac{\sqrt{0}}{2} = 0$ ve $x = 4$ 'te $\frac{\sqrt{4}}{2} = 1$ 'dir. Yani bu kez F_X hiç sıçramaz; $\mathbb{R}$ üzerinde süreklidir ve $[0, 4]$ boyunca 0'dan 1'e kesintisiz yükselir. Bir önceki örnekle aradaki fark çarpıcıdır: orada olasılık dört noktaya yığılmıştı, burada $[0, 4]$ aralığına yayılmıştır.

■

Aynı fonksiyonu simetrik olmayan bir aralıkta tanımlarsak $(X \leq x)$ olayı Ω 'nın sınırına takılır ve F_X bir parça daha kazanır.

Örnek 9.4 (Simetrik Olmayan Aralıkta Karenin Dağılım Fonksiyonu). $\Omega = [-1, 2]$, $\mathcal{U} = \mathcal{B}(\Omega)$ ve $A \in \mathcal{U}$ için

$$P(A) = \frac{A \text{ 'nın uzunluğu}}{3}$$

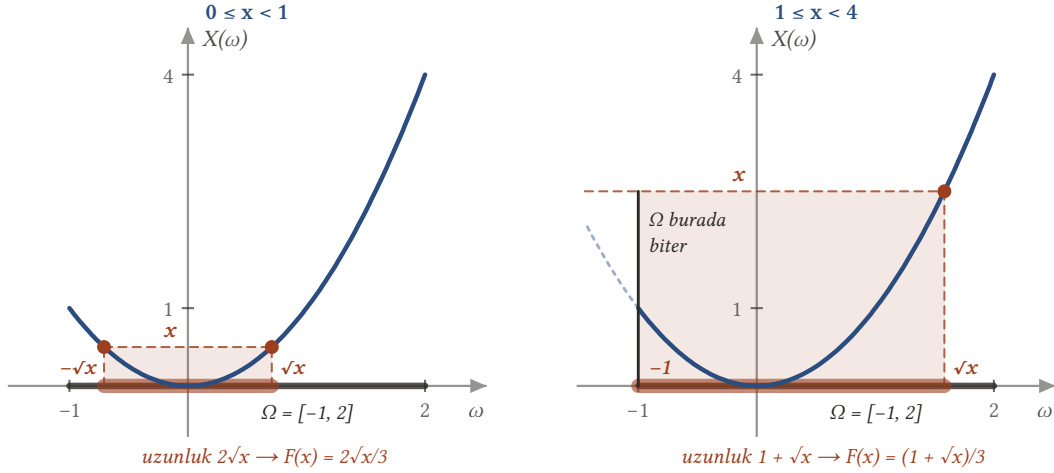
olmak üzere $(\Omega, \mathcal{U}, P)$ olasılık uzayı alınsın. $X(\omega) = \omega^2$ rastgele değişkeninin dağılım fonksiyonunu bulunuz.

Çözüm.

$\omega \in [-1, 2]$ iken ω^2 en küçük değerini $\omega = 0$ 'da (0), en büyük değerini $\omega = 2$ 'de (4) alır; $D_X = [0, 4]$ 'tür. $x \geq 0$ için

$$\omega^2 \leq x \Leftrightarrow -\sqrt{x} \leq \omega \leq \sqrt{x}$$

olduğundan $(X \leq x)$ olayı $[-\sqrt{x}, \sqrt{x}]$ aralığının $\Omega = [-1, 2]$ ile kesişimidir. Bu kesişimin biçimi, $-\sqrt{x}$ 'in -1 'in altına inip inmediğine bağlıdır; $\sqrt{x}$ ise $x < 4$ olduğu sürece 2'yi aşmaz. Uzunluk yönteminin geometrisi bunu doğrudan gösterir: parabolün x düzeyinin altında kalan parçası, x küçükken Ω 'nın içinde iki yana uzanan bir aralıktır; x büyüyünce bu aralığın sol ucu Ω 'nın sınırı olan -1 'de takılır.



Şekil 9.3: $\Omega = [-1, 2]$ üzerinde $X(\omega) = \omega^2$ için $(X \leq x)$ olayı, $[-\sqrt{x}, \sqrt{x}]$ aralığının Ω ile kesişimidir. $x < 1$ iken aralık Ω 'nın içine sığar ve olayın uzunluğu $2\sqrt{x}$ olur; $x \geq 1$ olunca sol uç Ω 'nın sınırı -1 'e takılır, olay yalnızca sağ uçtan büyür ve uzunluk $1 + \sqrt{x}$ 'e düşer. Dağılım fonksiyonunun $x = 1$ 'de formül değiştirmesi bu takılmanın sonucudur.

$x < 0$ için $(X \leq x) = \emptyset$ ve $F_X(x) = 0$.

$0 \leq x < 1$ için $\sqrt{x} < 1$, dolayısıyla $-1 < -\sqrt{x}$; aralığın tamamı Ω 'nın içindedir:

$$F_X(x) = P([- \sqrt{x}, \sqrt{x}]) = \frac{\sqrt{x} - (-\sqrt{x})}{3} = \frac{2\sqrt{x}}{3}.$$

$1 \leq x < 4$ için $-\sqrt{x} \leq -1$; aralığın sol ucu Ω 'nın dışına taşar ve kesişim $[-1, \sqrt{x}]$ olur:

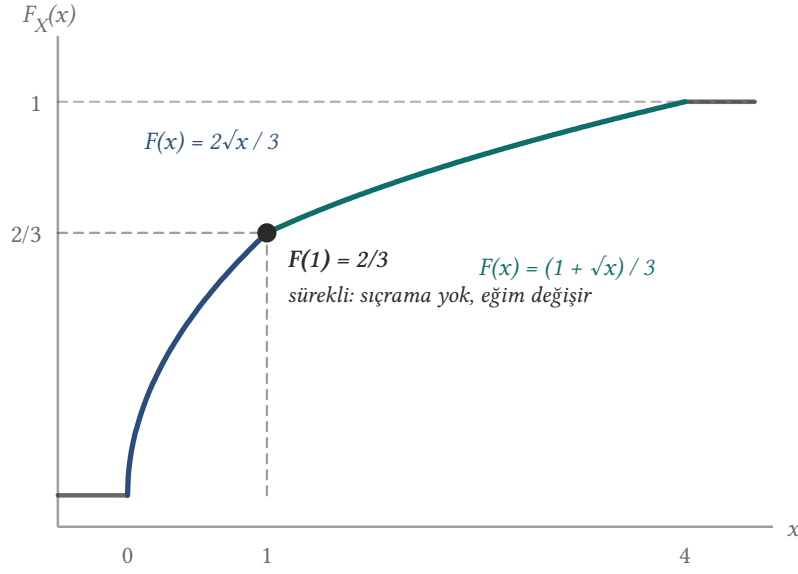
$$F_X(x) = P([-1, \sqrt{x}]) = \frac{\sqrt{x} + 1}{3}.$$

$x \geq 4$ için $(X \leq x) = \Omega$ ve $F_X(x) = 1$.

Sonuç:

$$F_X(x) = \begin{cases} 0, & x < 0 \\ \frac{2\sqrt{x}}{3}, & 0 \leq x < 1 \\ \frac{1+\sqrt{x}}{3}, & 1 \leq x < 4 \\ 1, & x \geq 4. \end{cases}$$

Birleşim noktalarını denetleyelim. $x = 1$ 'de birinci formül $\frac{2}{3}$, ikincisi $\frac{1+1}{3} = \frac{2}{3}$ verir; $x = 4$ 'te ikinci formül $\frac{1+2}{3} = 1$ verir. F_X yine her yerde süreklidir. $x = 1$ noktasında formülün değişmesi bir sıçrama değil, yalnızca eğimin değişmesidir: o ana kadar $(X \leq x)$ olayı iki uçtan birden büyürken, $x = 1$ 'den sonra sol uç -1 'e takılır ve olay yalnızca sağ uçtan büyür; bu yüzden F_X daha yavaş artar.



Şekil 9.4: $\Omega = [-1, 2]$ üzerinde $X(\omega) = \omega^2$ 'nin dağılım fonksiyonu. $0 \leq x < 1$ iken $(X \leq x)$ olayı iki uçtan birden büyür ve F hızla yükselir; $x \geq 1$ 'den sonra sol uç -1 'e takılır, olay yalnızca sağ uçtan büyür ve F yavaşlar. İki parça $x = 1$ 'de aynı değeri ($2/3$) verir: fonksiyon süreklidir, değişen yalnızca eğimdir.

■

Son örnekte X bir parabol olsun; $(X \leq x)$ olayını belirlemek için önce parabolü tam kareye tamamlamak gerekir.

Örnek 9.5 (Parabolün Dağılım Fonksiyonu). $\Omega = [-1, 2]$, $\mathcal{U} = \mathcal{B}(\Omega)$ ve $P(A) = \frac{A \text{ 'nin uzunluğu}}{3}$ olmak üzere $(\Omega, \mathcal{U}, P)$ olasılık uzayı alınsın. $X(\omega) = \omega^2 - 2\omega$ rastgele değişkeninin dağılım fonksiyonunu bulunuz.

Çözüm.

Değer kümesi. Tam kareye tamamlayalım: $X(\omega) = (\omega - 1)^2 - 1$. $\omega \in [-1, 2]$ iken $\omega - 1 \in [-2, 1]$ ve $(\omega - 1)^2 \in [0, 4]$ 'tür; en küçük değer $\omega = 1$ 'de ($X(1) = -1$), en büyük değer $\omega = -1$ 'de ($X(-1) = 3$) alınır. Öyleyse $D_X = [-1, 3]$.

$(X \leq x)$ olayı. $x \geq -1$ olsun ve $r = \sqrt{x+1} \geq 0$ yazalım. O zaman

$$X(\omega) \leq x \Leftrightarrow (\omega - 1)^2 \leq x + 1 \Leftrightarrow 1 - r \leq \omega \leq 1 + r$$

olur; yani $(X \leq x)$, $[1 - r, 1 + r]$ aralığının $\Omega = [-1, 2]$ ile kesişimidir. Bu aralığın merkezi 1'dir; sağ ucu $r \geq 1$ için Ω 'nın sağ sınırı olan 2'ye ulaşır ya da onu aşar, sol ucu ise ancak $r > 2$ için -1 'in altına iner. Bu iki eşik x cinsinden $x = 0$ ve $x = 3$ 'tür.

$x < -1$ için $(X \leq x) = \emptyset$ ve $F_X(x) = 0$.

$-1 \leq x < 0$ için $0 \leq r < 1$; $1 - r > 0$ ve $1 + r < 2$ olduğundan aralık Ω 'nın içindedir:

$$F_X(x) = P([1 - r, 1 + r]) = \frac{2r}{3} = \frac{2\sqrt{x+1}}{3}.$$

$0 \leq x < 3$ için $1 \leq r < 2$; sol uç $1 - r \in (-1, 0]$ hâlâ Ω 'nın içindedir, ama sağ uç $1 + r \geq 2$ olduğundan kesişim $[1 - r, 2]$ olur:

$$F_X(x) = P([1 - r, 2]) = \frac{2 - (1 - r)}{3} = \frac{1 + \sqrt{x+1}}{3}.$$

$x \geq 3$ için $r \geq 2$; $[1 - r, 1 + r] \supseteq [-1, 2]$ olduğundan $(X \leq x) = \Omega$ ve $F_X(x) = 1$.

Sonuç:

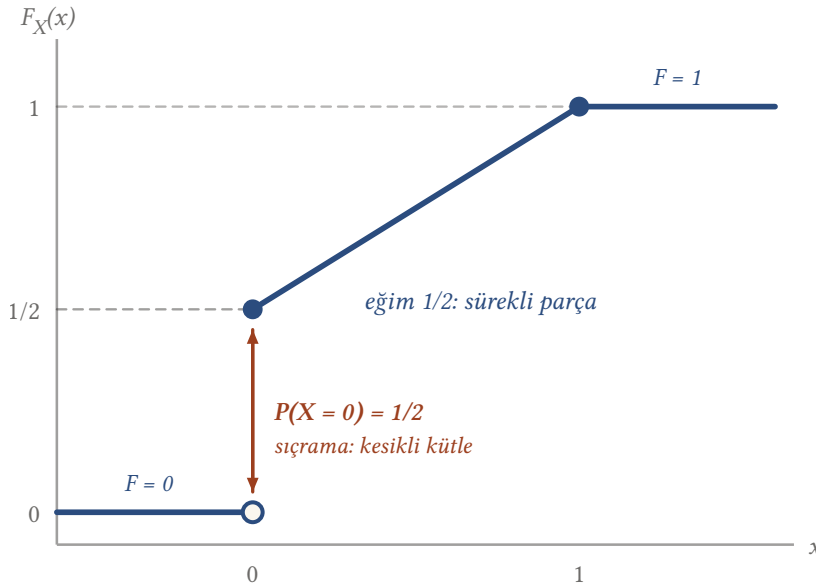
$$F_X(x) = \begin{cases} 0, & x < -1 \\ \frac{2\sqrt{x+1}}{3}, & -1 \leq x < 0 \\ \frac{1+\sqrt{x+1}}{3}, & 0 \leq x < 3 \\ 1, & x \geq 3. \end{cases}$$

Birleşim noktalarında $x = 0$ için $\frac{2\sqrt{1}}{3} = \frac{2}{3} = \frac{1+\sqrt{1}}{3}$ ve $x = 3$ için $\frac{1+\sqrt{4}}{3} = 1$ olduğundan F_X süreklidir. Sonucun [Örnek 9.4](#) ile benzerliği rastlantı değildir. $X = Y - 1$ yazılırsa, burada $Y(\omega) = (\omega - 1)^2$ 'dir ve $\omega \mapsto \omega - 1$ dönüşümü $[-1, 2]$ 'yi $[-2, 1]$ 'e taşır; $[-2, 1]$ üzerindeki kare fonksiyonu ise, $\omega \mapsto -\omega$ yansımasıyla, $[-1, 2]$ üzerindeki kare fonksiyonuyla aynı dağılıma sahiptir. Dolayısıyla F_Y , [Örnek 9.4](#)'teki fonksiyondur ve $F_X(x) = P(Y - 1 \leq x) = F_Y(x + 1)$ elde edilir; bu da F_Y 'nin formüllerinde x yerine $x + 1$ konmuş hâlidir.

■

Buraya kadar iki uç gördük: olasılığın birkaç noktaya yığıldığı basamaklı F_X ve bir aralığa yayıldığı sürekli F_X . İkisi aynı değişkende bir arada da bulunabilir.

Örnek 9.6 (Para ve Düzgün Sayı: Karma Bir Rastgele Değişken). Düzgün bir para atılıyor. Yazı gelirse $X = 0$ alınıyor; tura gelirse X , $(0, 1)$ aralığından uzunluk ölçüsüyle (düzgün olasılıkla) seçilen bir sayı oluyor. X 'in dağılım fonksiyonunu bulunuz; $P(X = 0)$, $(0 < X \leq \frac{1}{2})$ ve $(X = \frac{1}{2})$ olasılıklarını hesaplayınız.



Şekil 9.5: Karma bir rastgele değişkenin dağılım fonksiyonu: para yazı gelirse $X = 0$, tura gelirse X $(0, 1)$ üzerinde düzgün seçilir. $x = 0$ 'daki sıçrama $P(X = 0) = 1/2$ kesikli kütle; $(0, 1)$ üzerindeki eğimi $1/2$ olan rampa ise olasılığın geri kalan yarısının aralığa yayıldığını gösterir. Fonksiyon ne basamaklı ne de düzdür: sıçrama ile rampa bir aradadır.

Çözüm.

Model. Y (yazı) ve T (tura) olayları ayrık, $P(Y) = P(T) = \frac{1}{2}$. Tura geldiğinde X 'in $(0, 1)$ üzerinde düzgün seçilmesi, her $0 \leq x < 1$ için $P(X \leq x | T) = x$ demektir ([Örnek 9.2 \(a\)](#)); yazı geldiğinde ise $X = 0$ kesin olduğundan $x \geq 0$ için $P(X \leq x | Y) = 1$, $x < 0$ için 0'dır.

Dağılım fonksiyonu. $(X \leq x)$ olayını Y ve T 'ye göre parçalayalım ([Koşullu Olasılık](#) bölümündeki toplam olasılık kuralı):

$$F_X(x) = P(X \leq x) = P(Y) P(X \leq x | Y) + P(T) P(X \leq x | T).$$

$x < 0$ için iki koşullu olasılık da 0'dır: $F_X(x) = 0$.

$0 \leq x < 1$ için birincisi 1, ikincisi x 'tir:

$$F_X(x) = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot x = \frac{1}{2} + \frac{x}{2}.$$

$x \geq 1$ için ikisi de 1'dir: $F_X(x) = 1$. Sonuç:

$$F_X(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{2} + \frac{x}{2}, & 0 \leq x < 1 \\ 1, & x \geq 1. \end{cases}$$

Olasılıklar. $P(X = 0)$ için para yazı gelmeli ya da tura gelip düzgün sayı tam 0 çıkmalıdır; ikincisinin olasılığı tek bir noktanın uzunluğu, yani 0'dır:

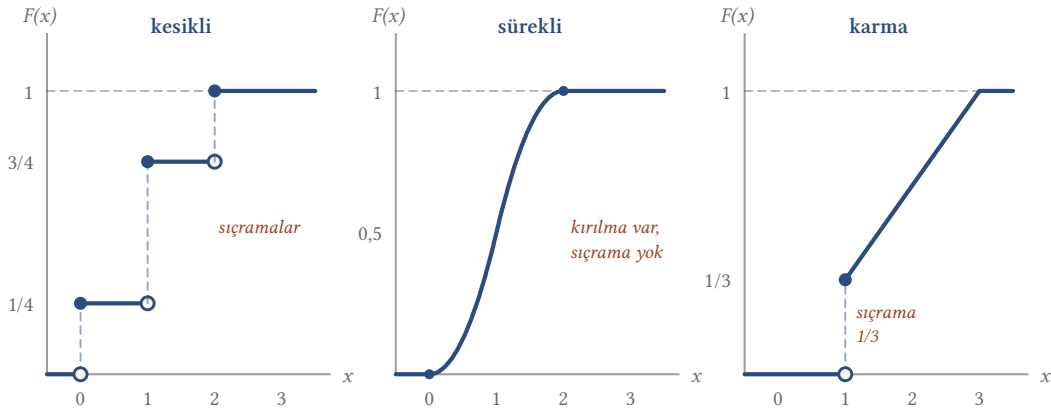
$$P(X = 0) = P(Y) \cdot 1 + P(T) \cdot 0 = \frac{1}{2}.$$

Bu değer grafikte $x = 0$ 'daki **sıçramanın boyu** olarak görünür: F_X , 0'ın hemen solunda 0, 0'da ise $\frac{1}{2}$ 'dir. Aralık olasılığı dağılım fonksiyonundan okunur:

$$P(0 < X \leq \frac{1}{2}) = F_X(\frac{1}{2}) - F_X(0) = \frac{3}{4} - \frac{1}{2} = \frac{1}{4};$$

denetim için doğrudan hesap: tura gelmeli ($\frac{1}{2}$) ve düzgün sayı $(0, \frac{1}{2}]$ 'ye düşmeli ($\frac{1}{2}$), çarpım $\frac{1}{4}$. Son olarak $X = \frac{1}{2}$ olması için tura gelip sayının tam $\frac{1}{2}$ çıkması gerekir; bunun olasılığı $\frac{1}{2} \cdot 0 = 0$ 'dir. Grafikte de $x = \frac{1}{2}$ 'de sıçrama yoktur.

Sınıflandırma. F_X ne basamaklı ne de sürekli: $x = 0$ 'da $\frac{1}{2}$ boyunda bir sıçrama yapar (kesikli kütle), ardından $(0, 1)$ boyunca $\frac{1}{2}$ eğimiyle düzgün yükselir (sürekli parça). Böyle bir X 'e **karma** rastgele değişken denir; olasılığının yarısı tek bir noktada, öteki yarısı bir aralığa yayılmış durumdadır.



Şekil 9.6: Üç dağılım fonksiyonu türü. Kesikli türde F basamaklıdır ve olasılık sıçramalarda toplanır; sürekli türde F hiç sıçramaz, olasılık aralığa yayılır; karma türde hem sıçrama hem eğimli parça vardır. Üçü de aynı üç özelliği taşır: azalmayan, sağdan sürekli, uç limitleri 0 ve 1. Fark yalnızca sıçramaların varlığındadır.

■

i Dağılım fonksiyonunun biçimi rastgele değişkenin türünü ele verir

Örnekler üç farklı davranış gösterdi. Sonlu değer kümeli X 'in dağılım fonksiyonu basamaklıdır: değerler arasında sabittir, her değerde sıçrar. Bir aralık üzerinde değer alan ω , $|\omega|$, ω^2 ve $\omega^2 - 2\omega$ 'nın dağılım fonksiyonları ise süreklidir ve olasılık bir aralığa yayılmıştır. Bu iki tür, sonraki iki bölümün konusu olan **kesikli** ve **süreklili** rastgele değişkenlerdir. Son örnekteki gibi ikisinin karışımı da mümkündür; **Örnek 9.8**'te bir karma örnek daha göreceğiz.

9.3 Dağılım Fonksiyonunun Temel Özellikleri

Bir fonksiyona bakınca onun bir rastgele değişkenin dağılım fonksiyonu olup olamayacağını söyleyebilmek isteriz; ayrıca ileride F_X 'ten olasılık hesaplarken sol limitlerin varlığına ihtiyacımız olacak. Bunun için F_X 'in her zaman taşıdığı yapısal özellikleri belirlemeliyiz. Bu özellikleri **Olasılık Uzaı Örnekleri ve Geometrik Olasılık** bölümünde, Borel olasılık ölçülerine karşılık gelen dağılım fonksiyonları için ispatlamıştık (**Teorem 4.1**); F_X de P_X ölçüsünün dağılım fonksiyonu olduğundan aynı sonuç ona da uygulanır. Aşağıda teoremi rastgele değişken diliyle yeniden ifade ediyor ve ispatı bu dile çeviriyoruz. İspatın motoru olasılık ölçüsünün sürekliliğidir (**Teorem 3.5**): iç içe olaylar dizisinin olasılığı, limit olayın olasılığına yakınsar.

Teorem 9.1 (Dağılım Fonksiyonunun Temel Özellikleri). X bir rastgele değişken ve F_X onun dağılım fonksiyonu olsun.

- (a) F_X azalmayıdır: $x_1 < x_2 \implies F_X(x_1) \leq F_X(x_2)$.
- (b) F_X sağdan süreklidir: her $x \in \mathbb{R}$ için $\lim_{h \rightarrow 0^+} F_X(x+h) = F_X(x)$.
- (c) $\lim_{x \rightarrow -\infty} F_X(x) = 0$ ve $\lim_{x \rightarrow +\infty} F_X(x) = 1$.

İspat.

(a) $x_1 < x_2$ olsun. $X(\omega) \leq x_1$ ise $X(\omega) \leq x_2$ de doğrudur; yani $(X \leq x_1) \subseteq (X \leq x_2)$. Olasılık ölçüsü monotondur (**Teorem 3.1**), dolayısıyla

$$F_X(x_1) = P(X \leq x_1) \leq P(X \leq x_2) = F_X(x_2).$$

(b) $x \in \mathbb{R}$ sabit olsun ve $n = 1, 2, 3, \dots$ için

$$A_n = \left(X \leq x + \frac{1}{n} \right) = \left\{ \omega \in \Omega : X(\omega) \leq x + \frac{1}{n} \right\}$$

olaylarını alalım. $x + \frac{1}{n+1} < x + \frac{1}{n}$ olduğundan (a)'daki gerekçeyle $A_1 \supseteq A_2 \supseteq A_3 \supseteq \dots$ olur; dizi azalandır. Kesişimleri $(X \leq x)$ 'tir: $X(\omega) \leq x + \frac{1}{n}$ eşitsizliği her n için sağlanıyorsa $n \rightarrow \infty$ alınarak $X(\omega) \leq x$ bulunur; tersine $X(\omega) \leq x$ ise her n için $X(\omega) \leq x + \frac{1}{n}$ 'dir. Azalan diziler için süreklilik (**Teorem 3.5**) gereği

$$\lim_{n \rightarrow \infty} P\left(x + \frac{1}{n}\right) = \lim_{n \rightarrow \infty} P(A_n) = P\left(\bigcap_{n=1}^{\infty} A_n\right) = P(X \leq x) = F_X(x).$$

Bu, sağdan limitin $\frac{1}{n}$ dizisi boyunca $F_X(x)$ 'e eşit olduğunu gösterir; keyfî $h \rightarrow 0^+$ için (a)'yı kullanırız. $\varepsilon > 0$ verilsin; yukarıdaki limit gereği $P\left(x + \frac{1}{N}\right) < F_X(x) + \varepsilon$ olacak biçimde bir N vardır. $0 < h < \frac{1}{N}$ için azalmayıktan

$$F_X(x) \leq F_X(x+h) \leq P\left(x + \frac{1}{N}\right) < F_X(x) + \varepsilon$$

olur. Demek ki $\lim_{h \rightarrow 0^+} F_X(x+h) = F_X(x)$.

(c) $n = 1, 2, 3, \dots$ için $B_n = (X \leq -n)$ olsun. $B_1 \supseteq B_2 \supseteq \dots$ azalan bir dizidir ve kesişimi boşdur: $X(\omega)$ bir reel sayı olduğundan her n için $X(\omega) \leq -n$ olamaz. Süreklilikten

$$\lim_{n \rightarrow \infty} F_X(-n) = \lim_{n \rightarrow \infty} P(B_n) = P(\emptyset) = 0.$$

Keyfî $x \rightarrow -\infty$ için yine (a)'ya başvururuz: $x \leq -n$ ise $0 \leq F_X(x) \leq F_X(-n)$ olur ve sağ taraf 0'a gittiğinden $\lim_{x \rightarrow -\infty} F_X(x) = 0$ 'dır.

Şimdi $C_n = (X \leq n)$ olsun. $C_1 \subseteq C_2 \subseteq \dots$ artan bir dizidir ve birleşimi Ω 'dır: her ω için $X(\omega)$ bir reel sayıdır, dolayısıyla $X(\omega) \leq n$ olacak biçimde bir n vardır. Artan diziler için süreklilikten

$$\lim_{n \rightarrow \infty} F_X(n) = \lim_{n \rightarrow \infty} P(C_n) = P\left(\bigcup_{n=1}^{\infty} C_n\right) = P(\Omega) = 1.$$

$x \geq n$ için $F_X(n) \leq F_X(x) \leq 1$ olduğundan $\lim_{x \rightarrow +\infty} F_X(x) = 1$ elde edilir.

■

Teoremin üç özelliği birlikte, F_X 'in grafiğinin nasıl görüldüğünü söyler: soldan 0 düzeyinden gelir, hiç aşağı inmeden yükselir, sağda 1 düzeyine yaklaşır; yükselirken sıçrayabilir, ama her sıçramada değeri üst basamağa aittir. Bu özellikler yalnızca gerekli değil, aynı zamanda **yeterlidir**: (a), (b), (c)'yi sağlayan her fonksiyon bir rastgele değişkenin dağılım fonksiyonudur. Bunu **Önerme 9.1**'de göstereceğiz.

⚠ Sağdan sürekli, ama soldan sürekli olmak zorunda değil

Örnek 9.1'ndeki dağılım fonksiyonunda $x \rightarrow 1^-$ iken $F_X(x) = \frac{1}{8}$, $x \rightarrow 1^+$ iken $F_X(x) = \frac{1}{2}$ ve $F_X(1) = \frac{1}{2}$ 'dir. Sağ limit fonksiyonun değerine eşittir, sol limit değildir. Sağdan sürekliliğin kaynağı, tanımdaki $(X \leq x)$ olayının x 'i **içermesidir**; x 'e sağdan yaklaşırken $(X \leq x + h)$ olayları $(X \leq x)$ 'e küçülür, ama soldan yaklaşırken $(X \leq x - h)$ olayları $(X = x)$ olayını hiç yakalayamaz. Dağılım fonksiyonu $P(X < x)$ diye tanımlansaydı soldan sürekli olurdu; bu bir gelenek seçimidir ve bu kitapta hep $P(X \leq x)$ kullanılır.

Sağdan sürekliliğin yanı sıra, her noktada sol limitin de var olduğunu bilmek gerekir; bu, sonraki teoremden kullanacağımız bir gösterimi tanımlamamıza olanak verir. F_X azalmayan ve $[0, 1]$ içinde sınırlı olduğundan, her $a \in \mathbb{R}$ için $x \rightarrow a^-$ iken $F_X(x)$ 'in limiti vardır ve

$$F_X(a^-) = \lim_{h \rightarrow 0^+} F_X(a - h) = \sup\{F_X(x) : x < a\}$$

ile gösterilir. Azalmayanlıktan $F_X(a^-) \leq F_X(a)$ 'dır; eşitlik tam olarak F_X 'in a 'da sürekli olduğu durumda sağlanır. $F_X(a) - F_X(a^-)$ farkı, grafiğin a noktasındaki **sıçramasıdır**.

9.4 Dağılım Fonksiyonundan Olasılık Hesabı

Dağılım fonksiyonu tanım gereği yalnızca $(X \leq x)$ biçimindeki olayların olasılığını verir. Uygulamada ise “ X 'in a ile b arasına düşme olasılığı”, “ X 'in tam olarak a olma olasılığı” gibi sorular sorulur. Aşağıdaki teorem bütün bu olasılıkların yalnızca F_X 'in değerleri ve sol limitleriyle hesaplandığını gösterir; aynı formülleri **Olasılık Uzayı Örnekleri ve Geometrik Olasılık** bölümünde ölçü dilinde görmüştük (**Önerme 4.4**); şimdi onları rastgele değişken diliyle ve eksiksiz olarak ispatlıyoruz. İspatlarda üç araç yeter: olayları ayrık parçalara bölüp toplamsallığı (A3) kullanmak, tümleyene geçmek ve nokta olasılığı için ölçünün sürekliliğine başvurmak.

Teorem 9.2 (Dağılım Fonksiyonuyla Aralık ve Nokta Olasılıkları). X bir rastgele değişken, $F = F_X$ onun dağılım fonksiyonu ve $a < b$ reel sayılar olsun. O zaman aşağıdaki eşitlikler sağlanır:

- (a) $P(a < X \leq b) = F(b) - F(a)$,
- (b) $P(X = a) = F(a) - F(a^-)$,
- (c) $P(X < a) = F(a^-)$,
- (d) $P(X > a) = 1 - F(a)$ ve $P(X \geq a) = 1 - F(a^-)$,
- (e) $P(a \leq X \leq b) = F(b) - F(a^-)$,
- (f) $P(a < X < b) = F(b^-) - F(a)$,
- (g) $P(a \leq X < b) = F(b^-) - F(a^-)$.

İspat.

(a) $(X \leq b)$ olayını, X 'in a 'yı aşıp aşmadığına göre ikiye bölelim:

$$(X \leq b) = (X \leq a) \cup (a < X \leq b).$$

Sağdaki iki olay ayrıktır, çünkü $X(\omega)$ aynı anda hem a 'yı aşmayıp hem de a 'dan büyük olamaz. Toplamsallık (A3) gereği

$$F(b) = P(X \leq b) = P(X \leq a) + P(a < X \leq b) = F(a) + P(a < X \leq b),$$

buradan $P(a < X \leq b) = F(b) - F(a)$.

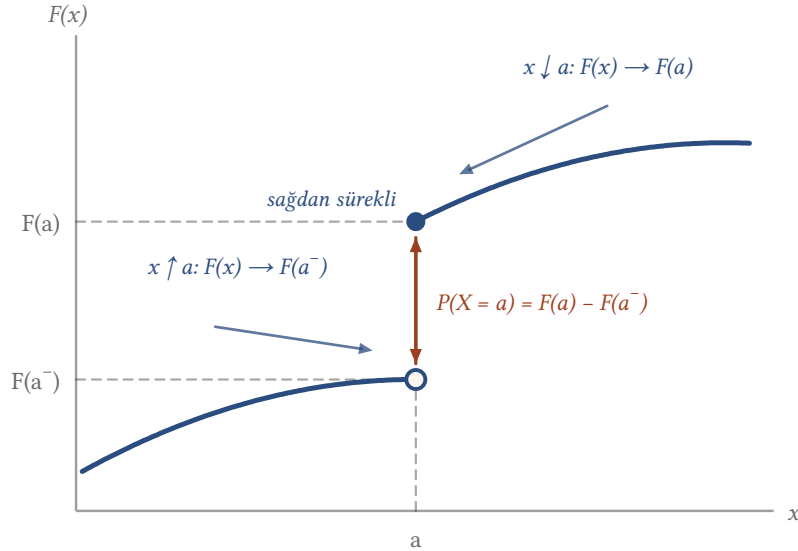
(b) $n = 1, 2, 3, \dots$ için

$$A_n = \left(a - \frac{1}{n} < X \leq a + \frac{1}{n} \right)$$

olaylarını alalım. n büyüdükçe aralık daralır; $A_1 \supseteq A_2 \supseteq \dots$ azalan bir dizidir. Kesişimi $(X = a)$ 'dır: her n için $a - \frac{1}{n} < X(\omega) \leq a + \frac{1}{n}$ ise $n \rightarrow \infty$ alınarak $a \leq X(\omega) \leq a$, yani $X(\omega) = a$ bulunur; tersine $X(\omega) = a$ ise her n için $a - \frac{1}{n} < a \leq a + \frac{1}{n}$ sağlanır. Süreklilik (Teorem 3.5) ve (a) gereği

$$P(X = a) = \lim_{n \rightarrow \infty} P(A_n) = \lim_{n \rightarrow \infty} \left[F\left(a + \frac{1}{n}\right) - F\left(a - \frac{1}{n}\right) \right].$$

Sağdan süreklilik (Teorem 9.1) gereği $F\left(a + \frac{1}{n}\right) \rightarrow F(a)$; $a - \frac{1}{n}$ dizisi a 'ya soldan yaklaştığından ve sol limit var olduğundan $F\left(a - \frac{1}{n}\right) \rightarrow F(a^-)$. İki limiti birleştirince $P(X = a) = F(a) - F(a^-)$.



Şekil 9.7: Bir dağılım fonksiyonunun a noktasındaki sıçraması. Soldan yaklaşırken $F(x)$, değer olarak alınmayan sol limit $F(a^-)$ 'ye (boş nokta) gider; sağdan yaklaşırken fonksiyonun gerçek değeri $F(a)$ 'ya (dolu nokta) gider: sağdan süreklilik budur. Sıçramanın yüksekliği $P(X = a) = F(a) - F(a^-)$ nokta olasılığıdır.

(c) $(X \leq a) = (X < a) \cup (X = a)$ ayrık bir birleşimdir; (A3) ve (b) ile

$$P(X < a) = P(X \leq a) - P(X = a) = F(a) - (F(a) - F(a^-)) = F(a^-).$$

(d) $(X > a)$ olayı $(X \leq a)$ olayının tümleyenidir; Teorem 3.1 gereği $P(X > a) = 1 - P(X \leq a) = 1 - F(a)$. Benzer biçimde $(X \geq a) = (X < a)^c$ olduğundan (c) ile $P(X \geq a) = 1 - F(a^-)$.

(e) $(a \leq X \leq b) = (X = a) \cup (a < X \leq b)$ ayrık birleşiminden, (b) ve (a) ile

$$P(a \leq X \leq b) = (F(a) - F(a^-)) + (F(b) - F(a)) = F(b) - F(a^-).$$

(f) $(a < X \leq b) = (a < X < b) \cup (X = b)$ ayrık birleşiminden, (a) ve (b) ile

$$P(a < X < b) = (F(b) - F(a)) - (F(b) - F(b^-)) = F(b^-) - F(a).$$

(g) $(a \leq X < b) = (X = a) \cup (a < X < b)$ ayrık birleşiminden, (b) ve (f) ile

$$P(a \leq X < b) = (F(a) - F(a^-)) + (F(b^-) - F(a)) = F(b^-) - F(a^-).$$

■

Formülleri ezberlemek yerine kuralı akılda tutmak daha kolaydır. Sağ uç b aralığa **dahilse** $F(b)$, **hariçse** $F(b^-)$ alınır; sol uç a dahilse $F(a^-)$, hariçse $F(a)$ çıkarılır. Örneğin $[a, b)$ aralığında sağ uç hariç, sol uç dahildir; formül $F(b^-) - F(a^-)$ olur.

! Nokta olasılığı, dağılım fonksiyonunun sıçramasıdır

Teoremin (b) maddesi bu bölümün en çok kullanılacak sonucudur: $P(X = a)$, F_X grafiğinin a noktasındaki sıçrama boyudur. Bundan iki sonuç çıkar. Birincisi, F_X bir a noktasında süreklirse $P(X = a) = 0$ 'dır; F_X her yerde süreklirse **hiçbir tek noktanın pozitif olasılığı yoktur. Örnek 9.3'**de $X(\omega) = \omega^2$ için $P(X = 1) = 0$ 'dır; oysa X 'in 1 değerini alması ($\omega = 1$) mümkündür. Olasılığı sıfır olan olay imkânsız olay demek değildir. İkincisi, F_X süreklirse (c)-(g)'deki bütün sol limitler fonksiyon değerlerine dönüşür ve $\leq$ ile $<$ arasındaki fark kaybolur: $P(a < X < b) = P(a \leq X \leq b) = F_X(b) - F_X(a)$. Sıçramalı dağılım fonksiyonlarında ise bu fark tam olarak uç noktaların olasılığı kadardır ve göz ardı edilemez.

Şimdi teoremi iki örnekte uygulayalım: önce sürekli bir dağılım fonksiyonunda, sonra sıçramalı bir dağılım fonksiyonunda.

Örnek 9.7 (Üstel Dağılım Fonksiyonuyla Olasılık Hesabı). X rastgele değişkeninin dağılım fonksiyonu

$$F_X(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-x}, & x \geq 0 \end{cases}$$

olsun. $P(X \leq -1)$, $P(X \leq 0)$, $P(X \leq 1)$, $P(X = 1)$, $P(0 < X \leq 1)$, $P(0 < X < 1)$ ve $P(X > 2)$ olasılıklarını hesaplayınız.

Çözüm.

Önce fonksiyonu tanıyalım. $x \geq 0$ için $1 - e^{-x}$, 0'dan başlayıp 1'e yaklaşan, kesin artan bir fonksiyondur; $x = 0$ 'da değeri $1 - e^0 = 0$ olduğundan iki parça sürekli biçimde birleşir. Dolayısıyla F_X bütün $\mathbb{R}$ üzerinde süreklidir ve her a için $F_X(a^-) = F_X(a)$ 'dır; **Teorem 9.2'**nin (b) maddesi gereği her nokta olasılığı sıfırdır. Bu fonksiyon ileride **üstel dağılım** adıyla karşımıza çıkacaktır.

Şimdi teoremi maddeler hâlinde uygulayalım.

$$P(X \leq -1) = F_X(-1) = 0, \text{ çünkü } -1 < 0.$$

$$P(X \leq 0) = F_X(0) = 1 - e^0 = 0.$$

$$P(X \leq 1) = F_X(1) = 1 - e^{-1} \approx 0,632.$$

$$P(X = 1) = F_X(1) - F_X(1^-) = 0, \text{ çünkü } F_X \text{ süreklidir.}$$

$$P(0 < X \leq 1) = F_X(1) - F_X(0) = (1 - e^{-1}) - 0 = 1 - e^{-1}.$$

$$P(0 < X < 1) = F_X(1^-) - F_X(0) = F_X(1) - F_X(0) = 1 - e^{-1}; \text{ süreklilik nedeniyle bir önceki olasılıkla aynıdır.}$$

$$P(X > 2) = 1 - F_X(2) = 1 - (1 - e^{-2}) = e^{-2} \approx 0,135.$$

İlk iki sonuç X 'in negatif değer alma olasılığının sıfır olduğunu, hatta $P(X \leq 0) = 0$ olduğunu söyler: bütün olasılık $(0, \infty)$ aralığına yayılmıştır. Son sonuç ise X 'in 2'yi aşma olasılığının yalnızca e^{-2} olduğunu, yani kütleinin büyük bölümünün 0'a yakın toplandığını gösterir.

■

Örnek 9.8 (Sıçramalı Dağılım Fonksiyonundan Olasılık Hesabı). $F : \mathbb{R} \rightarrow [0, 1]$ fonksiyonu

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{x+1}{4}, & 0 \leq x < 1 \\ \frac{3}{4}, & 1 \leq x < 2 \\ 1, & x \geq 2 \end{cases}$$

ile verilsin. F' 'nin Teorem 9.1'ndeki üç özelliği taşıdığını doğrulayınız. F' 'yi dağılım fonksiyonu kabul eden bir X rastgele değişkeni için $P(X = 0)$, $P(X = 1)$, $P(X = 2)$, $P(X < 1)$, $P(X \leq 1)$, $P(0 < X < 1)$, $P(\frac{1}{2} < X \leq \frac{3}{2})$, $P(X \geq 1)$ ve $P(X > 1)$ olasılıklarını hesaplayınız. Son olarak aşağıdaki toplamı bulunuz:

$$P(X = 0) + P(0 < X < 1) + P(X = 1) \\ + P(1 < X < 2) + P(X = 2).$$

Çözüm.

Özelliklerin doğrulanması. Parçalar sırasıyla $0, \frac{1}{4}$ 'ten $\frac{1}{2}$ 'ye artan doğru parçası, $\frac{3}{4}$ sabiti ve 1 sabitidir; her parça bir öncekinin bittiği yerin üstünden ya da aynı düzeyden başlar, öyleyse F azalmayandır. Sağdan süreklilik yalnızca parça sınırlarında denetlenmelidir: $x \rightarrow 0^+$ iken $\frac{x+1}{4} \rightarrow \frac{1}{4} = F(0)$; $x \rightarrow 1^+$ iken $F(x) = \frac{3}{4} = F(1)$; $x \rightarrow 2^+$ iken $F(x) = 1 = F(2)$. Uç limitler açıktır: $x < 0$ için $F(x) = 0$, $x \geq 2$ için $F(x) = 1$. Öyleyse F bir dağılım fonksiyonudur (Önerme 9.1 bunu bir rastgele değişkenle gerçekleştirmenin mümkün olduğunu garanti edecek).

Sol limitler. $F(0^-) = 0$, $F(1^-) = \lim_{x \rightarrow 1^-} \frac{x+1}{4} = \frac{1}{2}$, $F(2^-) = \frac{3}{4}$. Başka hiçbir noktada sıçrama yoktur.

Nokta olasılıkları. Teorem 9.2 (b) ile

$$P(X = 0) = F(0) - F(0^-) = \frac{1}{4}, \quad P(X = 1) = F(1) - F(1^-) = \frac{3}{4} - \frac{1}{2} = \frac{1}{4}, \quad P(X = 2) = F(2) - F(2^-) = 1 - \frac{3}{4} = \frac{1}{4}.$$

Öteki olasılıklar.

$P(X < 1) = F(1^-) = \frac{1}{2}$ ve $P(X \leq 1) = F(1) = \frac{3}{4}$; aradaki fark $P(X = 1) = \frac{1}{4}$ 'tür.

$P(0 < X < 1) = F(1^-) - F(0) = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$.

$P(\frac{1}{2} < X \leq \frac{3}{2}) = (F(\frac{3}{2}) - F(\frac{1}{2})) = \frac{3}{4} - \frac{3}{8} = \frac{3}{8}$.

$P(X \geq 1) = 1 - F(1^-) = \frac{1}{2}$ ve $P(X > 1) = 1 - F(1) = \frac{1}{4}$.

Toplam. $P(1 < X < 2) = F(2^-) - F(1) = \frac{3}{4} - \frac{3}{4} = 0$; F bu aralıkta sabit olduğundan X oraya hiç düşmez. Böylece

$$P(X = 0) + P(0 < X < 1) + P(X = 1) + P(1 < X < 2) + P(X = 2) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} + 0 + \frac{1}{4} = 1.$$

Bu beş olay ayırık olup birleşimleri $(0 \leq X \leq 2)$ 'dir; toplamın 1 çıkması, X 'in bütün olasılığının $[0, 2]$ aralığında olduğunu doğrular. Dağılımın yapısı da ortaya çıkmıştır: X , 0, 1 ve 2 değerlerinin her birini $\frac{1}{4}$ olasılıkla alır, geri kalan $\frac{1}{4}$ 'lük olasılık ise $(0, 1)$ aralığına düzgün biçimde yayılmıştır. Bu X ne kesikli ne de sürekli; karma türündendir.

■

9.5 Dağılım Fonksiyonu Olasılık Dağılımını Belirler

Bölüme, P_X 'in taşıdığı bilgiyi tek bir fonksiyona sıkıştırma vaadiyle başlamıştık. Teorem 9.2 bu vaadin aralıklar için tutulduğunu gösterdi. Ama Borel cebri aralıklardan çok daha zengindir; F_X bilinince $P_X(B)$ her Borel kümesi B için belirlenmiş olur mu? Bu sorunun yanıtı Olasılık Uzayı Örnekleri ve Geometrik Olasılık bölümündeki Borel–Stieltjes teoreminde saklıdır: (a), (b), (c) özelliklerini taşıyan bir fonksiyon, Borel cebri üzerinde $P((-\infty, x]) = F(x)$ koşulunu sağlayan bir tek olasılık ölçüsü belirler (Teorem 4.2). Bunu rastgele değişken diline çevirelim.

Önerme 9.1 (Dağılım Fonksiyonu Olasılık Dağılımını Tek Biçimde Belirler). (a) X ve Y , aynı ya da farklı olasılık uzayları üzerinde tanımlı iki rastgele değişken olsun. $F_X = F_Y$ ise $P_X = P_Y$ 'dir; yani her $B \in \mathcal{B}(\mathbb{R})$ için X 'in B 'ye düşme olasılığı ile Y 'nin B 'ye düşme olasılığı eşittir. Terside doğrudur.

(b) $F : \mathbb{R} \rightarrow [0, 1]$, Teorem 9.1'ndeki (a), (b), (c) özelliklerini taşıyan herhangi bir fonksiyon olsun. O zaman dağılım fonksiyonu F olan bir rastgele değişken vardır.

İspat.

(a) Önce kolay yön: $P_X = P_Y$ ise her x için $F_X(x) = P_X((-\infty, x]) = P_Y((-\infty, x]) = F_Y(x)$ 'tir.

Şimdi $F_X = F_Y$ olsun. P_X ve P_Y , $\mathcal{B}(\mathbb{R})$ üzerinde birer olasılık ölçüsüdür (Teorem 8.1) ve her x için

$$P_X((-\infty, x]) = F_X(x) = F_Y(x) = P_Y((-\infty, x])$$

sağlanır. F_X , **Teorem 9.1** gereği (a), (b), (c) özelliklerini taşır; **Teorem 4.2'**e göre bu özellikleri taşıyan bir fonksiyon için $P((-\infty, x]) = F_X(x)$ koşulunu sağlayan Borel olasılık ölçüsü tektir. P_X ve P_Y bu koşulu sağladığından $P_X = P_Y$ 'dir.

(b) **Teorem 4.2**, F 'ye karşılık gelen bir P_F Borel olasılık ölçüsü verir: her x için $P_F((-\infty, x]) = F(x)$. $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_F)$ olasılık uzayını alalım ve $X : \mathbb{R} \rightarrow \mathbb{R}$, $X(\omega) = \omega$ özdeşlik fonksiyonunu tanımlayalım. Her a için $(X \leq a) = (-\infty, a] \in \mathcal{B}(\mathbb{R})$ olduğundan X bir rastgele değişkendir ve

$$F_X(x) = P_F(X \leq x) = P_F((-\infty, x]) = F(x)$$

bulunur.

■

Önermenin (a) maddesi, “ X 'in dağılımını bilmek” ile “ F_X 'i bilmek” ifadelerinin eşanlamı olduğunu söyler; bundan böyle bir rastgele değişkenin dağılımını verirken çoğunlukla dağılım fonksiyonunu ya da ondan türeyen daha kullanışlı bir nesneyi (kesikli durumda olasılık fonksiyonunu, sürekli durumda yoğunluk fonksiyonunu) vereceğiz. (b) maddesi ise tersini söyler: (a), (b), (c) özellikleri dağılım fonksiyonu olmak için gerekli ve yeterlidir. İki madde birlikte, bu üç özelliği taşıyan fonksiyonların kümesi ile $\mathcal{B}(\mathbb{R})$ üzerindeki olasılık ölçülerinin kümesi arasında bire bir bir eşleme kurar: her dağılım fonksiyonuna bir tek Borel olasılık ölçüsü, her Borel olasılık ölçüsüne bir tek dağılım fonksiyonu karşılık gelir. Bir fonksiyonun bu üç özelliği taşıdığı doğrulanınca, “dağılım fonksiyonu şu olan bir X rastgele değişkeni” ifadesi meşrurdur; **Örnek 9.7** ve **Örnek 9.8**'te tam olarak bunu yaptık.

🔗 Aynı dağılım, farklı rastgele değişkenler

$F_X = F_Y$ olması $X = Y$ olması demek değildir. Üç para atışında X tura sayısı, $Y = 3 - X$ yazı sayısı olsun. $X \neq Y$ 'dir; hatta hiçbir ω için $X(\omega) = Y(\omega)$ olmaz, çünkü toplamları tek sayı 3'tür. Ama $P(Y = k) = P(X = 3 - k) = P(X = k)$ olduğundan ($k = 0, 3$ için $\frac{1}{8}$, $k = 1, 2$ için $\frac{3}{8}$) $F_X = F_Y$ 'dir. Böyle rastgele değişkenlere **aynı dağılımlı** denir. Olasılık teorisinin pek çok sonucu yalnızca dağılıma bağlıdır; bu yüzden bir rastgele değişkeni çoğu zaman üzerinde tanımlandığı Ω 'yı hiç anmadan, doğrudan dağılım fonksiyonuyla veririz.

9.6 Alıştırmalar

Alıştırma 9.1 (Dağılım Fonksiyonu Alıştırmaları). (a) Bir para bir kez atılıyor: $\Omega = \{Y, T\}$, $\mathcal{U} = \mathcal{P}(\Omega)$, $P(\{Y\}) = P(\{T\}) = \frac{1}{2}$. $X(Y) = 0$, $X(T) = 1$ olsun. F_X 'i bulunuz ve $P(X = 0)$, $P(X = 1)$ değerlerini F_X 'in sıçramalarından okuyunuz.

(b) $\Omega = [-1, 1]$, $\mathcal{U} = \mathcal{B}(\Omega)$, $P(A) = \frac{A \text{ 'nin uzunluğu}}{2}$ ve $X(\omega) = \omega^2$ olsun. F_X 'i bulunuz ve $(\frac{1}{4} < X \leq \frac{1}{2})$ olasılığını hesaplayınız.

(c) $\Omega = [0, 2]$, $\mathcal{U} = \mathcal{B}(\Omega)$, $P(A) = \frac{A \text{ 'nin uzunluğu}}{2}$ ve $X(\omega) = |\omega - 1|$ olsun. F_X 'i bulunuz.

(d) X rastgele değişkeninin dağılım fonksiyonu $x < 1$ için $F_X(x) = 0$, $x \geq 1$ için $F_X(x) = 1 - \frac{1}{x^2}$ olsun. $P(X \leq 2)$, $P(X > 3)$, $P(2 < X \leq 3)$, $P(X = 2)$ ve $P(X \geq 2)$ olasılıklarını hesaplayınız.

(e) Aşağıdaki fonksiyonlardan hangileri bir rastgele değişkenin dağılım fonksiyonudur? Olanlar için $P(X = 1)$ ve $P(X > 0)$ olasılıklarını bulunuz.

$$G_1(x) = \begin{cases} 0, & x < 0 \\ \frac{x}{2}, & 0 \leq x < 1 \\ 1, & x \geq 1, \end{cases} \quad G_2(x) = \begin{cases} 0, & x \leq 0 \\ 1, & x > 0, \end{cases}$$

$$G_3(x) = \frac{1}{1 + e^{-x}}, \quad G_4(x) = \begin{cases} 0, & x < 0 \\ 2x, & 0 \leq x < 1 \\ 1, & x \geq 1. \end{cases}$$

(f) **Örnek 9.1**'ndeki dağılım fonksiyonunu kullanarak $P(X < 2)$, $P(1 \leq X \leq 2)$, $P(X > 1)$ ve $P(0, 5 < X < 2, 5)$ olasılıklarını hesaplayınız.

(g) Bir dağılım fonksiyonunun süreksiz olduğu noktaların kümesinin en çok sayılabilir olduğunu gösteriniz.

Çözüm.

(a) $D_X = \{0, 1\}$. $x < 0$ için $(X \leq x) = \emptyset$; $0 \leq x < 1$ için $(X \leq x) = \{Y\}$; $x \geq 1$ için $(X \leq x) = \Omega$. Öyleyse

$$F_X(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{2}, & 0 \leq x < 1 \\ 1, & x \geq 1. \end{cases}$$

Sıçramalar: $P(X = 0) = F_X(0) - F_X(0^-) = \frac{1}{2} - 0 = \frac{1}{2}$ ve $P(X = 1) = F_X(1) - F_X(1^-) = 1 - \frac{1}{2} = \frac{1}{2}$.

(b) $\omega \in [-1, 1]$ için $\omega^2 \in [0, 1]$; $D_X = [0, 1]$. $0 \leq x < 1$ için $\omega^2 \leq x \Leftrightarrow -\sqrt{x} \leq \omega \leq \sqrt{x}$ ve $\sqrt{x} < 1$ olduğundan $[-\sqrt{x}, \sqrt{x}] \subseteq \Omega$; dolayısıyla

$$F_X(x) = P([- \sqrt{x}, \sqrt{x}]) = \frac{2\sqrt{x}}{2} = \sqrt{x}.$$

$x < 0$ için $F_X(x) = 0$, $x \geq 1$ için $F_X(x) = 1$ ($x = 1$ 'de $\sqrt{1} = 1$ olduğundan F_X sürekli). **Teorem 9.2** (a) ile

$$\left(\frac{1}{4} < X \leq \frac{1}{2}\right) = F_X\left(\frac{1}{2}\right) - F_X\left(\frac{1}{4}\right) = \sqrt{\frac{1}{2}} - \frac{1}{2} = \frac{\sqrt{2}-1}{2} \approx 0,207.$$

(c) $\omega \in [0, 2]$ için $|\omega - 1| \in [0, 1]$; $D_X = [0, 1]$. $0 \leq x < 1$ için

$$|\omega - 1| \leq x \Leftrightarrow 1 - x \leq \omega \leq 1 + x$$

ve $0 \leq 1 - x, 1 + x \leq 2$ olduğundan $[1 - x, 1 + x] \subseteq \Omega$ 'dir:

$$F_X(x) = P([1 - x, 1 + x]) = \frac{2x}{2} = x.$$

Böylece

$$F_X(x) = \begin{cases} 0, & x < 0 \\ x, & 0 \leq x < 1 \\ 1, & x \geq 1. \end{cases}$$

X , olasılığı $[0, 1]$ aralığına düzgün yayan bir rastgele değişkendir: $P(a < X \leq b) = b - a$ olur ($0 \leq a < b \leq 1$).

(d) $x \geq 1$ için $1 - \frac{1}{x^2}$ sürekli ve artandır, $x = 1$ 'de 0 değerini alır, $x \rightarrow \infty$ iken 1'e yaklaşır; F_X her yerde sürekli. O hâlde

$$P(X \leq 2) = F_X(2) = 1 - \frac{1}{4} = \frac{3}{4}, \quad P(X > 3) = 1 - F_X(3) = \frac{1}{9},$$

$$P(2 < X \leq 3) = F_X(3) - F_X(2) = \frac{8}{9} - \frac{3}{4} = \frac{32 - 27}{36} = \frac{5}{36},$$

$$P(X = 2) = F_X(2) - F_X(2^-) = 0, \quad P(X \geq 2) = 1 - F_X(2^-) = 1 - \frac{3}{4} = \frac{1}{4}.$$

(e) Üç özelliği tek tek deneriz.

G_1 : 0'dan $\frac{1}{2}$ 'ye artan bir doğru parçası, sonra 1 sabiti; azalmayandır. $G_1(0) = 0 = \lim_{x \rightarrow 0^+} \frac{x}{2}$ ve $G_1(1) = 1 = \lim_{x \rightarrow 1^+} G_1(x)$; sağdan sürekli. Uç limitler 0 ve 1'dir. **Dağılım fonksiyonudur.** $P(X = 1) = G_1(1) - G_1(1^-) = 1 - \frac{1}{2} = \frac{1}{2}$ ve $P(X > 0) = 1 - G_1(0) = 1$.

G_2 : azalmayandır ve uç limitleri doğrudur, ama $G_2(0) = 0$ iken $\lim_{h \rightarrow 0^+} G_2(h) = 1$ 'dir; 0'da sağdan sürekli değildir. **Dağılım fonksiyonu değildir.** (Tanımı $x < 0$ için 0, $x \geq 0$ için 1 olarak değiştirilirse olur: bu, $P(X = 0) = 1$ olan bir rastgele değişkenin dağılım fonksiyonudur.)

G_3 : her x için $0 < G_3(x) < 1$; e^{-x} azalan olduğundan G_3 artandır; sürekli olduğundan sağdan sürekli; $x \rightarrow -\infty$ iken $e^{-x} \rightarrow \infty$ ve $G_3(x) \rightarrow 0$, $x \rightarrow +\infty$ iken $e^{-x} \rightarrow 0$ ve $G_3(x) \rightarrow 1$. **Dağılım fonksiyonudur.** Sürekli olduğundan $P(X = 1) = 0$; $P(X > 0) = 1 - G_3(0) = 1 - \frac{1}{2} = \frac{1}{2}$.

$G_4: x \rightarrow 1^-$ iken $2x \rightarrow 2 > 1 = G_4(1)$; fonksiyon 1'de düşer, azalmayan değildir. Üstelik $\frac{1}{2} < x < 1$ için $G_4(x) > 1$ olduğundan değerleri $[0, 1]$ 'in dışına çıkar. **Dağılım fonksiyonu değildir.**

(f) F_X için **Örnek 9.1**'ndaki basamak fonksiyonunu ve sol limitleri $F_X(1^-) = \frac{1}{8}$, $F_X(2^-) = \frac{1}{2}$, $F_X(2, 5^-) = F_X(2, 5) = \frac{7}{8}$ kullanırız.

$$P(X < 2) = F_X(2^-) = \frac{1}{2}.$$

$$P(1 \leq X \leq 2) = F_X(2) - F_X(1^-) = \frac{7}{8} - \frac{1}{8} = \frac{3}{4}.$$

$$P(X > 1) = 1 - F_X(1) = 1 - \frac{1}{2} = \frac{1}{2}.$$

$$P(0, 5 < X < 2, 5) = F_X(2, 5^-) - F_X(0, 5) = \frac{7}{8} - \frac{1}{8} = \frac{3}{4}.$$

Sonuçlar doğrudan sayımla uyuşur: $P(1 \leq X \leq 2) = P(X = 1) + P(X = 2) = \frac{3}{8} + \frac{3}{8} = \frac{3}{4}$ ve $(0, 5 < X < 2, 5) = (1 \leq X \leq 2)$ 'dir.

(g) F bir dağılım fonksiyonu ve X dağılım fonksiyonu F olan bir rastgele değişken olsun (**Önerme 9.1**). F sağdan sürekli olduğundan, a noktasında süreksiz olması demek $p_a = F(a) - F(a^-) > 0$ olması, yani **Teorem 9.2** (b) gereği $P(X = a) > 0$ olması demektir. $n \in \mathbb{N}$ için

$$D_n = \left\{ a \in \mathbb{R} : p_a > \frac{1}{n} \right\}$$

kümesini alalım. D_n içinden birbirinden farklı $a_1, \dots, a_k$ noktaları seçilirse $(X = a_1), \dots, (X = a_k)$ olayları ayrık ve toplamsallık ile

$$\frac{k}{n} < p_{a_1} + \dots + p_{a_k} = P(X \in \{a_1, \dots, a_k\}) \leq 1$$

olur; yani $k < n$. Demek ki D_n en çok $n - 1$ elemanlıdır, sonludur. Süreksizlik noktalarının kümesi

$$D = \{a : p_a > 0\} = \bigcup_{n=1}^{\infty} D_n$$

sonlu kümelerin sayılabilir birleşimidir, dolayısıyla en çok sayılabilir. Sonuç olarak bir dağılım fonksiyonu en çok sayılabilir çoklukta noktada sıçrayabilir; geri kalan her noktada sürekli.

■

Dağılım fonksiyonu her rastgele değişken için vardır ve dağılımı tek biçimde belirler. Örneklerde onun iki temel biçimde ortaya çıktığını gördük: basamaklı ve sürekli. İkisinin karışımı da mümkündür. Bir sonraki bölümde birinci biçimi, yani sayılabilir çoklukta değer alan rastgele değişkenleri ele alıyor ve dağılım fonksiyonunun sıçramalarını bir tabloya dönüştüren olasılık fonksiyonunu tanımlıyoruz: **Kesikli Rastgele Değişkenler ve Olasılık Fonksiyonu.**

10. Kesikli Rastgele Değişkenler ve Olasılık Fonksiyonu

Dağılım Fonksiyonu bölümünde bir rastgele değişkenin olasılık dağılımını tek bir gerçel fonksiyonla, $F_X(x) = P(X \leq x)$ dağılım fonksiyonuyla betimlemeyi öğrendik. Dağılım fonksiyonu her rastgele değişken için tanımlıdır; azalmayan, sağdan sürekli ve uçlarda 0 ile 1'e yaklaşan bir fonksiyondur. Ne var ki uygulamada karşılaşılan rastgele değişkenlerin büyük bölümü iki ana sınıftan birine düşer ve her sınıf için dağılımı betimlemenin daha somut, daha kullanışlı bir yolu vardır.

Bu bölümde ilk sınıfı ele alıyoruz: değer kümesi sonlu ya da sayılabilir sonsuz olan **kesikli rastgele değişkenler**. Zarın gösterdiği sayı, üç para atışında gelen tura sayısı, bir kutudan çekilen kusurlu parça sayısı, bir santrale bir saatte gelen çağrı sayısı: bunların hepsi kesiklidir. Böyle bir değişkenin dağılımını anlatmak için, alabileceği her değere ne olasılıkla ulaşıldığını söylemek yeter. Bu bilgiyi taşıyan fonksiyona **olasılık fonksiyonu** diyeceğiz.

Bölümde şunları öğreneceğiz: kesikli rastgele değişkenin ve olasılık fonksiyonunun tanımı; olasılık fonksiyonunu niteleyen iki özellik ve bunların Kolmogorov aksiyomlarından nasıl çıktığı; olasılık fonksiyonuyla dağılım fonksiyonu arasında iki yönlü geçiş (F toplamla, f sıçramayla elde edilir); ve bu geçişleri somutlaştıran altı çözümlü örnek. Bölümün sonunda kesikli değişkenlerin ayırt edici özelliğini, tek bir noktanın olasılığının sıfırdan farklı olabilmesini, bir sonraki bölümün sürekli değişkenleriyle karşılaştırmaya hazır olacağız.

10.1 Kesikli Rastgele Değişken

Bir X rastgele değişkeninin değer kümesi $D_X = X(\Omega) = \{X(\omega) : \omega \in \Omega\}$ idi (Tanım 8.2). Rastgele değişkenleri sınıflandırmanın en doğal yolu bu kümenin büyüklüğüne bakmaktır.

Tanım 10.1 (Kesikli Rastgele Değişken). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı ve $X : \Omega \rightarrow \mathbb{R}$ bir rastgele değişken olsun. X 'in değer kümesi D_X **sonlu ya da sayılabilir sonsuz** ise X 'e **kesikli rastgele değişken**, X 'in belirlediği P_X olasılık dağılımına da (Tanım 8.3) **kesikli dağılım** denir.

Tanımdaki koşul, D_X 'in elemanlarının bir dizi hâlinde sıralanabilmesi demektir:

$$D_X = \{x_1, x_2, x_3, \dots\}.$$

Sonlu durumda dizi bir yerde biter ve $D_X = \{x_1, \dots, x_n\}$ olur; sayılabilir sonsuz durumda ise bitmez. Aşağıdaki gözlemler tanımı yerine oturtur.

- Sonlu ya da sayılabilir bir örnek uzay üzerinde tanımlı her rastgele değişken kesiklidir; çünkü $X(\Omega)$ görüntü kümesinin eleman sayısı Ω 'ninkini aşamaz. Zar atma, para atma, torbadan bilye çekme gibi bütün sayma deneyleri bu sınıfa girer.
- Örnek uzayın sonsuz olması değişkenin kesikli olmasına engel değildir. Bir para ilk tura gelene kadar atılsın ve X atış sayısını göstere. Örnek uzay sonsuzdur; ama $D_X = \{1, 2, 3, \dots\} = \mathbb{N}$ sayılabilir olduğundan X kesiklidir.
- Değer kümesinin tam sayılardan oluşması gerekmez. $D_X = \{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots\}$ gibi bir küme de sayılabilir. Önemli olan, değerlerin bir dizi hâlinde sıralanabilmesidir.

Kesikli bir değişkenin dağılımını ayırt eden şey, toplam 1 birimlik olasılığın sayılabilir sayıda noktaya “yığılmış” olmasıdır. Bu yığılmayı nokta nokta kaydeden fonksiyonu şimdi tanımlıyoruz.

10.2 Olasılık Fonksiyonu

Kesikli bir X değişkeninin alabileceği her x değeri için $(X = x) = \{\omega \in \Omega : X(\omega) = x\}$ kümesi bir olaydır: $\{x\}$ bir Borel kümesidir (Önerme 2.5) ve Borel ölçütü gereği (Önerme 8.2) her Borel kümesinin ters görüntüsü $\mathcal{U}$ 'ya aittir. Dolayısıyla $P(X = x)$ olasılığı anlamlıdır.

Tanım 10.2 (Olasılık Fonksiyonu). X kesikli bir rastgele değişken olsun.

$$f_X : D_X \rightarrow \mathbb{R}, \quad f_X(x) = P(X = x) = P(\{\omega \in \Omega : X(\omega) = x\})$$

fonksiyonuna X rastgele değişkeninin **olasılık fonksiyonu** (olasılık kütle fonksiyonu) denir. Bağlam açıkken f_X yerine kısaca f yazılır.

Adlandırmadaki “kütle” sözcüğü şu benzetmeye dayanır: toplam 1 birimlik olasılık kütle D_X ’in noktalarına dağıtılmıştır ve $f(x)$, tam olarak x noktasına konan kütledir.

Çoğu zaman f ’yi, D_X dışında 0 değerini alacak biçimde bütün $\mathbb{R}$ ’ye genişletmek uygun olur:

$$f(x) = \begin{cases} P(X = x), & x \in D_X \\ 0, & x \notin D_X. \end{cases}$$

Bu genişletme yeni bir bilgi eklemeyiz; $x \notin D_X$ için $(X = x) = \emptyset$ olduğundan $P(X = x) = 0$ zaten doğrudur. Ama f ’yi $\mathbb{R}$ üzerinde tanımlı bir fonksiyon olarak görmek, toplamaların sınırlarını yazmayı ve bir sonraki bölümdeki kesikli-süreklili karşılaştırmasını kolaylaştırır. Bundan sonra f ’yi ihtiyaca göre iki biçimde de kullanacağız.

Hangi fonksiyonlar bir olasılık fonksiyonu olabilir? Aşağıdaki önerme iki basit koşulun **gerekli** olduğunu söyler ve buna, olasılık fonksiyonunun dağılımı tümüyle belirlediğini gösteren üçüncü bir özellik ekler. Hemen ardından iki koşulun **yeterli** de olduğunu göreceğiz.

Önerme 10.1 (Olasılık Fonksiyonunun Özellikleri). X kesikli bir rastgele değişken ve f onun olasılık fonksiyonu olsun. O zaman:

- (1) Her $x \in D_X$ için $f(x) \geq 0$.
- (2) $\sum_{x \in D_X} f(x) = 1$.
- (3) Her $B \in \mathcal{B}(\mathbb{R})$ Borel kümesi için

$$P(X \in B) = P_X(B) = \sum_{x \in B \cap D_X} f(x).$$

Özel olarak X ’in olasılık dağılımı P_X , olasılık fonksiyonu f tarafından tümüyle belirlenir.

İspat.

$D_X = \{x_1, x_2, \dots\}$ yazalım ve $i = 1, 2, \dots$ için $A_i = (X = x_i)$ olaylarını düşünelim. Sonlu durumda dizi n terimde biter; aşağıdaki hesaplar sonlu toplamalarla aynen geçerlidir.

- (1) $f(x) = P(X = x)$ bir olayın olasılığıdır; Tanım 3.1’ndeki (A1) aksiyomu gereği negatif olamaz.
- (2) Önce A_i olaylarının **ikişer ikişer ayrık** olduğunu görelim. $\omega \in A_i \cap A_j$ olsaydı hem $X(\omega) = x_i$ hem $X(\omega) = x_j$ olurdu; X bir fonksiyon olduğundan bu ancak $x_i = x_j$, yani $i = j$ iken mümkündür. Demek ki $i \neq j$ için $A_i \cap A_j = \emptyset$.

Sonra A_i olaylarının **birleşiminin Ω** olduğunu görelim. Her $\omega \in \Omega$ için $X(\omega) \in X(\Omega) = D_X$ ’dir; yani $X(\omega) = x_i$ olan bir i vardır ve $\omega \in A_i$ ’dir. Öyleyse

$$\Omega = \bigcup_i A_i.$$

Sayılabılır toplamsallık aksiyomu (A3) ve $P(\Omega) = 1$ (A2) ile

$$1 = P(\Omega) = P\left(\bigcup_i A_i\right) = \sum_i P(A_i) = \sum_i f(x_i) = \sum_{x \in D_X} f(x).$$

- (3) B bir Borel kümesi olsun. $\omega \in (X \in B)$ olması, $X(\omega) \in B$ olması demektir; $X(\omega)$ zaten D_X ’te olduğundan bu, $X(\omega) \in B \cap D_X$ demekle aynıdır. Öyleyse

$$(X \in B) = \bigcup_{x_i \in B \cap D_X} (X = x_i) = \bigcup_{x_i \in B \cap D_X} A_i.$$

Sağdaki birleşim, D_X ’in bir alt kümesi üzerinden alındığından sayılabilir; olaylar ikişer ikişer ayrık olduğundan (A3) uygulanır:

$$P(X \in B) = \sum_{x_i \in B \cap D_X} P(A_i) = \sum_{x \in B \cap D_X} f(x).$$

$B \cap D_X = \emptyset$ ise sağdaki toplam boş toplamdır ve 0'a eşittir; bu da $(X \in B) = \emptyset$ ile uyumludur. Son olarak $P_X(B) = P(X \in B)$ eşitliği olasılık dağılımının tanımıdır (Tanım 8.3); dolayısıyla P_X 'in her Borel kümesindeki değeri yalnızca f 'ye bağlıdır.

■

Önermenin (3) maddesi kesikli değişkenlerle çalışmanın temel aracıdır: **bir olayın olasılığı, o olaya düşen değerlerdeki olasılık fonksiyonu değerlerinin toplamıdır**. Bir aralığın, bir tek sayılar kümesinin, bir eşitsizliğin olasılığı hep aynı yolla hesaplanır. Bu bölümdeki bütün örnekler bu ilkenin uygulamasıdır.

i Bu iki özellik yeterlidir

$D = \{x_1, x_2, \dots\}$ sayılabilir bir gerçel sayı kümesi ve $g : D \rightarrow \mathbb{R}$, her $x \in D$ için $g(x) \geq 0$ ve $\sum_{x \in D} g(x) = 1$ koşullarını sağlayan bir fonksiyon olsun. O zaman g , kesikli bir rastgele değişkenin olasılık fonksiyonudur. Gerçekten, $\Omega = D$, $\mathcal{U} = \mathcal{P}(D)$ alalım ve $A \subseteq D$ için

$$P(A) = \sum_{x \in A} g(x)$$

tanımlayalım. $g \geq 0$ olduğundan $P(A) \geq 0$; $P(\Omega) = P(D) = \sum_{x \in D} g(x) = 1$; ikişer ikişer ayrık $A_1, A_2, \dots$ kümeleri için $P(\bigcup A_i) = \sum_{x \in \bigcup A_i} g(x) = \sum_i \sum_{x \in A_i} g(x) = \sum_i P(A_i)$, çünkü negatif olmayan terimli bir seride toplama sırası değiştirilebilir. Demek ki P bir olasılık ölçüsüdür. $X(\omega) = \omega$ özdeşlik fonksiyonu alınırsa $D_X = D$ ve $P(X = x) = P(\{x\}) = g(x)$ olur.

Bu gözlemin pratik sonucu şudur: olasılık fonksiyonu, olasılık uzayından bağımsız olarak **kendi başına bir nesnedir**. Negatif olmayan ve toplamı 1 olan bir sayı tablosu vermek, kesikli bir dağılım vermek demektir. Bu yüzden ileride dağılımları çoğu zaman altta yatan deneyi anmadan, doğrudan olasılık fonksiyonuyla tanıtacağız.

10.3 Olasılık Fonksiyonu ile Dağılım Fonksiyonu

Bir rastgele değişkenin olasılık dağılımını dağılım fonksiyonu $F_X(x) = P(X \leq x)$ ile betimlemeyi öğrenmiştik (Tanım 9.1). Kesikli değişkende artık elimizde ikinci bir betimleme, olasılık fonksiyonu var. İkisinin aynı bilgiyi taşıdığını, birinden ötekine kolayca geçilebildiğini göstermek gerekir; yoksa aynı nesne için iki ayrı dil kullanmış oluruz. Aşağıdaki önerme bu geçişi iki yönde de verir. Gösterimde $F(x^-)$, F 'nin x noktasındaki **sol limitini** gösterir:

$$F(x^-) = \lim_{t \rightarrow x^-} F(t).$$

F azalmayan ve $[0, 1]$ içinde sınırlı olduğundan bu limit her noktada vardır (Teorem 9.1).

Önerme 10.2 (Kesikli Değişkende Dağılım Fonksiyonu). X kesikli bir rastgele değişken, $D_X = \{x_1, x_2, \dots\}$ değer kümesi, f olasılık fonksiyonu ve F dağılım fonksiyonu olsun. O zaman:

(1) Her $x \in \mathbb{R}$ için

$$F(x) = P(X \leq x) = \sum_{x_i \in D_X, x_i \leq x} f(x_i).$$

(2) Her $x \in \mathbb{R}$ için

$$F(x^-) = P(X < x) = \sum_{x_i \in D_X, x_i < x} f(x_i).$$

(3) Her $x \in \mathbb{R}$ için $P(X = x) = F(x) - F(x^-)$; özel olarak

$$f(x) = F(x) - F(x^-), \quad x \in D_X,$$

ve $x \notin D_X$ noktalarında F süreklidir.

(4) $D_X = \{x_1, \dots, x_n\}$ sonlu ve $x_1 < x_2 < \dots < x_n$ ise F bir basamak fonksiyonudur:

$$F(x) = \begin{cases} 0, & x < x_1 \\ f(x_1) + \dots + f(x_k), & x_k \leq x < x_{k+1}, k = 1, \dots, n-1 \\ 1, & x \geq x_n. \end{cases}$$

İspat.

(1) **Önerme 10.1'nin** (3) maddesini $B = (-\infty, x]$ Borel kümesine uygulayalım. $(X \in B) = (X \leq x)$ ve $B \cap D_X = \{x_i \in D_X : x_i \leq x\}$ olduğundan

$$F(x) = P(X \leq x) = \sum_{x_i \in D_X, x_i \leq x} f(x_i).$$

(2) Önce $F(x^-) = P(X < x)$ eşitliğini gösterelim. $n = 1, 2, \dots$ için $C_n = (X \leq x - \frac{1}{n})$ olaylarını alalım. $x - \frac{1}{n}$ sayıları n ile birlikte arttığından $C_1 \subseteq C_2 \subseteq \dots$ olur; yani (C_n) artan bir olay dizisidir. Birleşimi $(X < x)$ olayıdır: $X(\omega) < x$ ise $x - X(\omega) > 0$ 'dır ve $\frac{1}{n} \leq x - X(\omega)$ olan bir n vardır, yani $\omega \in C_n$; tersine $\omega \in C_n$ ise $X(\omega) \leq x - \frac{1}{n} < x$ 'tir. Olasılık ölçüsünün sürekliliği (**Teorem 3.5**) gereği

$$P(X < x) = P\left(\bigcup_n C_n\right) = \lim_{n \rightarrow \infty} P(C_n) = \lim_{n \rightarrow \infty} \left(Fx - \frac{1}{n}\right).$$

F azalmayan olduğundan sol limiti $F(x^-)$ vardır ve x 'e soldan yaklaşan her dizi boyunca F 'nin limiti bu değere eşittir; özel olarak $\lim_{n \rightarrow \infty} (Fx - \frac{1}{n}) = F(x^-)$. Böylece $F(x^-) = P(X < x)$ bulunur. Şimdi **Önerme 10.1'nin** (3) maddesini $B = (-\infty, x)$ kümesine uygulamak toplam ifadesini verir.

(3) $(X \leq x)$ olayı, ayrık $(X < x)$ ve $(X = x)$ olaylarının birleşimidir. Sonlu toplamsallık (**Teorem 3.1**) ile

$$F(x) = P(X \leq x) = P(X < x) + P(X = x) = F(x^-) + P(X = x),$$

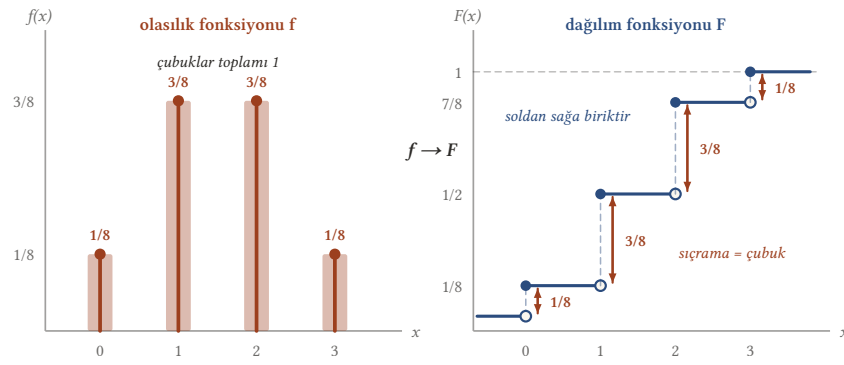
yani $P(X = x) = F(x) - F(x^-)$. Bu, **Teorem 9.2'de** her rastgele değişken için verilen eşitliğin ta kendisidir. $x \in D_X$ ise sol taraf $f(x)$ 'tir. $x \notin D_X$ ise $(X = x) = \emptyset$ olduğundan $F(x) - F(x^-) = 0$, yani $F(x^-) = F(x)$ olur; F zaten sağdan sürekli olduğundan (**Teorem 9.1**) F , x noktasında süreklidir.

(4) (1) maddesindeki toplamı üç durumda değerlendirelim. $x < x_1$ ise $x_i \leq x$ koşulunu sağlayan hiçbir x_i yoktur; toplam boştur ve $F(x) = 0$ olur. $x_k \leq x < x_{k+1}$ ise $x_i \leq x$ koşulunu tam olarak $x_1, \dots, x_k$ sağlar ve $F(x) = f(x_1) + \dots + f(x_k)$ olur. $x \geq x_n$ ise bütün x_i 'ler koşulu sağlar ve **Önerme 10.1'nin** (2) maddesi gereği $F(x) = \sum_{x_i \in D_X} f(x_i) = 1$ olur.

■

Önermenin söylediği şudur: f verilmişse F , değerleri **soldan sağa biriktirerek** bulunur; F verilmişse f , F 'nin **sıçramalarını okuyarak** bulunur. Sıçramanın büyüklüğü $F(x) - F(x^-)$ noktanın olasılığıdır; F 'nin sürekli olduğu noktaların olasılığı sıfırdır. Bu iki işlem birbirinin tersidir, dolayısıyla kesikli bir değişkenin dağılımını vermek için f 'yi (çoğu zaman bir tablo hâlinde) yazmak yeterlidir.

Grafiklerde durum şöyle görünür. f 'nin grafiği, D_X noktalarına dikilmiş, yükseklikleri $f(x_i)$ olan çubuklardan oluşur; çubukların toplam yüksekliği 1'dir. F 'nin grafiği ise bir **merdivendir**: iki ardışık değer arasında yataydır, her x_i noktasında $f(x_i)$ kadar yukarı sıçrar ve sağdan sürekli olduğu için her basamağın **sol ucu** dolu, bir önceki basamağın **sağ ucu** boştur. (4) maddesi, sonlu değer kümesi için bu merdivenin denklemini verir; D_X 'in noktaları artan bir dizi hâlinde sıralanabiliyorsa – bu bölümdeki bütün örneklerde durum budur – sayılabilir sonsuz değer kümesinde de merdivenin sonsuz sayıda basamağı vardır ve yükseklikleri 1'e yaklaşır.

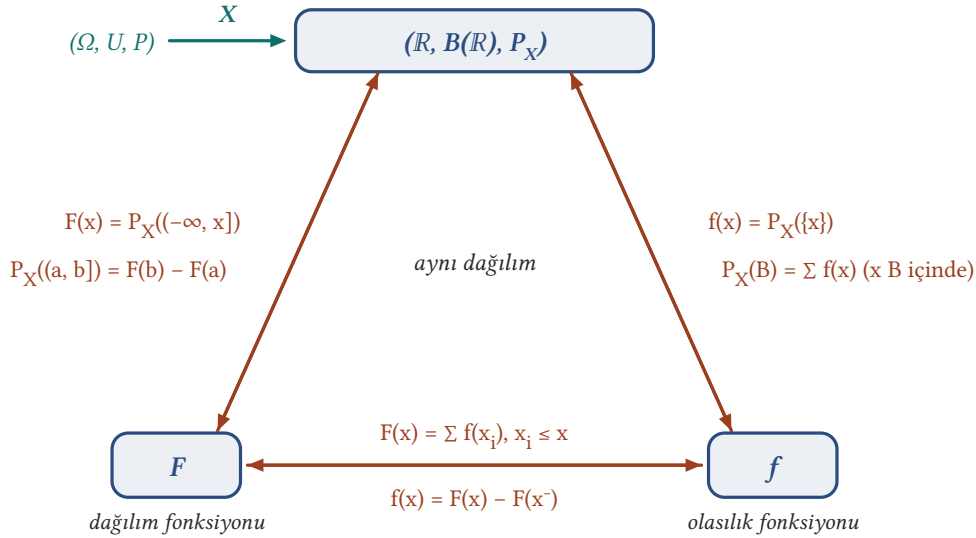


Şekil 10.1: Olasılık fonksiyonundan dağılım fonksiyonuna geçiş; üç para atışında tura sayısı. Soldaki çubuklar değerleri soldan sağa biriktirilince sağdaki merdiven çıkar: $F(1) = 1/8 + 3/8 = 1/2$. Her çubuk, sağda aynı noktadaki sıçramanın yüksekliği olarak yeniden görünür; bu yüzden F verildiğinde sıçramaları okumak f yi geri verir.

🔗 Kesikli değişkende kavramlar nasıl birbirine bağlanır

Bir kesikli X değişkeni için karşımıza çıkan nesneler ve aralarındaki geçişler şöyle özetlenebilir.

- $X : \Omega \rightarrow \mathbb{R}$ fonksiyonu $(\Omega, \mathcal{U}, P)$ olasılık uzayını $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$ uzayına taşır; taşıma kuralı $P_X(B) = P(X \in B)$ 'dir.
- P_X ile F birbirini belirler: $F(x) = P_X((-\infty, x])$ ve $P_X((a, b]) = F(b) - F(a)$.
- P_X ile f birbirini belirler: $f(x) = P_X(\{x\})$ ve $P_X(B) = \sum_{x \in B \cap D_X} f(x)$.
- F ile f birbirini belirler: $F(x) = \sum_{x_i \leq x} f(x_i)$ ve $f(x) = F(x) - F(x^-)$.



Şekil 10.2: Kesikli bir rastgele değişkenin dağılımını betimleyen üç nesne. X fonksiyonu olasılık uzayını sayı doğrusuna taşır ve P_X ölçüsünü doğurur; bu ölçü, dağılım fonksiyonu F ve olasılık fonksiyonu f birbirini belirler. Kenarlardaki formüller her geçişin iki yönünü verir: f den F ye biriktirerek, F den f ye sıçramaları okuyarak geçilir.

Üçü de aynı dağılımı betimler; hangisinin kullanılacağı soruya bağlıdır. Tek tek değerlerin olasılığı soruluyorsa f daha kullanışlıdır; “en çok”, “en az”, “arasında” biçimindeki sorularda ise F işi kolaylaştırır.

! Kesikli değişkende tek bir noktanın olasılığı sıfırdan farklı olabilir

Kesikli bir değişken için $P(X = x) = f(x)$ 'tir ve $x \in D_X$ noktalarında bu değer genellikle pozitiftir. Bunun iki önemli sonucu vardır.

Birincisi, **uç noktalar önemlidir**. $a < b$ için dört olasılık

$$P(a < X \leq b) = F(b) - F(a), \quad P(a \leq X \leq b) = F(b) - F(a^-),$$

$$P(a < X < b) = F(b^-) - F(a), \quad P(a \leq X < b) = F(b^-) - F(a^-)$$

birbirinden farklı olabilir (Teorem 9.2). Düzgün bir zar için $P(1 \leq X \leq 2) = \frac{2}{6}$ ama $P(1 < X < 2) = 0$ 'dır. Bir olasılık hesaplarken eşitsizliklerin “<” mi “≤” mi olduğuna her zaman bakılır.

İkincisi, **dağılım fonksiyonu süreksizdir**: F , olasılık taşıyan her noktada sıçrar ve grafiği bir merdivendir.

Bir sonraki bölümde göreceğimiz sürekli rastgele değişkenlerde durum tam tersidir: her x için $P(X = x) = 0$ olur, dört aralık olasılığı çakışır ve F sürekli bir fonksiyondur. Kesikli ile sürekli arasındaki ayrım, en kısa hâliyle, “tek bir noktaya kütle konabilir mi?” sorusunun yanıtıdır.

10.4 Örnekler

Önce en basit kesikli dağılımla, bir zarın dağılımıyla başlayalım; olasılık fonksiyonu ve dağılım fonksiyonu burada yan yana görülür.

Örnek 10.1 (Düzgün Bir Zarın Olasılık Fonksiyonu). Düzgün bir zar atılıyor ve gelen nokta sayısı X rastgele değişkeniyle gösteriliyor. X 'in olasılık fonksiyonunu ve dağılım fonksiyonunu bulunuz; ikisinin grafiğini çiziniz.

Çözüm.

Deneyin örnek uzayı $\Omega = \{1, 2, 3, 4, 5, 6\}$ ve zar düzgün olduğundan olasılık uzayı bir Laplace uzayıdır (Tanım 4.1): her ω için $P(\{\omega\}) = \frac{1}{6}$. $X(\omega) = \omega$ alınırsa değer kümesi

$$D_X = X(\Omega) = \{1, 2, 3, 4, 5, 6\}$$

sonludur; dolayısıyla X kesikli bir rastgele değişkendir. Her $x \in D_X$ için $(X = x) = \{x\}$ olduğundan olasılık fonksiyonu

$$f(x) = P(X = x) = P(\{x\}) = \frac{1}{6}, \quad x \in \{1, 2, 3, 4, 5, 6\}$$

olur. Tablo hâlinde:

x	1	2	3	4	5	6
$f(x) =$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
$P(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

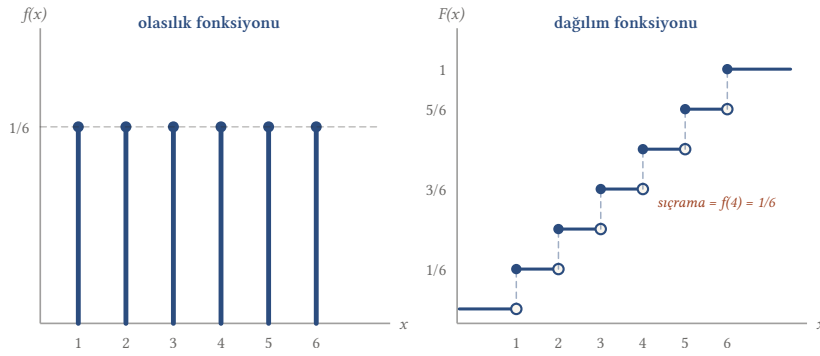
Bütün değerler negatif olmayan sayılardır ve toplamaları $6 \cdot \frac{1}{6} = 1$ 'dir; **Önerme 10.1**'nin iki koşulu sağlanır. Bütün değerlerin eşit olasılıklı olduğu bu dağılıma $\{1, \dots, 6\}$ üzerinde **kesikli düzgün dağılım** denir.

Dağılım fonksiyonu için **Önerme 10.2**'nin (4) maddesini kullanalım. $x < 1$ için hiçbir değer x 'i aşmaz, $F(x) = 0$. $k \leq x < k + 1$ ($k = 1, \dots, 5$) için x 'i aşmayan değerler $1, \dots, k$ 'dir ve $F(x) = \frac{k}{6}$ olur. $x \geq 6$ için bütün değerler toplanır ve $F(x) = 1$ 'dir:

$$F(x) = \begin{cases} 0, & x < 1 \\ \frac{1}{6}, & 1 \leq x < 2 \\ \frac{2}{6}, & 2 \leq x < 3 \\ \frac{3}{6}, & 3 \leq x < 4 \\ \frac{4}{6}, & 4 \leq x < 5 \\ \frac{5}{6}, & 5 \leq x < 6 \\ 1, & x \geq 6. \end{cases}$$

Kısaca, $1 \leq x < 6$ için $F(x) = \frac{\lfloor x \rfloor}{6}$ 'dir; burada $\lfloor x \rfloor$, x 'i aşmayan en büyük tam sayıdır. Örneğin $F(2.7) = P(X \leq 2.7) = P(X \in \{1, 2\}) = \frac{2}{6}$.

F 'nin sıçramalarından f 'yi geri okuyabiliriz: $f(3) = F(3) - F(3^-) = \frac{3}{6} - \frac{2}{6} = \frac{1}{6}$. Aralık olasılıkları da F 'den okunur: $P(2 < X \leq 5) = F(5) - F(2) = \frac{5}{6} - \frac{2}{6} = \frac{3}{6}$, gerçekten de bu olaya 3, 4, 5 değerleri düşer.



Şekil 10.3: Düzgün bir zarda gelen sayı X . Solda olasılık fonksiyonu: altı noktada eşit kütle $1/6$. Sağda dağılım fonksiyonu: her tam sayıda $1/6$ 'lık bir basamak. İki gösterim birbirini belirler — basamağın yüksekliği kütle, kütlelerin birikimi basamakları verir.

■

Şimdi değer kümesi doğrudan örnek uzay olmayan, değişkenin sonuçları “sıkıştırdığı” bir örneğe geçelim.

Örnek 10.2 (İşaretli Paranın Son İki Atışının Toplamı). Düzgün bir paranın bir yüzü 1, diğer yüzü 2 ile işaretlenmiştir. Bu para üç kez atılıyor ve X , son iki atışta gelen sayıların toplamını gösteriyor. X 'in olasılık fonksiyonunu ve dağılım fonksiyonunu bulunuz.

Çözüm.

Her atışın sonucu 1 ya da 2'dir; deneyin örnek uzayı

$$\Omega = \{1, 2\}^3 = \{(\omega_1, \omega_2, \omega_3) : \omega_i \in \{1, 2\}\}$$

kümesidir ve $2^3 = 8$ elemanı vardır. Para düzgün ve atışlar birbirinden bağımsız olduğundan her üçlü eşit olasılıklıdır: $P(\{\omega\}) = \frac{1}{8}$. Rastgele değişken

$$X(\omega_1, \omega_2, \omega_3) = \omega_2 + \omega_3$$

ile tanımlıdır; ilk atış X 'i etkilemez. $\omega_2 + \omega_3$ toplamı en az $1 + 1 = 2$, en çok $2 + 2 = 4$ olabildiğinden

$$D_X = \{2, 3, 4\}$$

bulunur; X kesiklidir. Her değer için olayı elemanlarıyla yazıp sayalım:

- $(X = 2) = \{\omega : \omega_2 = 1, \omega_3 = 1\} = \{(1, 1, 1), (2, 1, 1)\}$, iki eleman;
- $(X = 3) = \{\omega : \{\omega_2, \omega_3\} = \{1, 2\}\}$, yani $\{(1, 1, 2), (1, 2, 1), (2, 1, 2), (2, 2, 1)\}$, dört eleman;
- $(X = 4) = \{\omega : \omega_2 = 2, \omega_3 = 2\} = \{(1, 2, 2), (2, 2, 2)\}$, iki eleman.

Laplace formülü (Önerme 4.2) ile

$$f(2) = \frac{2}{8} = \frac{1}{4}, \quad f(3) = \frac{4}{8} = \frac{1}{2}, \quad f(4) = \frac{2}{8} = \frac{1}{4}.$$

x	2	3	4
$f(x) = P(X = x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Toplam $\frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1$ 'dir. Aynı sonuca bağımsızlıkla da varılır: ikinci atışın 1 gelmesi ile üçüncü atışın 1 gelmesi bağımsız olaylardır (Tanım 7.1) ve her birinin olasılığı $\frac{1}{2}$ 'dir; dolayısıyla $P(X = 2) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$. $X = 3$ için iki ayrık yol, $(\omega_2, \omega_3) = (1, 2)$ ve $(\omega_2, \omega_3) = (2, 1)$, vardır ve $P(X = 3) = 2 \cdot \frac{1}{4} = \frac{1}{2}$. Dağılım fonksiyonu, değerleri soldan sağa biriktirerek bulunur:

$$F(x) = \begin{cases} 0, & x < 2 \\ \frac{1}{4}, & 2 \leq x < 3 \\ \frac{3}{4}, & 3 \leq x < 4 \\ 1, & x \geq 4. \end{cases}$$

■

Sıradaki örnekte olayları tek tek listelemek yerine sayma tekniklerini kullanacağız.

Örnek 10.3 (Torbadan İadesiz Bilye Çekme). Bir torbada 4 sarı ve 2 kırmızı bilye vardır. Torbadan, çekilen bilye geri konmaksızın art arda 3 bilye çekiliyor ve X , çekilen kırmızı bilye sayısını gösteriyor. X 'in olasılık fonksiyonunu ve dağılım fonksiyonunu bulunuz.

Çözüm.

Bilyeler geri konmadığından üç çekilişin sonucu, torbadaki 6 bilyeden seçilmiş 3 elemanlı bir alt kümedir; çekiliş sırası X 'i etkilemediğinden örnek uzayı bu alt kümelerin kümesi olarak alabiliriz. Kombinasyon sayısı (Teorem 5.2) gereği

$$|\Omega| = \binom{6}{3} = \frac{6 \cdot 5 \cdot 4}{3 \cdot 2 \cdot 1} = 20$$

ve bilyeler rastgele çekildiğinden her alt küme eşit olasılıklıdır; olasılık uzayı Laplace uzayıdır.

Torbada yalnızca 2 kırmızı bilye olduğundan X en çok 2 olabilir; $X \geq 0$ da açıktır. Öte yandan $X = 0$ mümkündür, çünkü üç sarı bilye çekilebilir. Demek ki

$$D_X = \{0, 1, 2\}.$$

$X = x$ olayı, 2 kırmızıdan x tanesinin ve 4 sarıdan $3 - x$ tanesinin seçilmesi demektir. Kırmızılardan seçimi sarıların seçimini kısıtlamadığından çarpma kuralıyla (Önerme 5.2) bu olayın eleman sayısı $\binom{2}{x} \binom{4}{3-x}$ 'tir. Laplace formülü (Önerme 4.2) ile

$$f(x) = P(X = x) = \frac{\binom{2}{x} \binom{4}{3-x}}{\binom{6}{3}}, \quad x \in \{0, 1, 2\}.$$

Değerleri hesaplayalım:

$$f(0) = \frac{\binom{2}{0} \binom{4}{3}}{20} = \frac{1 \cdot 4}{20} = \frac{1}{5}, \quad f(1) = \frac{\binom{2}{1} \binom{4}{2}}{20} = \frac{2 \cdot 6}{20} = \frac{3}{5}, \quad f(2) = \frac{\binom{2}{2} \binom{4}{1}}{20} = \frac{1 \cdot 4}{20} = \frac{1}{5}.$$

x	0	1	2
$f(x) = P(X = x)$	$\frac{1}{5}$	$\frac{3}{5}$	$\frac{1}{5}$

Toplam $\frac{1}{5} + \frac{3}{5} + \frac{1}{5} = 1$ 'dir. Bu biçimdeki dağılıma **hipergeometrik dağılım** denir; genel hâlini ileride ayrıca inceleyeceğiz.

Sonucu koşullu olasılıkla (Tanım 6.1) da denetleyebiliriz. $X = 0$ olması için üç çekilişin de sarı gelmesi gerekir. İlk çekilişte sarı gelme olasılığı $\frac{4}{6}$; sarı çıktıysa torbada 3 sarı, 2 kırmızı kalır ve ikinci çekilişte sarı olasılığı $\frac{3}{5}$; iki sarı çıktıysa 2 sarı, 2 kırmızı kalır ve üçüncüde sarı olasılığı $\frac{2}{4}$. Çarpma teoremiyle (Teorem 6.3)

$$P(X = 0) = \frac{4}{6} \cdot \frac{3}{5} \cdot \frac{2}{4} = \frac{24}{120} = \frac{1}{5},$$

sayma yoluyla bulduğumuzla aynıdır.

Dağılım fonksiyonu:

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{5}, & 0 \leq x < 1 \\ \frac{4}{5}, & 1 \leq x < 2 \\ 1, & x \geq 2. \end{cases}$$

Örneğin “en çok bir kırmızı” olasılığı $P(X \leq 1) = F(1) = \frac{4}{5}$, “en az bir kırmızı” olasılığı $P(X \geq 1) = 1 - P(X = 0) = \frac{4}{5}$ ’tir.

■

Değer kümesi biraz daha kalabalık olan bir örnek: bir tablo yardımıyla saymak burada en güvenli yoldur.

Örnek 10.4 (İki Zarın Mutlak Farkı). Düzgün iki zar atılıyor ve gelen nokta sayıları arasındaki mutlak fark X ile gösteriliyor. X ’in olasılık fonksiyonunu ve dağılım fonksiyonunu bulunuz. $P(X \geq 2)$ ve $P(1 \leq X \leq 3)$ olasılıklarını hesaplayınız.

Çözüm.

Örnek uzay, sıralı ikililerden oluşur:

$$\Omega = \{(i, j) : i, j \in \{1, \dots, 6\}\}, \quad |\Omega| = 36,$$

ve zarlar düzgün olduğundan her ikili $\frac{1}{36}$ olasılıklıdır. Rastgele değişken $X(i, j) = |i - j|$ olarak tanımlıdır. Fark en az 0 (iki zar aynı), en çok $|6 - 1| = 5$ olabildiğinden

$$D_X = \{0, 1, 2, 3, 4, 5\}.$$

Her ikili için $|i - j|$ değerini bir tabloya yazalım; satırlar birinci zarı, sütunlar ikinci zarı gösterir.

$i \setminus j$	1	2	3	4	5	6
1	0	1	2	3	4	5
2	1	0	1	2	3	4
3	2	1	0	1	2	3
4	3	2	1	0	1	2
5	4	3	2	1	0	1
6	5	4	3	2	1	0

Tablo, köşegene göre simetriktir. Köşegen üzerinde 6 tane 0 vardır. $k \geq 1$ için k değeri, köşegenin k adım üstündeki ve k adım altındaki köşegenlerde yer alır; bu köşegenlerin her birinde $6 - k$ hücre bulunur. Öyleyse k değerini alan hücre sayısı $2(6 - k)$ ’dir:

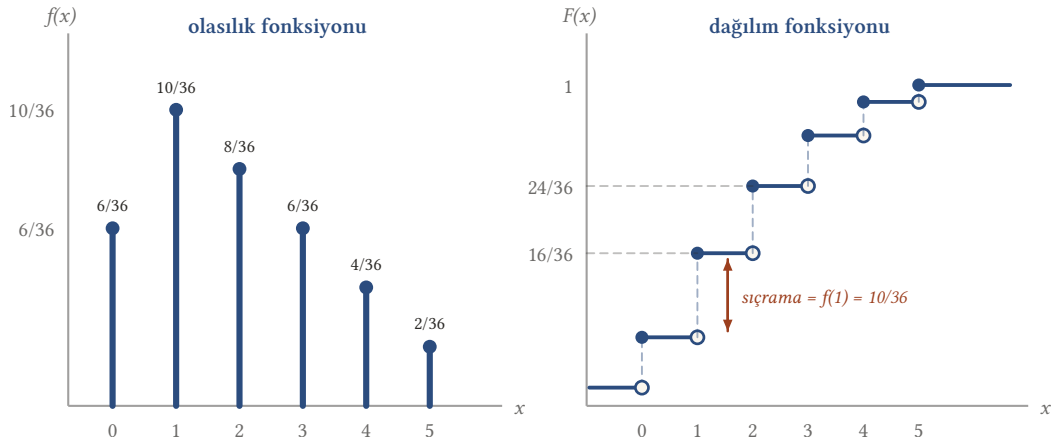
$$f(0) = \frac{6}{36}, \quad f(k) = \frac{2(6 - k)}{36} = \frac{12 - 2k}{36}, \quad k = 1, 2, 3, 4, 5.$$

x	0	1	2	3	4	5
$f(x) =$	$\frac{6}{36}$	$\frac{10}{36}$	$\frac{8}{36}$	$\frac{6}{36}$	$\frac{4}{36}$	$\frac{2}{36}$
$P(X = x)$						

Sadeleştirilmiş halleriyle $f(0) = \frac{1}{6}$, $f(1) = \frac{5}{18}$, $f(2) = \frac{2}{9}$, $f(3) = \frac{1}{6}$, $f(4) = \frac{1}{9}$, $f(5) = \frac{1}{18}$. Toplam $\frac{6+10+8+6+4+2}{36} = \frac{36}{36} = 1$ ’dir. En olası fark 1’dir, 0 değil; çünkü iki zarın eşit gelmesinin altı yolu varken farkın 1 olmasının on yolu vardır.

Dağılım fonksiyonu, olasılıkları biriktirerek bulunur; ara toplamalar 6, 16, 24, 30, 34, 36’dır:

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{6}{36}, & 0 \leq x < 1 \\ \frac{16}{36}, & 1 \leq x < 2 \\ \frac{24}{36}, & 2 \leq x < 3 \\ \frac{30}{36}, & 3 \leq x < 4 \\ \frac{34}{36}, & 4 \leq x < 5 \\ 1, & x \geq 5. \end{cases}$$



Şekil 10.4: Düzgün iki zarın gösterdiği sayıların mutlak farkı $X = |i - j|$. Solda olasılık fonksiyonu: kütleler eşit değildir, en olası fark 0 değil 1'dir ($10/36$). Sağda dağılım fonksiyonu: basamak yükseklikleri de eşit değildir; her basamak o noktanın kütlesi kadardır ve merdiven $6/36$ 'dan başlayıp $16/36$, $24/36$, $30/36$, $34/36$ düzeylerinden geçerek 1'e ulaşır.

İstenen olasılıklar:

$$P(X \geq 2) = 1 - P(X \leq 1) = 1 - F(1) = 1 - \frac{16}{36} = \frac{20}{36} = \frac{5}{9},$$

$$P(1 \leq X \leq 3) = f(1) + f(2) + f(3) = \frac{10 + 8 + 6}{36} = \frac{24}{36} = \frac{2}{3}.$$

İkincisi F ile de hesaplanabilir:

$$P(1 \leq X \leq 3) = F(3) - F(1^-) = F(3) - F(0) = \frac{30}{36} - \frac{6}{36} = \frac{24}{36}.$$

Burada sol uç 1 olaya dâhil olduğu için $F(1)$ değil $F(1^-) = F(0)$ çıkarılmıştır.

■

Buraya kadar hep f 'den F 'ye gittik. Şimdi ters yönü görelim: verilen bir merdiven grafiğinden olasılık fonksiyonunu okumak.

Örnek 10.5 (Dağılım Fonksiyonundan Olasılık Fonksiyonuna). X rastgele değişkeninin dağılım fonksiyonu

$$F(x) = \begin{cases} 0, & x < -1 \\ \frac{1}{4}, & -1 \leq x < 1 \\ \frac{1}{2}, & 1 \leq x < 3 \\ \frac{3}{4}, & 3 \leq x < 5 \\ 1, & x \geq 5 \end{cases}$$

olarak veriliyor. X 'in olasılık fonksiyonunu bulunuz; $P(X \leq 2)$, $P(X > 1)$ ve $P(-1 \leq X < 3)$ olasılıklarını hesaplayınız.

Çözüm.

F bir basamak fonksiyonudur: dört noktada sıçrar, aralarda sabittir. Önerme 10.2'nun (3) maddesine göre her x için $P(X = x) = F(x) - F(x^-)$ 'dir. F 'nin sabit olduğu bir aralığın iç noktalarında $F(x^-) = F(x)$ olduğundan bu noktaların olasılığı sıfırdır; olasılık yalnızca sıçrama noktalarında toplanır. Sıçrama noktaları $-1, 1, 3, 5$ 'tir.

Her sıçramanın büyüklüğünü hesaplayalım. F sağdan sürekli olduğundan $F(x)$ basamağın sol ucundaki (yeni) değer, $F(x^-)$ ise bir önceki basamağın değeridir:

$$P(X = -1) = F(-1) - F((-1)^-) = \frac{1}{4} - 0 = \frac{1}{4},$$

$$P(X = 1) = F(1) - F(1^-) = \frac{1}{2} - \frac{1}{4} = \frac{1}{4},$$

$$P(X = 3) = F(3) - F(3^-) = \frac{3}{4} - \frac{1}{2} = \frac{1}{4},$$

$$P(X = 5) = F(5) - F(5^-) = 1 - \frac{3}{4} = \frac{1}{4}.$$

Bu dört olasılığın toplamı 1'dir; demek ki bütün kütle bu dört noktadadır, başka hiçbir noktaya kütle kalmaz ve $D_X = \{-1, 1, 3, 5\}$ 'tir. Olasılık fonksiyonu:

x	-1	1	3	5
$f(x) = P(X = x)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

Yani X , dört noktası eşit olasılıklı bir kesikli düzgün dağılıma sahiptir. Sağlama olarak f 'den F 'yi yeniden kurabiliriz: $-1 \leq x < 1$ için $F(x) = f(-1) = \frac{1}{4}$, $1 \leq x < 3$ için $F(x) = f(-1) + f(1) = \frac{1}{2}$ ve böyle devam eder; verilen F elde edilir.

İstenen olasılıklar:

$$P(X \leq 2) = F(2) = \frac{1}{2} \quad (2 \text{ sayısı } [1, 3) \text{ basamağındadır}),$$

$$P(X > 1) = 1 - P(X \leq 1) = 1 - F(1) = 1 - \frac{1}{2} = \frac{1}{2},$$

$$P(-1 \leq X < 3) = F(3^-) - F((-1)^-) = \frac{1}{2} - 0 = \frac{1}{2}.$$

Sonuncusu, olaya düşen değerlerin olasılıklarını toplayarak da doğrulanır: $-1 \leq x < 3$ koşulunu D_X 'ten -1 ve 1 sağlar, $f(-1) + f(1) = \frac{1}{2}$.

■

Son örnek, tablo hâlinde verilen bir olasılık fonksiyonundan dağılım fonksiyonunu kurmayı ve ikisi arasındaki tutarlılığı denetlemeyi gösterir.

Örnek 10.6 (Olasılık Fonksiyonundan Dağılım Fonksiyonuna). X rastgele değişkeninin olasılık fonksiyonu aşağıdaki tabloyla veriliyor.

x	1	2	3
$f(x) = P(X = x)$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{6}$

X 'in dağılım fonksiyonunu bulunuz; $P(X \geq 2)$, $P(X < 3)$ ve $P(1 < X \leq 3)$ olasılıklarını hesaplayınız.

Çözüm.

Önce tablonun gerçekten bir olasılık fonksiyonu tanımladığını denetleyelim: değerler negatif değildir ve

$$\frac{1}{2} + \frac{1}{3} + \frac{1}{6} = \frac{3}{6} + \frac{2}{6} + \frac{1}{6} = 1.$$

“Bu iki özellik yeterlidir” notunda görüldüğü gibi, bu iki koşul tablonun kesikli bir dağılım tanımlaması için yeterlidir; $D_X = \{1, 2, 3\}$ 'tür.

Dağılım fonksiyonunu **Önerme 10.2**'nin (4) maddesiyle kuralım. $x < 1$ için $F(x) = 0$. $1 \leq x < 2$ için yalnızca 1 değeri x 'i aşmaz: $F(x) = f(1) = \frac{1}{2}$. $2 \leq x < 3$ için 1 ve 2 değerleri toplanır: $F(x) = \frac{1}{2} + \frac{1}{3} = \frac{5}{6}$. $x \geq 3$ için bütün değerler toplanır: $F(x) = 1$.

$$F(x) = \begin{cases} 0, & x < 1 \\ \frac{1}{2}, & 1 \leq x < 2 \\ \frac{5}{6}, & 2 \leq x < 3 \\ 1, & x \geq 3. \end{cases}$$

Sağlama: sıçramalar $F(1) - F(1^-) = \frac{1}{2}$, $F(2) - F(2^-) = \frac{5}{6} - \frac{1}{2} = \frac{1}{3}$, $F(3) - F(3^-) = 1 - \frac{5}{6} = \frac{1}{6}$; tablo geri gelir.

İstenen olasılıklar:

$$P(X \geq 2) = 1 - P(X \leq 1) = 1 - F(1) = \frac{1}{2} \quad (= f(2) + f(3)),$$

$$P(X < 3) = F(3^-) = \frac{5}{6} \quad (= f(1) + f(2)),$$

$$P(1 < X \leq 3) = F(3) - F(1) = 1 - \frac{1}{2} = \frac{1}{2} \quad (= f(2) + f(3)).$$

Son iki hesapta uç noktalara dikkat edilmelidir: $P(X < 3)$ için $F(3) = 1$ değil, $F(3^-) = \frac{5}{6}$ alınır; $P(1 < X \leq 3)$ ise 1 değerini dışarıda bıraktığı için $F(3) - F(1^-)$ 'dir, $F(3) - F(1^-) = 1$ değil.

■

10.5 Alıştırmalar

Alıştırma 10.1 (Kesikli Rastgele Değişkenler). (a) Düzgün iki para atılıyor ve X gelen tura sayısını gösteriyor. X 'in olasılık fonksiyonunu ve dağılım fonksiyonunu bulunuz.

(b) Bir yüzü 1, diğer yüzü 2 ile işaretlenmiş düzgün bir para üç kez atılıyor ve X gelen üç sayının toplamını gösteriyor. X 'in olasılık fonksiyonunu bulunuz ve $F(4, 5)$ değerini hesaplayınız.

(c) X rastgele değişkeninin olasılık fonksiyonu $f(x) = cx$, $x \in \{1, 2, 3, 4\}$ biçimindedir. c sabitini bulunuz; $P(X \geq 3)$ ve $F(2)$ değerlerini hesaplayınız.

(d) X rastgele değişkeninin olasılık fonksiyonu $f(x) = \left(\frac{1}{2}\right)^x$, $x \in \{1, 2, 3, \dots\}$ olsun. Bunun bir olasılık fonksiyonu olduğunu gösteriniz; dağılım fonksiyonunu bulunuz; $P(X > 3)$ ve $P(X \text{ çift})$ olasılıklarını hesaplayınız.

(e) **Örnek 10.3**'teki torbadan bu kez çekilen bilye her seferinde **geri konarak** üç bilye çekiliyor ve X çekilen kırmızı bilye sayısını gösteriyor. X 'in olasılık fonksiyonunu bulunuz ve iadesiz durumla karşılaştırınız.

(f) Düzgün iki zar atılıyor ve X gelen sayıların büyük olanını (ikisi eşitse ortak değeri) gösteriyor. Önce $F(k) = P(X \leq k)$ değerlerini $k \in \{1, \dots, 6\}$ için hesaplayınız, sonra buradan olasılık fonksiyonunu bulunuz.

(g) X rastgele değişkeninin dağılım fonksiyonu

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{8}, & 0 \leq x < 1 \\ \frac{1}{2}, & 1 \leq x < 2 \\ \frac{7}{8}, & 2 \leq x < 3 \\ 1, & x \geq 3 \end{cases}$$

ise X 'in olasılık fonksiyonunu bulunuz; $P(X \geq 1)$ ve $P(0 < X < 3)$ olasılıklarını hesaplayınız.

Çözüm.

(a) Örnek uzay $\Omega = \{YY, YT, TY, TT\}$ dört eşit olasılıklı sonuçtan oluşur ve $D_X = \{0, 1, 2\}$ 'dir. $(X = 0) = \{YY\}$, $(X = 1) = \{YT, TY\}$, $(X = 2) = \{TT\}$ olduğundan

x	0	1	2
$f(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Dağılım fonksiyonu: $x < 0$ için $F(x) = 0$; $0 \leq x < 1$ için $F(x) = \frac{1}{4}$; $1 \leq x < 2$ için $F(x) = \frac{3}{4}$; $x \geq 2$ için $F(x) = 1$.

(b) $\Omega = \{1, 2\}^3$ sekiz eşit olasılıklı üçlünden oluşur ve $X(\omega) = \omega_1 + \omega_2 + \omega_3$ 'tür. Toplam en az 3, en çok 6 olabilir: $D_X = \{3, 4, 5, 6\}$. Toplamın $3 + m$ olması için tam m atışta 2 gelmelidir; bunun $\binom{3}{m}$ yolu vardır. Öyleyse

$$f(3) = \frac{1}{8}, \quad f(4) = \frac{3}{8}, \quad f(5) = \frac{3}{8}, \quad f(6) = \frac{1}{8}.$$

Toplam $\frac{1+3+3+1}{8} = 1$ 'dir. $F(4, 5) = P(X \leq 4, 5) = f(3) + f(4) = \frac{1}{8} + \frac{3}{8} = \frac{1}{2}$.

(c) Toplam koşulu $c(1 + 2 + 3 + 4) = 10c = 1$ verir; $c = \frac{1}{10}$. $c > 0$ olduğundan negatif olmama koşulu da sağlanır. $P(X \geq 3) = f(3) + f(4) = \frac{3}{10} + \frac{4}{10} = \frac{7}{10}$ ve $F(2) = f(1) + f(2) = \frac{1}{10} + \frac{2}{10} = \frac{3}{10}$. Sağlama: $P(X \geq 3) = 1 - F(2) = \frac{7}{10}$.

(d) Değerler pozitifdir ve geometrik seri toplamı

$$\sum_{x=1}^{\infty} \left(\frac{1}{2}\right)^x = \frac{\frac{1}{2}}{1 - \frac{1}{2}} = 1$$

olduğundan f bir olasılık fonksiyonudur; $D_X = \mathbb{N}$ sayılabilir sonsuzdur. $x < 1$ için $F(x) = 0$. $n \leq x < n + 1$ ($n \in \mathbb{N}$) için x 'i aşmayan değerler $1, \dots, n$ 'dir ve

$$F(x) = \sum_{k=1}^n \left(\frac{1}{2}\right)^k = \frac{1}{2} \cdot \frac{1 - \left(\frac{1}{2}\right)^n}{1 - \frac{1}{2}} = 1 - \left(\frac{1}{2}\right)^n.$$

Kısaca, $x \geq 1$ için $F(x) = 1 - \left(\frac{1}{2}\right)^{\lfloor x \rfloor}$. Merdivenin sonsuz sayıda basamağı vardır ve yükseklikleri 1'e yaklaşır; $F(x) = 1$ olan hiçbir x yoktur ama $\lim_{x \rightarrow \infty} F(x) = 1$ 'dir. $P(X > 3) = 1 - F(3) = \left(\frac{1}{2}\right)^3 = \frac{1}{8}$. Çift değerler için

$$P(X \text{ çift}) = \sum_{k=1}^{\infty} \left(\frac{1}{2}\right)^{2k} = \sum_{k=1}^{\infty} \left(\frac{1}{4}\right)^k = \frac{\frac{1}{4}}{1 - \frac{1}{4}} = \frac{1}{3}.$$

(e) Bilye geri konduğundan her çekilişte torbada 4 sarı, 2 kırmızı vardır; her çekilişte kırmızı gelme olasılığı $\frac{2}{6} = \frac{1}{3}$, sarı gelme olasılığı $\frac{2}{3}$ 'tür ve çekilişler bağımsızdır. Artık üç kırmızı da mümkündür: $D_X = \{0, 1, 2, 3\}$. Tam x kırmızı içeren belirli bir sıralamanın olasılığı $\left(\frac{1}{3}\right)^x \left(\frac{2}{3}\right)^{3-x}$ ve bu sıralamaların sayısı $\binom{3}{x}$ olduğundan

$$f(x) = \binom{3}{x} \left(\frac{1}{3}\right)^x \left(\frac{2}{3}\right)^{3-x}, \quad x \in \{0, 1, 2, 3\}.$$

x	0	1	2	3
$f(x)$	$\frac{8}{27}$	$\frac{12}{27}$	$\frac{6}{27}$	$\frac{1}{27}$

Toplam $\frac{8+12+6+1}{27} = 1$ 'dir. İadesiz durumda değerler $\frac{1}{5}, \frac{3}{5}, \frac{1}{5}$ idi ve $X = 3$ imkânsızdı. İadeli çekilişte kütle daha çok yayılır: $X = 0$ olasılığı $\frac{1}{5} = 0,20$ 'den $\frac{8}{27} \approx 0,30$ 'a çıkar, $X = 1$ olasılığı $0,60$ 'tan $\approx 0,44$ 'e düşer. Bu dağılıma ileride **binom dağılımı** adını vereceğiz.

(f) $(X \leq k)$ olayı, iki zarın da k 'yi aşmaması demektir. Bu olaya düşen ikililer $\{1, \dots, k\}^2$ kümesidir, k^2 tane; öyleyse

$$F(k) = P(X \leq k) = \frac{k^2}{36}, \quad k \in \{1, \dots, 6\},$$

$x < 1$ için $F(x) = 0$ ve $x \geq 6$ için $F(x) = 1$. Olasılık fonksiyonu sıçramalardan okunur:

$$f(k) = F(k) - F(k-1) = \frac{k^2 - (k-1)^2}{36} = \frac{2k-1}{36}, \quad k \in \{1, \dots, 6\}.$$

x	1	2	3	4	5	6
$f(x)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

Toplam $\frac{1+3+5+7+9+11}{36} = 1$ 'dir. Bu alıştırma, bazı sorularda F 'yi doğrudan hesaplayıp f 'ye sonradan geçmenin daha kolay olduğunu gösterir.

(g) Sıçrama noktaları 0, 1, 2, 3'tür ve sıçramalar

$$f(0) = \frac{1}{8} - 0 = \frac{1}{8}, \quad f(1) = \frac{1}{2} - \frac{1}{8} = \frac{3}{8}, \quad f(2) = \frac{7}{8} - \frac{1}{2} = \frac{3}{8}, \quad f(3) = 1 - \frac{7}{8} = \frac{1}{8}.$$

Toplam 1 olduğundan $D_X = \{0, 1, 2, 3\}$ 'tür; bu, düzgün üç para atışında gelen tura sayısının dağılımıdır. $P(X \geq 1) = 1 - F(0) = 1 - \frac{1}{8} = \frac{7}{8}$ ve $P(0 < X < 3) = f(1) + f(2) = \frac{3}{8} + \frac{3}{8} = \frac{3}{4}$; ikincisi F ile $F(3^-) - F(0) = \frac{7}{8} - \frac{1}{8} = \frac{3}{4}$ olarak da bulunur.

■

Kesikli değişkenlerde olasılık kütlesi sayılabilir noktalara yığılmıştı ve her hesap bir toplama indirgeniyordu. Olasılığın bir aralık boyunca sürekli biçimde yayıldığı durumda toplamaların yerini integraller, olasılık fonksiyonunun yerini yoğunluk fonksiyonu alır: [Sürekli Rastgele Değişkenler ve Yoğunluk Fonksiyonu](#).

11. Sürekli Rastgele Değişkenler ve Yoğunluk Fonksiyonu

Kesikli Rastgele Değişkenler ve Olasılık Fonksiyonu bölümünde değer kümesi sonlu ya da sayılabilir sonsuz olan rastgele değişkenleri inceledik. Böyle bir değişkenin bütün olasılık bilgisi, $f_X(x) = P(X = x)$ olasılık fonksiyonunda toplanıyordu: dağılım fonksiyonu bu değerlerin toplamıyla elde ediliyor, olaylar da ilgili noktaların olasılıkları toplanarak hesaplanıyordu. Ne var ki uygulamada karşılaştığımız pek çok rastgele değişken bu çerçeveye sığmaz. Bir elektronik parçanın ömrü, bir yolcunun durağa geldiği an, rastgele seçilen bir noktanın koordinatı ya da bir ölçümün hatası bir aralıktaki **her** gerçel değeri alabilir; değer kümesi sayılamazdır.

Bu bölümde işte bu tür değişkenleri ele alıyoruz. Toplamların yerini integraller, olasılık fonksiyonunun yerini **olasılık yoğunluk fonksiyonu** alacak. Yoğunluk fonksiyonu f , tek bir noktanın olasılığını değil, “birim uzunluğa düşen olasılığı” ölçer: X ’in $[a, b]$ aralığına düşme olasılığı, f ’nin grafiği altında a ile b arasında kalan alandır. Bu bakış açısının ilk şaşırtıcı sonucu, sürekli bir rastgele değişkenin herhangi bir sabit değeri alma olasılığının sıfır olmasıdır; ikincisi de yoğunluğun kendisinin bir olasılık olmamasıdır, hatta 1’i aşabilmesidir.

Bölüm boyunca şunları öğreneceğiz: sürekli ve mutlak sürekli rastgele değişkenin tanımı; yoğunluk fonksiyonunun iki belirleyici özelliği; dağılım fonksiyonunun sürekli olduğu ve nokta olasılıklarının sıfıra eşit olduğu; aralık olasılıklarının integralle hesabı; dağılım fonksiyonu verildiğinde yoğunluğun türevle bulunması. Sonra bu araçları sinüs, parçalı sabit, üçgen ve üstel biçimli yoğunluklar üzerinde sonuna kadar çalıştıracacağız.

11.1 Sürekli Rastgele Değişken ve Yoğunluk Fonksiyonu

Bir rastgele değişkenin (Tanım 8.1) değer kümesi $D_X = X(\Omega)$ idi (Tanım 8.2). Kesikli değişken için bu küme en çok sayılabilir sonsuzdu (Tanım 10.1). Bunun karşıtıyla başlıyoruz.

Tanım 11.1 (Sürekli Rastgele Değişken). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı ve $X : \Omega \rightarrow \mathbb{R}$ bir rastgele değişken olsun. X ’in değer kümesi D_X **sayılamaz** bir küme ise X ’e **sürekli rastgele değişken** denir.

Bu tanım yalnızca değer kümesinin büyüklüğüne bakar; olasılıkların bu küme üzerinde nasıl dağıldığı hakkında bir şey söylemez. Kesikli durumda dağılımı $P(X = x)$ sayıları belirliyordu. Sayılamaz bir değer kümesinde bu yol kapanır: birazdan göreceğimiz gibi, yoğunluğu olan bir değişkende $P(X = x)$ sayılarının hepsi sıfırdır ve toplanacak bir şey kalmaz. Onun yerine olasılığı “yayılmış” biçimde, bir yoğunlukla anlatırız.

Tanım 11.2 (Mutlak Sürekli Rastgele Değişken ve Olasılık Yoğunluk Fonksiyonu). X bir rastgele değişken ve $F : \mathbb{R} \rightarrow [0, 1]$ onun dağılım fonksiyonu olsun. Eğer

$$(i) f(x) \geq 0 (x \in \mathbb{R}), \quad (ii) \int_{-\infty}^{\infty} f(x) dx = 1$$

özelliklerini taşıyan bir $f : \mathbb{R} \rightarrow \mathbb{R}$ fonksiyonu için her $x \in \mathbb{R}$ ’de

$$F(x) = \int_{-\infty}^x f(t) dt$$

yazılabiliyorsa X ’e **mutlak sürekli rastgele değişken**, f ’ye de X ’in **olasılık yoğunluk fonksiyonu** (kısaca **yoğunluk fonksiyonu** ya da **yoğunluğu**) denir. X ’in yoğunluğu f_X ile de gösterilir.

İki noktaya dikkat edelim. Birincisi, integral değişkeni olarak t kullandık; üst sınır x olduğundan integralin içinde de x yazmak yanıltıcı olur. İkincisi, mutlak sürekli bir değişken zorunlu olarak sürekli: F ’nin

bir integral olması onu sürekli kılar (bunu az sonra ispatlayacağız), sürekli bir F ise sayılabilir bir değer kümesiyle uyumsuz, çünkü kesikli bir değişkende $\sum_{x \in D_X} f_X(x) = 1$ olduğundan $f_X(x) > 0$ olan en az bir $x \in D_X$ vardır ve dağılım fonksiyonu bu noktada $f_X(x)$ kadar sıçrar (Önerme 10.2). Tersisi doğru değildir: dağılım fonksiyonu sürekli olduğu hâlde hiçbir yoğunluğa sahip olmayan rastgele değişkenler vardır; ama bunlar bu kitabın dışında kalan, yapay örneklerdir. **Bu kitapta “sürekli rastgele değişken” dendiğinde her zaman yoğunluğu olan, yani mutlak sürekli bir değişken kastedilir.**

Yoğunluğun anlamını sezmek için F 'nin tanımını küçük bir aralıkta okuyalım. Kısa bir $[x, x + h]$ aralığı için

$$P(x < X \leq x + h) = F(x + h) - F(x) = \int_x^{x+h} f(t) dt \approx f(x) h$$

olur. Yani $f(x)$, x civarındaki birim uzunluğa düşen olasılıktır; f büyükse X 'in oraya düşme ihtimali büyüktür. Bu yüzden adı “yoğunluk”tur. Aynı okuma bir uyarıyı da beraberinde getirir.

⚠ Yoğunluk bir olasılık değildir

Kesikli durumda $f_X(x) = P(X = x)$ bir olasılıktı; dolayısıyla $0 \leq f_X(x) \leq 1$ idi. Sürekli durumda $f(x)$ **bir olasılık değildir** ve 1'i aşabilir. Örneğin

$$f(x) = \begin{cases} 3x^2, & 0 < x < 1 \\ 0, & \text{diğer durumlarda} \end{cases}$$

fonksiyonu $f \geq 0$ ve $\int_{-\infty}^{\infty} f(x) dx = \int_0^1 3x^2 dx = 1$ koşullarını sağlar, dolayısıyla bir yoğunluk fonksiyonudur; ama $f(0,9) = 2,43 > 1$ 'dir. Hatta yoğunluk sınırsız bile olabilir: $(0, 1]$ üzerinde $\frac{1}{2\sqrt{x}}$ ile verilen fonksiyon 0'a yaklaşırken sonsuza gider, yine de bir yoğunluktur (bkz. Alıştırma 11.1). Olasılık olan şey f 'nin değeri değil, f 'nin **grafiki altında kalan alandır**: her aralık üzerindeki alan 0 ile 1 arasındadır, toplam alan tam 1'dir. Yukarıdaki $3x^2$ yoğunluğu aşağıda ayrıntılı olarak incelenecektir (Örnek 11.1).

11.2 Dağılım Fonksiyonunun Sürekliliği ve Nokta Olasılıkları

Dağılım fonksiyonunun genel özelliklerini biliyoruz (Teorem 9.1): artan (azalmayan), sağdan sürekli, $-\infty$ 'da 0'a ve $+\infty$ 'da 1'e giden bir fonksiyon. Kesikli değişkende F basamaklıydı ve basamakların yüksekliği nokta olasılıklarını veriyordu. Yoğunluğu olan bir değişkende ise F bir integraldir ve integraller “sıçramaz”. Bu gözlem, sürekli değişkenlerin bütün hesap kurallarının kaynağıdır; o yüzden onu dikkatle ispatlayalım.

Teorem 11.1 (Sürekli Rastgele Değişkenin Dağılım Fonksiyonu Süreklidir). X , yoğunluğu f olan sürekli bir rastgele değişken olsun. O zaman X 'in dağılım fonksiyonu

$$F(x) = \int_{-\infty}^x f(t) dt$$

$\mathbb{R}$ 'nin her noktasında **süreklidir**. Aynı sonuç, $f \geq 0$ ve $\int_{-\infty}^{\infty} f = 1$ koşullarını sağlayan **her** f için, bu integralle tanımlanan F hakkında geçerlidir.

İspat.

$a \in \mathbb{R}$ sabit bir nokta ve $h > 0$ olsun. İntegralin aralık üzerinde toplanabilirliğinden

$$F(a) - F(a - h) = \int_{a-h}^a f(t) dt, \quad F(a + h) - F(a) = \int_a^{a+h} f(t) dt$$

yazılır ve $f \geq 0$ olduğundan iki fark da negatif değildir. Göstermemiz gereken, $h \rightarrow 0^+$ iken bu iki integralin de 0'a gitmesidir; o zaman $F(a^-) = F(a) = F(a^+)$ olur ve F , a 'da sürekli çıkar. İki farkın toplamı $\int_{a-h}^{a+h} f(t) dt$ olduğundan bu toplamın 0'a gittiğini göstermek yeter.

Durum 1: f , a 'nın bir komşuluğunda sınırlıdır. Öyle bir $\delta > 0$ ve $M > 0$ olsun ki $|t - a| \leq \delta$ için $0 \leq f(t) \leq M$ olsun. O zaman $0 < h \leq \delta$ için

$$0 \leq \int_{a-h}^{a+h} f(t) dt \leq \int_{a-h}^{a+h} M dt = 2Mh$$

olur ve sağ taraf $h \rightarrow 0^+$ iken 0'a gider. Sıkıştırma ile integral de 0'a gider.

Durum 2: f , a 'nın hiçbir komşuluğunda sınırlı değildir. Bu durumda f 'nin $[a-1, a+1]$ üzerindeki integrali a noktasında bir has olmayan (improper) integraldir; $\int_{-\infty}^{\infty} f$ yakınsak olduğundan bu integral de yakınsaktır. Has olmayan integralin yakınsak olması, tanım gereği,

$$\lim_{h \rightarrow 0^+} \int_{a-1}^{a-h} f(t) dt = \int_{a-1}^a f(t) dt, \quad \lim_{h \rightarrow 0^+} \int_{a+h}^{a+1} f(t) dt = \int_a^{a+1} f(t) dt$$

demektir. Birinci eşitlik $\int_{a-h}^a f(t) dt \rightarrow 0$, ikincisi $\int_a^{a+h} f(t) dt \rightarrow 0$ anlamına gelir; toplamı da 0'a gider.

Her iki durumda da $F(a-h) \rightarrow F(a)$ ve $F(a+h) \rightarrow F(a)$ bulundu; F , a 'da süreklidir. a keyfi olduğundan F , $\mathbb{R}$ üzerinde süreklidir. İspatta f hakkında yalnızca $f \geq 0$ ve integralin yakınsaklığı kullanıldı; bu, son cümledeki genel ifadeyi de verir.

■

Bu teoremin ilk ve en önemli sonucu, kesikli durumla en keskin ayrılığı gösterir: sürekli bir değişkende hiçbir tek nokta pozitif olasılık taşımaz. Bunun pratik yansıması, aralık uçlarının olasılık hesabında hiç önem taşımamasıdır.

Önerme 11.1 (Nokta Olasılığı Sıfırdır; Aralık Olasılıkları). X , yoğunluğu f ve dağılım fonksiyonu F olan sürekli bir rastgele değişken olsun. Her $a \in \mathbb{R}$ için

$$P(X = a) = 0$$

dır. Bunun sonucu olarak, her $a \in \mathbb{R}$ için

$$P(X \leq a) = P(X < a) = F(a) = \int_{-\infty}^a f(x) dx, \quad P(X \geq a) = P(X > a) = 1 - F(a) = \int_a^{\infty} f(x) dx$$

ve her $a < b$ için

$$P(a < X \leq b) = P(a < X < b) = P(a \leq X \leq b) = P(a \leq X < b) = F(b) - F(a) = \int_a^b f(x) dx$$

olur.

İspat.

Dağılım fonksiyonu ile nokta olasılığı arasındaki bağıntı (Teorem 9.2)

$$P(X = a) = F(a) - F(a^-), \quad F(a^-) = \lim_{h \rightarrow 0^+} F(a - h)$$

biçimindedir. Teorem 11.1 gereği F , a 'da süreklidir; dolayısıyla $F(a^-) = F(a)$ ve $P(X = a) = F(a) - F(a) = 0$ olur.

Aralık formülleri için $(X \leq a) = (X < a) \cup (X = a)$ ayrık birleşimini yazalım. Toplama kuralıyla

$$P(X \leq a) = P(X < a) + P(X = a) = P(X < a)$$

bulunur; $P(X \leq a) = F(a)$ ise dağılım fonksiyonunun tanımıdır (Tanım 9.1) ve $F(a) = \int_{-\infty}^a f$, yoğunluğun tanımıdır. Tümleyene geçerek $P(X > a) = 1 - F(a)$ ve $P(X \geq a) = P(X > a) + P(X = a) = 1 - F(a)$ elde edilir; $1 - F(a) = \int_{-\infty}^{\infty} f - \int_{-\infty}^a f = \int_a^{\infty} f$ olur.

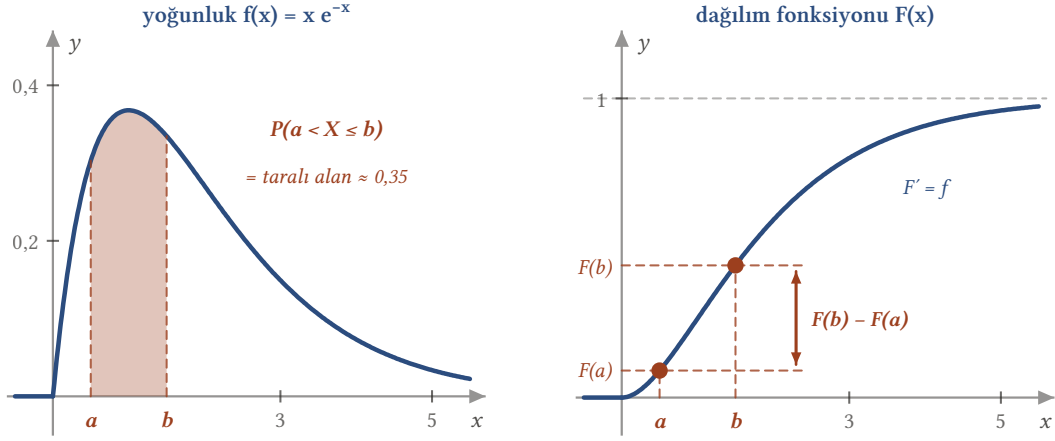
$a < b$ olsun. $P(a < X \leq b) = F(b) - F(a)$ eşitliği Teorem 9.2'den bilinmektedir. Öteki üç olay bu olaydan yalnızca a ya da b uç noktasının eklenip çıkarılmasıyla elde edilir:

$$(a \leq X \leq b) = (a < X \leq b) \cup (X = a), \quad (a < X < b) = (a < X \leq b) \setminus (X = b), \quad (a \leq X < b) = (a < X < b) \cup (X = a).$$

Eklenen ve çıkarılan olayların olasılığı 0 olduğundan dördünün olasılığı da $F(b) - F(a)$ 'dır. Son olarak

$$F(b) - F(a) = \int_{-\infty}^b f(x) dx - \int_{-\infty}^a f(x) dx = \int_a^b f(x) dx$$

integralin toplanabilirliğinden gelir.

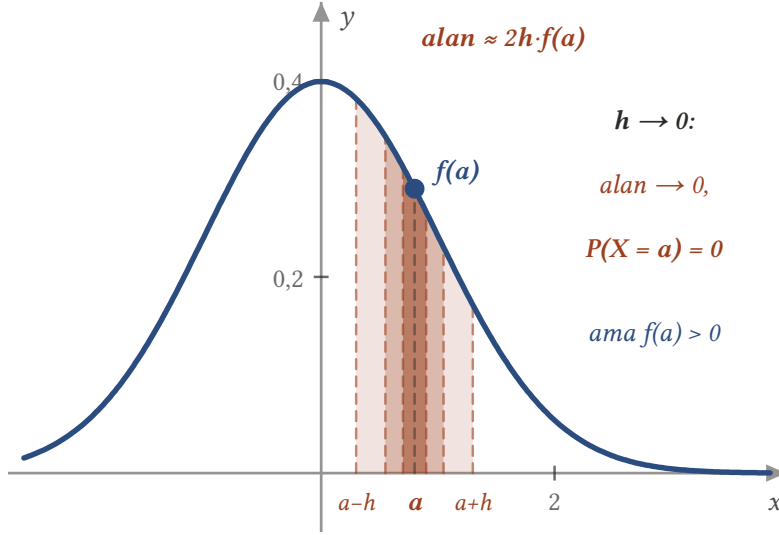


Şekil 11.1: Aralık olasılığının iki görünümü. Solda $P(a < X \leq b)$, yoğunluk eğrisinin altında a ile b arasında kalan alandır; sağda aynı sayı dağılım fonksiyonunun iki değeri arasındaki fark $F(b) - F(a)$ olarak okunur. Sürekli değişkende uç noktalar sıfır olasılık taşıdığından aralığın açık ya da kapalı olması sonucu değiştirmez; dört aralık olasılığı da bu alana eşittir.

■

i Sıfır olasılık ile imkânsızlık aynı şey değildir

X her deneyde bir değer alır; alacağı değer a olsun. $P(X = a) = 0$ olması “ $X = a$ olayı gerçekleşemez” demek değildir; sadece, sayılamaz çoklukta aday değer arasında **tam olarak** a ’nın seçilmesine olasılık ölçüsünün sıfır ağırlık verdiği anlamına gelir. Geometrik olasılıkta (Tanım 4.5) bir noktanın alanının sıfır olmasıyla aynı durumdur. Buna karşılık a ’nın çevresindeki her aralık pozitif olasılık taşıyabilir: $P(a - h < X < a + h) = \int_{a-h}^{a+h} f > 0$ olabilir, ama $h \rightarrow 0$ iken bu sayı 0’a çekilir.



Şekil 11.2: Sürekli bir değişkende tek noktanın olasılığı sıfırdır. a çevresindeki $[a - h, a + h]$ aralığının olasılığı eğri altındaki taralı alandır, yaklaşık $2h \cdot f(a)$; h küçüldükçe (0,5; 0,25; 0,1) bu alan sıfıra çekilir ve limitte $P(X = a) = 0$ kalır. Oysa $f(a)$ sıfır değildir: yoğunluk bir olasılık değil, birim uzunluğa düşen olasılıktır.

Kesikli ve sürekli değişkenlerin karşılaştırması şöyle özetlenebilir:

	kesikli X	sürekli X
dağılımı belirleyen fonksiyon	olasılık fonksiyonu $f_X(x) = P(X = x)$	yoğunluk fonksiyonu f_X , $f_X(x)$ bir olasılık değil
normalleştirme	$\sum_{x \in D_X} f_X(x) = 1$	$\int_{-\infty}^{\infty} f_X(x) dx = 1$
dağılım fonksiyonu	$F(x) = \sum_{x_i \leq x} f_X(x_i)$, basamaklı	$F(x) = \int_{-\infty}^x f_X(t) dt$, sürekli
nokta olasılığı	$P(X = a) = F(a) - F(a^-)$, D_X 'te pozitif	$P(X = a) = 0$
aralık olasılığı	$P(a < X \leq b) = \sum_{a < x_i \leq b} f_X(x_i)$	$P(a < X \leq b) = \int_a^b f_X(x) dx$, uçlar önemsiz
F 'den f_X 'e dönüş	$f_X(x) = F(x) - F(x^-)$	$f_X(x) = F'(x)$ (türevin var olduğu yerde)

Tablonun son satırı bu bölümün ikinci ana sonucudur ve birazdan ispatlanacaktır.

Nokta olasılıklarının sıfır olmasının nasıl kullanıldığını ilk tam örnekte görelim. Örnek aynı zamanda yukarıdaki uyarıyı da somutlaştırır.

Örnek 11.1 (Bir Yoğunlukla Olasılık Hesapları). X rastgele değişkeninin yoğunluk fonksiyonu

$$f(x) = \begin{cases} 3x^2, & 0 < x < 1 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun.

- a) f 'nin bir yoğunluk fonksiyonu olduğunu doğrulayınız ve $f(x) > 1$ olan x değerlerini bulunuz.
b) X 'in dağılım fonksiyonunu bulunuz.
c) $(PX > \frac{1}{2})$, $(\frac{1}{4} < X \leq \frac{3}{4})$, $(PX = \frac{1}{2})$ ve $(PX > \frac{1}{2} \mid X > \frac{1}{4})$ olasılıklarını hesaplayınız.
d) $P(X \leq m) = \frac{1}{2}$ olan m sayısını (dağılımın **medyanını**) bulunuz.

Çözüm.

a) Her x için $f(x) \geq 0$ 'dır ve

$$\int_{-\infty}^{\infty} f(x) dx = \int_0^1 3x^2 dx = [x^3]_0^1 = 1$$

olur; f bir yoğunluk fonksiyonudur. $0 < x < 1$ için $3x^2 > 1 \Leftrightarrow x^2 > \frac{1}{3} \Leftrightarrow x > \frac{1}{\sqrt{3}} \approx 0,577$ 'dir. Demek ki $(\frac{1}{\sqrt{3}}, 1)$ aralığında yoğunluk 1'den büyüktür; $f(1^-) = 3$ 'e kadar çıkar. Bu, f 'nin olasılık olmadığını bir kez daha gösterir.

b) $x \leq 0$ için $F(x) = \int_{-\infty}^x 0 dt = 0$. $0 < x < 1$ için

$$F(x) = \int_{-\infty}^0 0 dt + \int_0^x 3t^2 dt = [t^3]_0^x = x^3.$$

$x \geq 1$ için $F(x) = \int_0^1 3t^2 dt + \int_1^x 0 dt = 1$. Böylece

$$F(x) = \begin{cases} 0, & x \leq 0 \\ x^3, & 0 < x < 1 \\ 1, & x \geq 1. \end{cases}$$

F 'nin 0 ve 1 noktalarında sürekli olduğuna dikkat edelim: $F(0) = 0 = 0^3$ ve $F(1) = 1 = 1^3$.

c) **Önerme 11.1**'i kullanıyoruz:

$$(PX > \frac{1}{2}) = 1 - (F_{\frac{1}{2}}) = 1 - \frac{1}{8} = \frac{7}{8},$$

$$(\frac{1}{4} < X \leq \frac{3}{4}) = (F_{\frac{3}{4}}) - (F_{\frac{1}{4}}) = \frac{27}{64} - \frac{1}{64} = \frac{26}{64} = \frac{13}{32}.$$

$(PX = \frac{1}{2}) = 0$ 'dır; bunu aynı formülle de görebiliriz: $(\frac{1}{2} \leq X \leq \frac{1}{2}) = \int_{1/2}^{1/2} 3x^2 dx = 0$. Koşullu olasılık için (**Tanım 6.1**) $(X > \frac{1}{2}) \subseteq (X > \frac{1}{4})$ olduğundan

$$(PX > \frac{1}{2} \mid X > \frac{1}{4}) = \frac{(PX > \frac{1}{2})}{(PX > \frac{1}{4})} = \frac{\frac{7}{8}}{1 - \frac{1}{64}} = \frac{7}{8} \cdot \frac{64}{63} = \frac{8}{9}.$$

d) F , $(0, 1)$ üzerinde kesin artan ve sürekli olduğundan $F(m) = \frac{1}{2}$ denkleminin bu aralıkta tek bir çözümü vardır: $m^3 = \frac{1}{2}$, yani $m = 2^{-1/3} = \frac{1}{\sqrt[3]{2}} \approx 0,794$. X bu değer altına ve üstüne eşit olasılıkla düşer.

■

11.3 Hangi Fonksiyonlar Yoğunluk Olabilir?

Örneklerde sık sık “ $f(x) = c g(x)$ bir yoğunluk olacak biçimde c 'yi bulunuz” sorusuyla karşılaşacağız. Böyle bir soruda yaptığımız şey $\int f = 1$ denklemini çözmektir; ama bu denklem çözüldüğünde gerçekten bir rastgele değişkenin var olduğundan emin miyiz? Yani tanımdaki (i) ve (ii) koşulları, bir fonksiyonun yoğunluk olması için yalnızca gerekli mi, yoksa yeterli mi? Aşağıdaki önerme yeterli olduğunu söyler ve c bulma alışkanlığımızı sağlam bir zemine oturtur.

Önerme 11.2 (Yoğunluk Fonksiyonunun Belirleyici Özellikleri). $f : \mathbb{R} \rightarrow \mathbb{R}$ bir fonksiyon olsun. f 'nin bir rastgele değişkenin yoğunluk fonksiyonu olması için gerek ve yeter koşul

$$(i) f(x) \geq 0 (x \in \mathbb{R}) \quad \text{ve} \quad (ii) \int_{-\infty}^{\infty} f(x) dx = 1$$

olmasıdır. Bu durumda $F(x) = \int_{-\infty}^x f(t) dt$ fonksiyonu azalmayan, sürekli, $\lim_{x \rightarrow -\infty} F(x) = 0$ ve $\lim_{x \rightarrow \infty} F(x) = 1$ olan bir dağılım fonksiyonudur ve yoğunluğu f olan bir rastgele değişkenin dağılım fonksiyonudur.

İspat.

Gerekliklik tanımın kendisidir: yoğunluk, tanım gereği (i) ve (ii)'yi sağlar.

Yeterlilik için (i) ve (ii)'yi sağlayan bir f alalım ve $F(x) = \int_{-\infty}^x f(t) dt$ koyalım. F 'nin dağılım fonksiyonu özelliklerini (Teorem 9.1) taşıdığını gösterelim.

Azalmayan: $x < y$ için $F(y) - F(x) = \int_x^y f(t) dt \geq 0$, çünkü $f \geq 0$.

Sınırlar: $\int_{-\infty}^{\infty} f = 1$ integralinin yakınsak olması, tanım gereği, $\lim_{x \rightarrow \infty} \int_{-\infty}^x f(t) dt = 1$ demektir; yani $F(x) \rightarrow 1$. Aynı yakınsaklık $\lim_{x \rightarrow -\infty} \int_{-\infty}^x f(t) dt = 0$ verir; yani $F(x) \rightarrow 0$. Ayrıca $f \geq 0$ olduğundan her x için $0 \leq F(x) \leq \int_{-\infty}^{\infty} f = 1$ 'dir.

Sağdan süreklilik: Teorem 11.1'in ispatı yalnızca (i) ve (ii)'yi kullandı; dolayısıyla F her noktada, özellikle sağdan, sürekli.

Böylece F bir dağılım fonksiyonudur. Borel–Stieltjes teoremine (Teorem 4.2) göre $\mathcal{B}(\mathbb{R})$ üzerinde her x için $P_F((-\infty, x]) = F(x)$ olan tek bir olasılık ölçüsü P_F vardır. $\Omega = \mathbb{R}$, $\mathcal{U} = \mathcal{B}(\mathbb{R})$, $P = P_F$ olasılık uzayında $X(\omega) = \omega$ birim fonksiyonunu alalım. X bir rastgele değişkendir, çünkü her $B \in \mathcal{B}(\mathbb{R})$ için $X^{-1}(B) = B \in \mathcal{U}$ 'dur; dağılım fonksiyonu da

$$F_X(x) = P(X \leq x) = P_F((-\infty, x]) = F(x) = \int_{-\infty}^x f(t) dt$$

olur. Tanım gereği f , X 'in yoğunluk fonksiyonudur.

■

Önermenin kullanım biçimi şudur: negatif olmayan bir g için $\int_{-\infty}^{\infty} g(x) dx = I$ sonlu ve pozitifse $f = \frac{1}{I} g$ bir yoğunluktur; “ c 'yi bul” sorularının yanıtı hep $c = \frac{1}{I}$ 'dir. $\int g = \infty$ ise hiçbir sabit işe yaramaz; örneğin $g(x) = \frac{1}{x}$ ($x \geq 1$) fonksiyonundan bir yoğunluk elde edilemez.

11.4 Yoğunluğun Dağılım Fonksiyonundan Elde Edilmesi

Tanım, yoğunluktan dağılım fonksiyonuna geçmenin yolunu verdi: integral al. Ters yönde, dağılım fonksiyonu verildiğinde yoğunluğu nasıl buluruz? Kesikli durumda yanıt sıçrama yükseklikleriydi: $f_X(x) = F(x) - F(x^-)$ (Önerme 10.2). Sürekli durumda sıçrama yoktur; onun yerini **türev** alır. Bu, analizin temel teoreminin olasılık dilindeki karşılığıdır: “değişken üst sınırlı integralin türevi, integrandın kendisidir.” Teoremi kitap boyunca karşılaşacağımız bütün yoğunlukları kapsayan bir varsayım altında ispatlayacağız: yoğunluğun **her sınırlı aralıkta sonlu sayıda nokta dışında sürekli** olması. Parçaları sürekli formüllerle verilen her yoğunluk, parça uçları dışında sürekli ve bu varsayımı sağlar. Herhangi bir yoğunluk için teorem yine doğrudur, ama ispatı Lebesgue integrali gerektirir.

Teorem 11.2 (Yoğunluk Dağılım Fonksiyonunun Türevidir). X , dağılım fonksiyonu F olan sürekli bir rastgele değişken ve A , F 'nin türevlenebildiği noktaların kümesi olsun. O zaman

$$g(x) = \begin{cases} F'(x), & x \in A \\ 0, & x \notin A \end{cases}$$

funksiyonu X 'in bir yoğunluk fonksiyonudur; yani $g \geq 0$, $\int_{-\infty}^{\infty} g = 1$ ve her x için $F(x) = \int_{-\infty}^x g(t) dt$ 'dir. Özel olarak X 'in herhangi bir f yoğunluğu x_0 noktasında sürekli ise $x_0 \in A$ ve $F'(x_0) = f(x_0)$ 'dir.

İspat.

f , X 'in bir yoğunluğu olsun; varsayımımıza göre f her sınırlı aralıkta sonlu sayıda nokta dışında sürekli. E , f 'nin süreksiz olduğu noktaların kümesi olsun.

Adım 1: f 'nin sürekli olduğu her x_0 noktasında $F'(x_0) = f(x_0)$ 'dır. $\varepsilon > 0$ verilsin. Süreklilikten öyle bir $\delta > 0$ vardır ki $|t - x_0| < \delta$ için $|f(t) - f(x_0)| < \varepsilon$ olur. $0 < |h| < \delta$ alalım. $h > 0$ ise

$$\frac{F(x_0 + h) - F(x_0)}{h} - f(x_0) = \frac{1}{h} \int_{x_0}^{x_0+h} f(t) dt - \frac{1}{h} \int_{x_0}^{x_0+h} f(x_0) dt = \frac{1}{h} \int_{x_0}^{x_0+h} (f(t) - f(x_0)) dt$$

olur ve mutlak değeri $\frac{1}{h} \int_{x_0}^{x_0+h} \varepsilon dt = \varepsilon$ 'u aşmaz. $h < 0$ için aynı hesap $\int_{x_0+h}^{x_0}$ ile yapılır ve yine ε sınırı çıkar. Demek ki fark oranı $h \rightarrow 0$ iken $f(x_0)$ 'a gider: F, x_0 'da türevlenebilirdir ve $F'(x_0) = f(x_0)$. Bu, analizin temel teoremidir; yalnızca türev alınan noktada süreklilik gerektirdiğine dikkat edelim.

Adım 2: g ile f yalnızca E üzerinde farklı olabilir. Adım 1'e göre $\mathbb{R} \setminus E \subseteq A$ ve $x \notin E$ için $g(x) = F'(x) = f(x)$ 'tir. Yani $\{x : g(x) \neq f(x)\} \subseteq E$.

Adım 3: g bir yoğunluktur. Her sınırlı $[u, v]$ aralığında E sonlu olduğundan g ile f bu aralıkta sonlu sayıda nokta dışında çakışır; sonlu sayıda noktada değeri değiştirilen bir fonksiyonun integrali değişmez. Öyleyse her $u < v$ için $\int_u^v g = \int_u^v f$ olur ve $u \rightarrow -\infty$ limitiyle her x için

$$\int_{-\infty}^x g(t) dt = \int_{-\infty}^x f(t) dt = F(x)$$

elde edilir; $x \rightarrow \infty$ limitiyle de $\int_{-\infty}^{\infty} g = 1$. Son olarak $g \geq 0$ 'dır: $x \notin A$ için $g(x) = 0$; $x \in A$ için F azalmayan olduğundan fark oranları $\frac{F(x+h)-F(x)}{h} \geq 0$ 'dır ve limitleri $F'(x) \geq 0$ olur. Böylece g tanımın (i) ve (ii) koşullarını sağlar ve F 'yi üretir; X 'in bir yoğunluğudur. Son cümledeki özel durum Adım 1'dir. ■

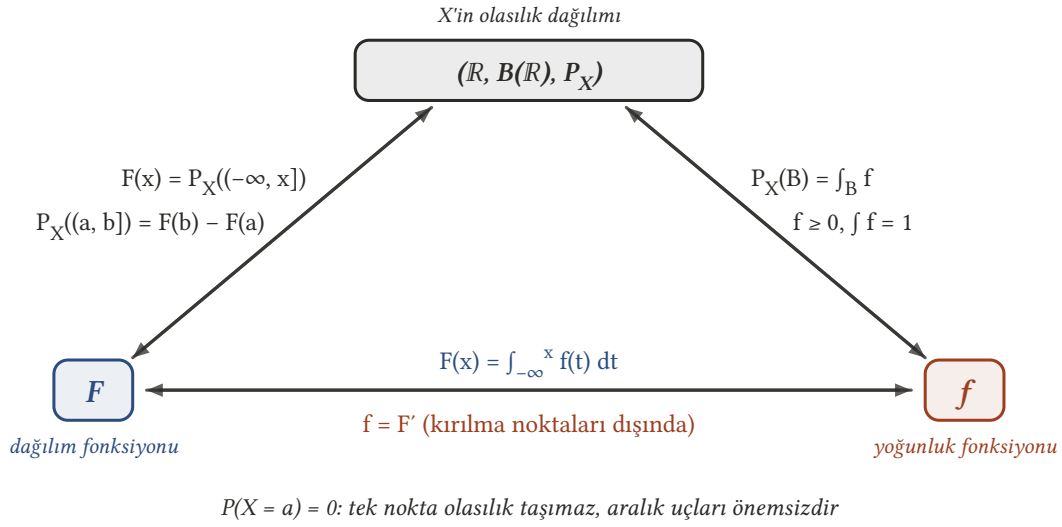
Teoremin üç noktasını vurgulayalım.

- **Yoğunluk tek değildir.** Bir yoğunluğun sonlu sayıda noktadaki değerini değiştirirsek integraller, dolayısıyla F ve bütün olasılıklar aynı kalır; yeni fonksiyon da X 'in bir yoğunluğudur. Örneğin $[0, 1]$ üzerinde 1, dışında 0 olan yoğunlukla $(0, 1)$ üzerinde 1, dışında 0 olan yoğunluk aynı değişkeni anlatır. Bu yüzden parçalı tanımlarda uç noktaları hangi parçaya kattığımız önemsizdir; teoremdaki “ A dışında 0” anlaşıması yalnızca g 'yi her yerde tanımlı kılmak içindir. Türevin var olmadığı noktalar tipik olarak F 'nin grafiğinin köşeleridir ve bunlar her sınırlı aralıkta sonlu sayıdadır.
- **Türev yalnızca F 'nin türevlenebildiği yerde alınır.** Uygulamada F parçalı formüllerle verilir; her parçanın içinde türev alınır. Bir parça ucunda sol ve sağ türev çakışıyorsa F orada türevlenebilirdir ve g bu ortak değeri alır; çakışmıyorsa F orada türevlenemez ve $g = 0$ konur. Sonlu sayıdaki bu uç noktalarda hangi değerin seçildiği hiçbir olasılığı değiştirmez.
- **Sürekli F tek başına yetmez.** Teorem “ X sürekli (yoğunluğu var)” varsayımıyla başlar. Her yerde sürekli olduğu ve türevinin var olduğu her noktada türevi 0 olduğu hâlde 0'dan 1'e yükselen dağılım fonksiyonları vardır; böyle bir F için $\int_{-\infty}^x F'(t) dt = 0 \neq F(x)$ olur ve türev bir yoğunluk vermez. Bu tür değişkenler bu kitapta karşımıza çıkmayacaktır.

Sonuç olarak, yoğunluğu olan bir değişken için F ile f aynı bilgiyi taşır: f 'den F 'ye integrale, F 'den f 'ye türevle geçilir; olasılıklar da ya F 'nin farkı ya f 'nin integrali olarak hesaplanır. Bu zincir

$$(\Omega, \mathcal{U}, P) \rightarrow P_X \rightarrow F \xrightleftharpoons[\text{türev}]{\text{integral}} f, \quad P(a < X \leq b) = F(b) - F(a) = \int_a^b f(x) dx$$

biçiminde özetlenir; P_X, X 'in olasılık dağılımıdır (Tanım 8.3).



Şekil 11.3: Sürekli bir X için üç nesne aynı dağılımı anlatır: olasılık dağılımı P_X , dağılım fonksiyonu F ve yoğunluk f . Kenarlardaki kurallar birinden ötekine nasıl geçileceğini söyler: F , f 'nin integrali; f , F 'nin türevidir; olasılıklar ya F 'nin farkı ya f 'nin integralidir. Kesikli durumdan farkı, hiçbir tek noktanın olasılık taşımasıdır.

11.5 Örnekler

Aşağıdaki örneklerde iki yönü de çalışacağız: yoğunluktan dağılım fonksiyonuna (integral) ve dağılım fonksiyonundan yoğunluğa (türev). Her örnekte F 'nin sürekliliğini ve $\int f = 1$ eşitliğini denetlemek, hesap hatalarını yakalamanın en kestirme yoludur.

11.5.1 Yoğunluktan Dağılım Fonksiyonuna

Örnek 11.2 (Sinüs Biçimli Yoğunluk). X rastgele değişkeninin yoğunluk fonksiyonu

$$f(x) = \begin{cases} c \sin x, & 0 \leq x \leq \pi \\ 0, & \text{diğer durumlarda} \end{cases}$$

biçiminde veriliyor.

a) c sabitini bulunuz.

b) X 'in dağılım fonksiyonunu bulunuz.

c) $(X \leq \frac{\pi}{3})$ ve $(\frac{\pi}{4} \leq X \leq \frac{3\pi}{4})$ olasılıklarını hesaplayınız.

Çözüm.

a) $[0, \pi]$ üzerinde $\sin x \geq 0$ olduğundan $f \geq 0$ koşulu $c \geq 0$ gerektirir. Normalleştirme koşulu

$$1 = \int_{-\infty}^{\infty} f(x) dx = \int_0^{\pi} c \sin x dx = c \left[-\cos x \right]_0^{\pi} = c (-\cos \pi + \cos 0) = c(1 + 1) = 2c$$

verir; $c = \frac{1}{2}$. **Önerme 11.2** gereği $f(x) = \frac{1}{2} \sin x$ ($0 \leq x \leq \pi$) gerçekten bir yoğunluktur.

b) $x < 0$ için f sıfır olduğundan $F(x) = 0$. $0 \leq x \leq \pi$ için

$$F(x) = \int_0^x \frac{1}{2} \sin t dt = \frac{1}{2} \left[-\cos t \right]_0^x = \frac{1}{2} (1 - \cos x).$$

$x > \pi$ için $F(x) = \int_0^{\pi} \frac{1}{2} \sin t dt + 0 = 1$. Demek ki

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{1-\cos x}{2}, & 0 \leq x \leq \pi \\ 1, & x > \pi. \end{cases}$$

Denetim: $F(0) = \frac{1-1}{2} = 0$ ve $F(\pi) = \frac{1-(-1)}{2} = 1$; parça uçlarında değerler çakışır, F süreklidir. $(0, \pi)$ üzerinde $F'(x) = \frac{\sin x}{2} = f(x)$ 'tir (Teorem 11.2).

c) $\cos \frac{\pi}{3} = \frac{1}{2}$ olduğundan

$$\left(PX \leq \frac{\pi}{3}\right) = \left(F\frac{\pi}{3}\right) = \frac{1 - \frac{1}{2}}{2} = \frac{1}{4}.$$

İkinci olasılıkta uçların dahil olması sonucu değiştirmez (Önerme 11.1); $\cos \frac{\pi}{4} = \frac{\sqrt{2}}{2}$ ve $\cos \frac{3\pi}{4} = -\frac{\sqrt{2}}{2}$ ile

$$\left(F\frac{\pi}{4} \leq X \leq \frac{3\pi}{4}\right) = \left(F\frac{3\pi}{4}\right) - \left(F\frac{\pi}{4}\right) = \frac{\cos \frac{\pi}{4} - \cos \frac{3\pi}{4}}{2} = \frac{\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}}{2} = \frac{\sqrt{2}}{2} \approx 0,707.$$

Sinüs eğrisi $\frac{\pi}{2}$ çevresinde yoğunlaştığından, uzunluğu toplam aralığın yarısı olan bu orta parça olasılığın %70'inden fazlasını taşır.

■

Örnek 11.3 (Parçalı Sabit Yoğunluk). X rastgele değişkeninin yoğunluk fonksiyonu

$$f(x) = \begin{cases} \frac{1}{2}, & 0 \leq x \leq 1 \\ \frac{1}{4}, & 3 \leq x \leq 5 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun. X 'in dağılım fonksiyonunu bulunuz; $\left(\frac{\pi}{2} \leq X \leq 4\right)$, $P(1 < X < 3)$ ve $P(X > 4)$ olasılıklarını hesaplayınız.

Çözüm.

Önce f 'nin bir yoğunluk olduğunu doğrulayalım: $f \geq 0$ ve

$$\int_{-\infty}^{\infty} f(x) dx = \int_0^1 \frac{1}{2} dx + \int_3^5 \frac{1}{4} dx = \frac{1}{2} + \frac{2}{4} = 1.$$

Şimdi $F(x) = \int_{-\infty}^x f(t) dt$ 'yi f 'nin parçalarına göre hesaplayalım.

- $x < 0$: $F(x) = 0$.
- $0 \leq x \leq 1$: $F(x) = \int_0^x \frac{1}{2} dt = \frac{x}{2}$.
- $1 < x < 3$: $F(x) = \int_0^1 \frac{1}{2} dt + \int_1^x 0 dt = \frac{1}{2}$.
- $3 \leq x \leq 5$: $F(x) = \frac{1}{2} + \int_3^x \frac{1}{4} dt = \frac{1}{2} + \frac{x-3}{4} = \frac{x-1}{4}$.
- $x > 5$: $F(x) = \frac{1}{2} + \frac{2}{4} = 1$.

Toplu hâlde

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{x}{2}, & 0 \leq x \leq 1 \\ \frac{1}{2}, & 1 < x < 3 \\ \frac{x-1}{4}, & 3 \leq x \leq 5 \\ 1, & x > 5. \end{cases}$$

Denetim: $F(1) = \frac{1}{2}$, $F(3) = \frac{3-1}{4} = \frac{1}{2}$, $F(5) = \frac{5-1}{4} = 1$; F süreklidir. F 'nin grafiği $[0, 1]$ üzerinde eğimi $\frac{1}{2}$ olan bir doğru parçası, $(1, 3)$ üzerinde yatay bir çizgi, $[3, 5]$ üzerinde eğimi $\frac{1}{4}$ olan bir doğru parçasıdır; yoğunluğun sıfır olduğu yerde F **sabit kalır**, yoğunluğun büyük olduğu yerde daha dik yükselir.

Olasılıklar:

$$\left(\frac{1}{2} \leq X \leq 4\right) = F(4) - \left(\frac{1}{2}\right) = \frac{3}{4} - \frac{1}{4} = \frac{1}{2},$$

$$P(1 < X < 3) = F(3) - F(1) = \frac{1}{2} - \frac{1}{2} = 0, \quad P(X > 4) = 1 - F(4) = 1 - \frac{3}{4} = \frac{1}{4}.$$

$P(1 < X < 3) = 0$ sonucu doğaldır: yoğunluk bu aralıkta sıfırdır, dolayısıyla X 'in buraya düşme olasılığı sıfırdır. Sürekli bir değişkende sıfır olasılıklı olay yalnızca tek noktalar değildir; yoğunluğun sıfır olduğu her bölge de sıfır olasılık taşır.

■

Örnek 11.4 (Üçgen Yoğunluk). X rastgele değişkeninin yoğunluk fonksiyonu

$$f(x) = \begin{cases} 1 - |x|, & |x| < 1 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun. X 'in dağılım fonksiyonunu bulunuz ve $P(X < \frac{1}{2})$ olasılığını hesaplayınız.

Çözüm.

f 'nin grafiği köşeleri $(-1, 0)$, $(0, 1)$ ve $(1, 0)$ olan bir üçgendir; altında kalan alan $\frac{1}{2} \cdot 2 \cdot 1 = 1$ 'dir, dolayısıyla f bir yoğunluktur. Mutlak değeri açalım: $-1 < x < 0$ için $f(x) = 1 + x$, $0 \leq x < 1$ için $f(x) = 1 - x$.

- $x < -1$: $F(x) = 0$.
- $-1 \leq x \leq 0$:

$$F(x) = \int_{-1}^x (1+t) dt = \left[t + \frac{t^2}{2} \right]_{-1}^x = x + \frac{x^2}{2} - \left(-1 + \frac{1}{2} \right) = \frac{x^2}{2} + x + \frac{1}{2} = \frac{(1+x)^2}{2}.$$

- $0 \leq x \leq 1$: $F(0) = \frac{1}{2}$ üzerine sağ parçanın integrali eklenir:

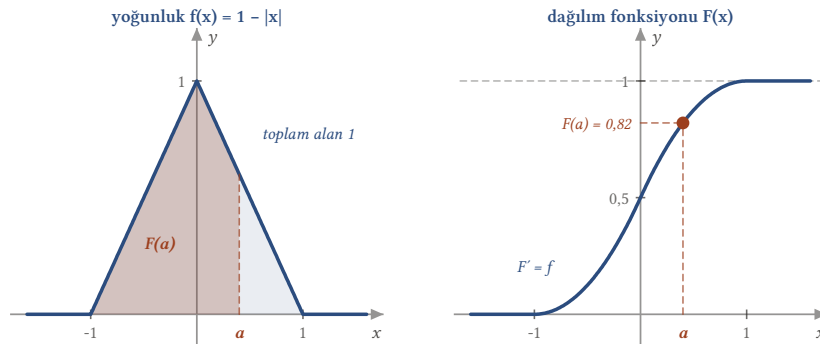
$$F(x) = \frac{1}{2} + \int_0^x (1-t) dt = \frac{1}{2} + x - \frac{x^2}{2} = 1 - \frac{(1-x)^2}{2}.$$

- $x > 1$: $F(x) = 1$.

Böylece

$$F(x) = \begin{cases} 0, & x < -1 \\ \frac{(1+x)^2}{2}, & -1 \leq x \leq 0 \\ 1 - \frac{(1-x)^2}{2}, & 0 < x \leq 1 \\ 1, & x > 1. \end{cases}$$

Denetim: $F(-1) = 0$, iki orta parçanın formülleri $x = 0$ 'da aynı $\frac{1}{2}$ değerini verir (bu yüzden 0 noktasını hangi parçaya kattığımızı önemsizdir), $F(1) = 1$. F süreklidir; $(-1, 0)$ üzerinde $F'(x) = 1 + x$, $(0, 1)$ üzerinde $F'(x) = 1 - x$, yani $F' = f$ (Teorem 11.2). F 'nin grafiği $[-1, 0]$ üzerinde yukarı bükey, $[0, 1]$ üzerinde aşağı bükey bir "S" eğrisidir; en dik olduğu yer, yoğunluğun tepe noktası olan $x = 0$ 'dır. Ayrıca f çift fonksiyon olduğundan $F(-x) = 1 - F(x)$ bakışımı vardır.



Şekil 11.4: Üçgen yoğunluk ve dağılım fonksiyonu. Solda $F(a) = P(X \leq a)$, yoğunluğun a 'nın solunda kalan alanıdır; sağda aynı değer eğri üzerinde okunur. Yoğunluk kırık olduğu hâlde dağılım fonksiyonu sürekli ve $F = f$ bağıntısı kırılma noktaları dışında geçerlidir.

Olasılık için uçları katmak katmamak fark etmez:

$$\left(P \mid X \mid < \frac{1}{2}\right) = \left(P - \frac{1}{2} < X < \frac{1}{2}\right) = \left(F\left(\frac{1}{2}\right) - F\left(-\frac{1}{2}\right)\right) = \left(1 - \frac{1}{8}\right) - \frac{1}{8} = \frac{3}{4}.$$

Geometrik denetim: $-\frac{1}{2} < x < \frac{1}{2}$ üzerindeki alan, büyük üçgenden iki uçtaki küçük üçgenler çıkarılarak bulunur; küçük üçgenlerin her birinin tabanı $\frac{1}{2}$, yüksekliği $\frac{1}{2}$, alanı $\frac{1}{8}$ 'dir. $1 - 2 \cdot \frac{1}{8} = \frac{3}{4}$.

■

11.5.2 Dağılım Fonksiyonundan Yoğunluğa

Örnek 11.5 (Verilen Dağılım Fonksiyonundan Yoğunluk). X rastgele değişkeninin dağılım fonksiyonu

$$F(x) = \begin{cases} 0, & x < -1 \\ \frac{1+x}{2}, & -1 \leq x < 0 \\ \frac{1+x^2}{2}, & 0 \leq x < 1 \\ 1, & x \geq 1 \end{cases}$$

ise X 'in sürekli olduğunu gösteriniz, yoğunluk fonksiyonunu bulunuz ve $\left(P - \frac{1}{2} < X < \frac{1}{2}\right)$ olasılığını iki yolla hesaplayınız.

Çözüm.

F 'nin incelenmesi. Parça uçlarında soldan ve sağdan değerlere bakalım:

- $x = -1$: sol parça 0, sağ parça $\frac{1+(-1)}{2} = 0$.
- $x = 0$: sol parça $\frac{1+0}{2} = \frac{1}{2}$, sağ parça $\frac{1+0}{2} = \frac{1}{2}$.
- $x = 1$: sol parça $\frac{1+1}{2} = 1$, sağ parça 1.

Her uçta değerler çakışır; F sürekli, dolayısıyla hiçbir noktada sıçramaz ve $P(X = a) = F(a) - F(a^-) = 0$ 'dır. Parçaların her biri azalmayan olduğundan F azalmayan; $x \rightarrow -\infty$ iken 0, $x \rightarrow \infty$ iken 1'dir. Yani F , sıçraması olmayan bir dağılım fonksiyonudur.

Yoğunluk. Teorem 11.2'yi uyguluyoruz. Parçaların içinde türev alalım:

$$F'(x) = \begin{cases} 0, & x < -1 \\ \frac{1}{2}, & -1 < x < 0 \\ x, & 0 < x < 1 \\ 0, & x > 1. \end{cases}$$

Köşe noktalarında F türevlenebilir değildir: $x = -1$ 'de sol türev 0, sağ türev $\frac{1}{2}$; $x = 0$ 'da sol türev $\frac{1}{2}$, sağ türev $\frac{d}{dx}\left(\frac{1+x^2}{2}\right)|_{x=0} = 0$; $x = 1$ 'de sol türev 1, sağ türev 0. Dolayısıyla $A = \mathbb{R} \setminus \{-1, 0, 1\}$ ve bu üç noktada yoğunluğa 0 değeri verilir:

$$f(x) = \begin{cases} \frac{1}{2}, & -1 < x < 0 \\ x, & 0 < x < 1 \\ 0, & \text{diğer durumlarda.} \end{cases}$$

Denetim: $\int_{-\infty}^{\infty} f(x) dx = \int_{-1}^0 \frac{1}{2} dx + \int_0^1 x dx = \frac{1}{2} + \frac{1}{2} = 1$. Ayrıca $\int_{-1}^x \frac{1}{2} dt = \frac{x+1}{2}$ ve $\frac{1}{2} + \int_0^x t dt = \frac{1+x^2}{2}$ olduğundan bu f 'nin integrali gerçekten verilen F 'yi geri verir; X sürekli ve yoğunluğu f 'dir. Yoğunluğun $-1, 0$ ve 1 noktalarındaki değerlerinin hiçbir olasılığı etkilemediğini bir kez daha not edelim.

Olasılık, iki yolla. Dağılım fonksiyonuyla:

$$\left(P - \frac{1}{2} < X < \frac{1}{2}\right) = \left(F - \frac{1}{2}\right) - \left(F - \frac{1}{2}\right) = \frac{1 + \frac{1}{4}}{2} - \frac{1 - \frac{1}{2}}{2} = \frac{5}{8} - \frac{1}{4} = \frac{3}{8}.$$

Yoğunlukla:

$$\left(P - \frac{1}{2} < X < \frac{1}{2}\right) = \int_{-1/2}^0 \frac{1}{2} dx + \int_0^{1/2} x dx = \frac{1}{4} + \frac{1}{8} = \frac{3}{8}.$$

İki yol aynı sonucu verir; hangisinin kullanılacağı, elde F 'nin mi f 'nin mi olduğuna bağlıdır.

■

11.5.3 Bir Ömür Modeli

Sürekli değişkenlerin en yaygın kullanım alanlarından biri, bir sürecin ne kadar süreceğinin modellenmesidir. Aşağıdaki yoğunluk, **üstel dağılımın** bir örneğidir; ilerleyen bölümlerde ayrıntılı olarak incelenecektir. Burada ana özelliğini bir hesaplama keşfedeceğiz.

Örnek 11.6 (Bir Elektronik Parçanın Ömrü). X rastgele değişkeni belli bir tür elektronik parçanın yıl olarak dayanma süresini gösterebilir ve yoğunluk fonksiyonu

$$f(x) = \begin{cases} c e^{-x/5}, & x \geq 0 \\ 0, & \text{diğer durumlarda} \end{cases}$$

biçiminde verilsin.

- f 'nin bir yoğunluk fonksiyonu olması için c değerini bulunuz.
- X 'in dağılım fonksiyonunu bulunuz.
- Parçanın 5 yıldan uzun dayanma olasılığını ve 2 ile 8 yıl arasında bozulma olasılığını hesaplayınız.
- 5 yıl dayanmış bir parçanın en az 5 yıl daha dayanma olasılığını hesaplayınız ve sonucu yorumlayınız.

Çözüm.

a) $e^{-x/5} > 0$ olduğundan $f \geq 0$ için $c \geq 0$ gerekir. Has olmayan integrali hesaplayalım:

$$\int_{-\infty}^{\infty} f(x) dx = \int_0^{\infty} c e^{-x/5} dx = c \lim_{M \rightarrow \infty} \left[-5e^{-x/5} \right]_0^M = c \lim_{M \rightarrow \infty} (5 - 5e^{-M/5}) = 5c.$$

$5c = 1$ olmalı; $c = \frac{1}{5}$. **Önerme 11.2** gereği $f(x) = \frac{1}{5}e^{-x/5}$ ($x \geq 0$) bir yoğunluktur.

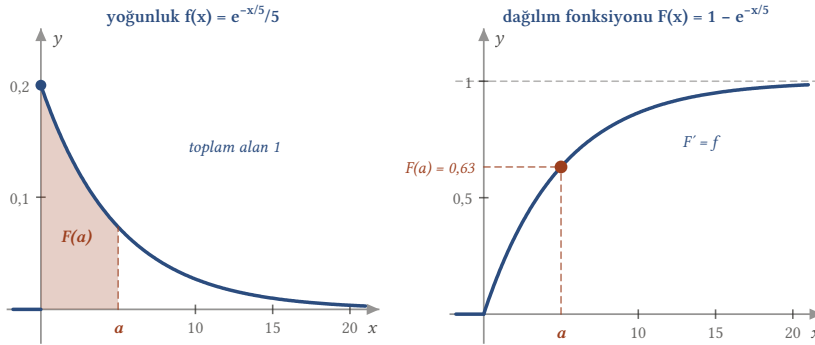
b) $x < 0$ için $F(x) = 0$. $x \geq 0$ için

$$F(x) = \int_0^x \frac{1}{5} e^{-t/5} dt = \left[-e^{-t/5} \right]_0^x = 1 - e^{-x/5}.$$

Böylece

$$F(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-x/5}, & x \geq 0. \end{cases}$$

Denetim: $F(0) = 1 - 1 = 0$, F sürekli; $x \rightarrow \infty$ iken $F(x) \rightarrow 1$; $x > 0$ için $F'(x) = \frac{1}{5}e^{-x/5} = f(x)$.



Şekil 11.5: Üstel yoğunluk $f(x) = e^{-x/5}/5$ ($x \geq 0$) ve dağılım fonksiyonu $F(x) = 1 - e^{-x/5}$. Solda $F(a)$, yoğunluğun a 'nın solunda kalan alanıdır; sağda aynı değer eğri üzerinde okunur. Yoğunluk sıfırda $1/5$ 'ten başlayıp söner; F sürekli olarak artar, 1'e yaklaşır ama hiçbir noktada 1'e ulaşmaz.

c) Parçanın 5 yıldan uzun dayanması ($X > 5$) olayıdır:

$$P(X > 5) = 1 - F(5) = e^{-1} \approx 0,368.$$

2 ile 8 yıl arasında bozulma ($2 \leq X \leq 8$) olayıdır; uçlar önemsizdir:

$$P(2 \leq X \leq 8) = F(8) - F(2) = (1 - e^{-8/5}) - (1 - e^{-2/5}) = e^{-2/5} - e^{-8/5} \approx 0,670 - 0,202 = 0,468.$$

d) İstenen, koşullu olasılık (Tanım 6.1) $P(X \geq 10 \mid X > 5)$ 'tir; uçlar önemsiz olduğundan (Önerme 11.1) bunu $P(X > 10 \mid X > 5)$ olarak hesaplayabiliriz. ($X > 10$) $\subseteq$ ($X > 5$) olduğundan

$$P(X > 10 \mid X > 5) = \frac{P(X > 10)}{P(X > 5)} = \frac{1 - F(10)}{1 - F(5)} = \frac{e^{-2}}{e^{-1}} = e^{-1} \approx 0,368.$$

Bu, tam olarak $P(X > 5)$ 'e eşittir: 5 yıl çalışmış bir parçanın 5 yıl daha dayanma olasılığı, yeni bir parçanın 5 yıl dayanma olasılığıyla aynıdır. Aynı hesap her $s, t \geq 0$ için

$$P(X > s + t \mid X > s) = \frac{e^{-(s+t)/5}}{e^{-s/5}} = e^{-t/5} = P(X > t)$$

verir. Parça “yaşlanmaz”, geçmişini unuttur; üstel dağılımın bu **hafızasızlık** özelliği, onu yıpranmanın ihmal edilebildiği ömür ve bekleme süresi modellerinin standart seçimi yapar.

■

11.6 Alıştırmalar

Alıştırma 11.1 (Sürekli Rastgele Değişkenler ve Yoğunluk Üzerine). a) $f(x) = cx$ ($0 \leq x \leq 2$), diğer durumlarda 0, bir yoğunluk fonksiyonu olacak biçimde c 'yi bulunuz. Dağılım fonksiyonunu yazınız; $P(X > 1)$ ve $(\frac{1}{2} < X < \frac{3}{2})$ olasılıklarını hesaplayınız.

b) $f(x) = \frac{c}{x^2}$ ($x \geq 1$), diğer durumlarda 0, için c 'yi bulunuz. Dağılım fonksiyonunu yazınız ve $P(2 < X < 4)$ 'ü hesaplayınız.

c) $f(x) = \frac{1}{2\sqrt{x}}$ ($0 < x \leq 1$), diğer durumlarda 0, fonksiyonunun sınırsız olmasına karşın bir yoğunluk olduğunu gösteriniz. Dağılım fonksiyonunu bulunuz ve $(X < \frac{1}{4})$ 'ü hesaplayınız.

d) $F(x) = 1 - e^{-x^2}$ ($x \geq 0$), $F(x) = 0$ ($x < 0$) fonksiyonunun sürekli bir rastgele değişkenin dağılım fonksiyonu olduğunu gösteriniz; yoğunluğu bulunuz; $P(X > 1)$ ve $P(1 < X < 2)$ 'yi hesaplayınız.

e) $g(x) = |x|$ ($-1 \leq x \leq 1$), diğer durumlarda 0, bir yoğunluk mudur? Öyleyse dağılım fonksiyonunu bulunuz ve $(\frac{1}{2} < X < \frac{1}{2})$ 'yi hesaplayınız.

f) $F(x) = 0$ ($x < 0$), $F(x) = \frac{x}{2}$ ($0 \leq x < 1$), $F(x) = 1$ ($x \geq 1$) bir dağılım fonksiyonudur. Bu dağılıma sahip bir X sürekli midir? $P(X = 1)$ ve $P(X < 1)$ 'i bulunuz.

g) $f(x) = c e^{-|x|}$ ($x \in \mathbb{R}$) için c 'yi bulunuz, dağılım fonksiyonunu yazınız ve $P(|X| > 1)$ 'i hesaplayınız.

Çözüm.

a) $\int_0^2 c x dx = c \left[\frac{x^2}{2} \right]_0^2 = 2c = 1$ verir; $c = \frac{1}{2}$ ve $f \geq 0$ sağlanır. $0 \leq x \leq 2$ için $F(x) = \int_0^x \frac{t}{2} dt = \frac{x^2}{4}$; dolayısıyla

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{x^2}{4}, & 0 \leq x \leq 2 \\ 1, & x > 2. \end{cases}$$

$P(X > 1) = 1 - F(1) = 1 - \frac{1}{4} = \frac{3}{4}$ ve $P(\frac{1}{2} < X < \frac{3}{2}) = (F(\frac{3}{2}) - F(\frac{1}{2})) = \frac{9}{16} - \frac{1}{16} = \frac{1}{2}$.

b) $\int_1^\infty \frac{c}{x^2} dx = c \lim_{M \rightarrow \infty} \left[-\frac{1}{x} \right]_1^M = c \lim_{M \rightarrow \infty} \left(1 - \frac{1}{M} \right) = c$; öyleyse $c = 1$. $x \geq 1$ için $F(x) = \int_1^x \frac{1}{t^2} dt = 1 - \frac{1}{x}$, $x < 1$ için $F(x) = 0$. $F(1) = 0$ olduğundan F süreklidir.

$$P(2 < X < 4) = F(4) - F(2) = \left(1 - \frac{1}{4} \right) - \left(1 - \frac{1}{2} \right) = \frac{1}{4}.$$

c) $x \rightarrow 0^+$ iken $\frac{1}{2\sqrt{x}} \rightarrow \infty$; f sınırsızdır ve integral 0'da has olmayan integraldir:

$$\int_0^1 \frac{1}{2\sqrt{x}} dx = \lim_{\varepsilon \rightarrow 0^+} \int_\varepsilon^1 \frac{1}{2\sqrt{x}} dx = \lim_{\varepsilon \rightarrow 0^+} \left[\sqrt{x} \right]_\varepsilon^1 = \lim_{\varepsilon \rightarrow 0^+} (1 - \sqrt{\varepsilon}) = 1.$$

$f \geq 0$ da sağlandığından **Önerme 11.2** gereği f bir yoğunluktur. $0 \leq x \leq 1$ için $F(x) = \lim_{\varepsilon \rightarrow 0^+} \int_\varepsilon^x \frac{1}{2\sqrt{t}} dt = \sqrt{x}$; dolayısıyla

$$F(x) = \begin{cases} 0, & x < 0 \\ \sqrt{x}, & 0 \leq x \leq 1 \\ 1, & x > 1. \end{cases}$$

F süreklidir (yoğunluğun sınırsızlığı F 'nin sürekliliğini bozmaz, yalnızca 0'daki eğimini sonsuz yapar). $P(X < \frac{1}{4}) = (F(\frac{1}{4})) = \frac{1}{2}$.

d) $F(0) = 1 - e^0 = 0$ olduğundan F , 0'da süreklidir; $x \geq 0$ üzerinde $1 - e^{-x^2}$ sürekli ve kesin artandır; $x \rightarrow \infty$ iken $e^{-x^2} \rightarrow 0$, yani $F(x) \rightarrow 1$. F sürekli bir dağılım fonksiyonudur. $x \neq 0$ noktalarında türev hemen bulunur: $x > 0$ için $F'(x) = 2xe^{-x^2}$, $x < 0$ için $F'(x) = 0$. Yoğunluk

$$f(x) = \begin{cases} 2xe^{-x^2}, & x > 0 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olur; $f \geq 0$ 'dır ve $\int_0^\infty 2xe^{-x^2} dx = \lim_{M \rightarrow \infty} \left[-e^{-x^2} \right]_0^M = 1$ olduğundan F gerçekten f 'nin integratördür; X süreklidir. ($x = 0$ 'da aslında $F'(0) = 0$ da vardır, çünkü sağ türev $2 \cdot 0 \cdot e^0 = 0$ ve sol türev 0'dır; bu noktadaki değer zaten önemsizdir.)

$$P(X > 1) = 1 - F(1) = e^{-1} \approx 0,368, \quad P(1 < X < 2) = F(2) - F(1) = e^{-1} - e^{-4} \approx 0,350.$$

e) $g \geq 0$ 'dır ve $\int_{-1}^1 |x| dx = 2 \int_0^1 x dx = 1$; evet, g bir yoğunluktur. Dağılım fonksiyonu: $-1 \leq x \leq 0$ için

$$F(x) = \int_{-1}^x (-t) dt = \left[-\frac{t^2}{2} \right]_{-1}^x = \frac{1-x^2}{2};$$

$0 \leq x \leq 1$ için $F(x) = \frac{1}{2} + \int_0^x t dt = \frac{1+x^2}{2}$; $x < -1$ için 0, $x > 1$ için 1. Uçlarda $F(-1) = 0$, $F(0) = \frac{1}{2}$, $F(1) = 1$ değerleri iki taraftan da çıkarılır. Bu F , **Örnek 11.5**'teki dağılım fonksiyonuna benzer ama aynı değildir: orada sol parça doğrusal, burada paraboliktir.

$$P\left(\frac{1}{2} < X < \frac{1}{2}\right) = \left(F\left(\frac{1}{2}\right) - F\left(-\frac{1}{2}\right)\right) = \frac{1+\frac{1}{4}}{2} - \frac{1-\frac{1}{4}}{2} = \frac{5}{8} - \frac{3}{8} = \frac{1}{4}.$$

Geometrik denetim: $|x|$ 'in $[-\frac{1}{2}, \frac{1}{2}]$ üzerindeki grafiği altında iki küçük üçgen vardır, her birinin alanı $\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$; toplam $\frac{1}{4}$. Yoğunluk merkezde sıfır olduğundan bu aralık, uzunluğu toplam aralığın yarısı olmasına karşın olasılığın yalnızca dörtte birini taşır.

f) $x = 1$ noktasında $F(1^-) = \lim_{x \rightarrow 1^-} \frac{x}{2} = \frac{1}{2}$ ve $F(1) = 1$ 'dir; F burada $\frac{1}{2}$ yüksekliğinde sıçrar. Teorem 9.2 gereği

$$P(X = 1) = F(1) - F(1^-) = 1 - \frac{1}{2} = \frac{1}{2}, \quad P(X < 1) = F(1^-) = \frac{1}{2}.$$

$P(X = 1) > 0$ olduğundan X 'in yoğunluğu yoktur; yani X , bu kitaptaki anlamıyla sürekli **değildir** (Önerme 11.1). Değer kümesi $[0, 1]$ sayılamaz olduğu hâlde bu böyledir; Tanım 11.1'deki koşul tek başına yoğunluğun varlığını güvence altına almaz. Kesikli de değildir: $[0, 1]$ aralığındaki her nokta değer kümesindedir ve $P(0 \leq X < 1) = \frac{1}{2}$ olasılığı bu sayılamaz kümeye yayılmıştır; $[0, 1]$ üzerinde $F'(x) = \frac{1}{2}$ 'dir ama $\int_0^1 \frac{1}{2} dx = \frac{1}{2} \neq 1$ olduğundan türev bir yoğunluk vermez. X , olasılığının yarısını $[0, 1]$ üzerine düzgün yayan, öteki yarısını 1 noktasında toplayan **karma** bir değişkendir. (Böyle bir değişkeni şöyle düşünebiliriz: $[0, 2]$ aralığında düzgün seçilen bir U sayısı için $X = \min(U, 1)$.)

g) $e^{-|x|}$ çift fonksiyondur:

$$\int_{-\infty}^{\infty} c e^{-|x|} dx = 2c \int_0^{\infty} e^{-x} dx = 2c \lim_{M \rightarrow \infty} (1 - e^{-M}) = 2c;$$

$c = \frac{1}{2}$. $x < 0$ için $F(x) = \int_{-\infty}^x \frac{1}{2} e^t dt = \frac{e^x}{2}$; $x \geq 0$ için $F(x) = \frac{1}{2} + \int_0^x \frac{1}{2} e^{-t} dt = \frac{1}{2} + \frac{1 - e^{-x}}{2} = 1 - \frac{e^{-x}}{2}$. Yani

$$F(x) = \begin{cases} \frac{e^x}{2}, & x < 0 \\ 1 - \frac{e^{-x}}{2}, & x \geq 0. \end{cases}$$

$x = 0$ 'da iki parça da $\frac{1}{2}$ verir; F sürekli. ($|X| > 1$) = $(X < -1) \cup (X > 1)$ ayrık birleşimi için

$$P(|X| > 1) = F(-1) + (1 - F(1)) = \frac{e^{-1}}{2} + \frac{e^{-1}}{2} = e^{-1} \approx 0,368.$$

■

Bu bölümde sürekli bir rastgele değişkenin bütün olasılık bilgisinin yoğunluk fonksiyonunda saklı olduğunu, F ile f arasında integral ve türevle gidilip gelindiğini gördük. Sıradaki doğal soru şudur: X 'in yoğunluğu biliniyorsa $Y = g(X)$ dönüşümünün, örneğin X^2 'nin ya da e^X 'in, yoğunluğu nedir? Yanıt, bu bölümdeki türev alma fikrinin bir adım ötesidir: **Rastgele Değişkenin Fonksiyonunun Dağılımı**.

12. Rastgele Değişkenin Fonksiyonunun Dağılımı

Sürekli Rastgele Değişkenler ve Yoğunluk Fonksiyonu bölümünde bir rastgele değişkenin dağılımını iki biçimde tanımayı öğrendik: kesikli değişkende olasılık fonksiyonu f_X , sürekli değişkende olasılık yoğunluk fonksiyonu f_X ve her iki durumda da dağılım fonksiyonu F_X . Uygulamada çoğu zaman elimizdeki değişkenin kendisi değil, ondan türetilmiş bir büyüklük ilgimizi çeker: bir ölçümün karesi, bir sıcaklığın başka bir birime çevrilmiş hâli, bir bekleme süresinin logaritması, rastgele bir açının tanjantı. Bunların hepsi $Y = g(X)$ biçimindedir.

Bu bölümde iki soruya yanıt vereceğiz. Birincisi kuramsal: $g(X)$ gerçekten bir rastgele değişken midir, yani $(g(X) \leq a)$ kümeleri olay mıdır? Yanıtın “evet” olması için g 'nin Borel-ölçülebilir olması yeter; sürekli, monoton ya da parçalı sürekli her fonksiyon bu koşulu sağlar. İkincisi hesaba dönük: X 'in dağılımı biliniyorken $Y = g(X)$ 'in dağılımı nasıl bulunur? Kesikli durumda yanıt bir toplamdır. Yoğunluğu olan sürekli durumda iki yol vardır: her zaman işleyen **dağılım fonksiyonu tekniği** ve g parçalı monoton olduğunda doğrudan bir formül veren **dönüşüm teoremi**.

Yol boyunca kare, doğrusal, üstel, tam değer ve tanjant dönüşümlerini tek tek işleyeceğiz; sürekli bir değişkenin dönüşümünün kesikli çıkabileceğini, kesikli bir değişkenin dönüşümününse asla sürekli olamayacağını göreceğiz. Tanjant dönüşümü bizi olasılık teorisinin ünlü dağılımlarından birine, Cauchy dağılımına götürecektir.

12.1 Rastgele Değişkenin Fonksiyonu

Rastgele değişkenin tanımını hatırlayalım (**Tanım 8.1**): $X : \Omega \rightarrow \mathbb{R}$ fonksiyonu, her $a \in \mathbb{R}$ için $(X \leq a) = \{\omega \in \Omega : X(\omega) \leq a\}$ kümesi $\mathcal{U}$ 'da bir olay olduğunda rastgele değişkendir; bu koşul, her Borel kümesi $B \in \mathcal{B}(\mathbb{R})$ için $X^{-1}(B) \in \mathcal{U}$ olmasına denktir. Şimdi $g : \mathbb{R} \rightarrow \mathbb{R}$ herhangi bir fonksiyon olsun ve

$$Y = g \circ X : \Omega \rightarrow \mathbb{R}, \quad Y(\omega) = g(X(\omega))$$

bileşkesini alalım. Bir ω sonucu önce X ile $X(\omega)$ sayısına, sonra g ile $g(X(\omega))$ sayısına gider. Y 'nin rastgele değişken olması için $(Y \leq a)$ kümelerinin olay olması gerekir. Bunu X 'e değil, g 'ye bir koşul koyarak garanti edeceğiz: g , Borel kümelerinin ters görüntüsünü Borel kümesine götürsün. Neden bu koşul? Çünkü $(Y \leq a)$ kümesini X 'in bir ters görüntüsü olarak yazabiliyoruz ve X yalnızca Borel kümelerinin ters görüntüsünü olay yapmayı taahhüt eder.

Teorem 12.1 (Rastgele Değişkenin Fonksiyonu Rastgele Değişkendir). $X, (\Omega, \mathcal{U}, P)$ olasılık uzayında bir rastgele değişken ve $g : \mathbb{R} \rightarrow \mathbb{R}$, her Borel kümesi $B \in \mathcal{B}(\mathbb{R})$ için

$$g^{-1}(B) = \{x \in \mathbb{R} : g(x) \in B\} \in \mathcal{B}(\mathbb{R})$$

koşulunu sağlayan bir fonksiyon olsun (böyle bir g 'ye **Borel-ölçülebilir fonksiyon** denir). O zaman $Y = g \circ X = g(X)$ de $(\Omega, \mathcal{U}, P)$ uzayında bir rastgele değişkendir. Ayrıca Y 'nin olasılık dağılımı, her $B \in \mathcal{B}(\mathbb{R})$ için

$$P_Y(B) = P(Y \in B) = P(X \in g^{-1}(B)) = P_X(g^{-1}(B))$$

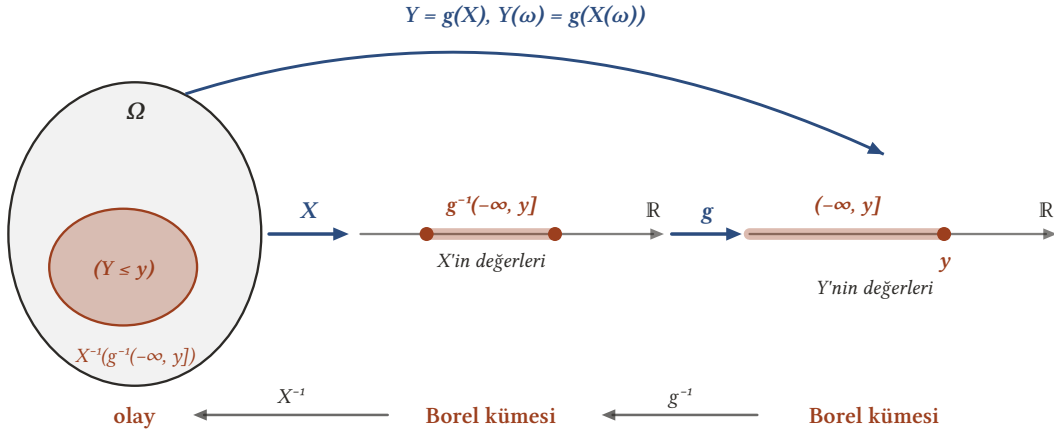
ile verilir.

İspat.

$a \in \mathbb{R}$ sabit olsun. $Y(\omega) \leq a$ olması $g(X(\omega)) \leq a$, yani $g(X(\omega)) \in (-\infty, a]$ demektir; bu da $X(\omega) \in g^{-1}((-\infty, a])$ ile aynı şeydir. Öyleyse

$$(Y \leq a) = \{\omega \in \Omega : g(X(\omega)) \leq a\} = \{\omega \in \Omega : X(\omega) \in g^{-1}((-\infty, a])\} = X^{-1}(g^{-1}((-\infty, a])).$$

$(-\infty, a]$ bir Borel kümesidir (Önerme 2.5). g üzerindeki koşul gereği $C = g^{-1}((-\infty, a])$ de bir Borel kümesidir. X rastgele değişken olduğundan $X^{-1}(C) \in \mathcal{U}$ olur. Böylece her $a \in \mathbb{R}$ için $(Y \leq a) \in \mathcal{U}$; Tanım 8.1 gereği Y bir rastgele değişkendir.



Şekil 12.1: Bileşke $Y = g \circ X$, bir sonucu önce X ile bir sayıya, sonra g ile ikinci bir sayıya götürür. $(Y \leq y)$ olayı sağdan sola geri çekilerek bulunur: $(-\infty, y]$ Borel kümesidir, g Borel-ölçülebilir olduğundan $g^{-1}((-\infty, y])$ de Borel kümesidir ve X rastgele değişken olduğundan onun ters görüntüsü bir olaydır. Koşulun X 'e değil g 'ye konmasının nedeni budur.

Aynı hesap $(-\infty, a]$ yerine herhangi bir $B \in \mathcal{B}(\mathbb{R})$ ile yapılırsa

$$(Y \in B) = \{\omega : g(X(\omega)) \in B\} = \{\omega : X(\omega) \in g^{-1}(B)\} = (X \in g^{-1}(B))$$

bulunur. İki tarafın olasılığı alınınca, olasılık dağılımının tanımıyla (Tanım 8.3), $P_Y(B) = P(Y \in B) = P(X \in g^{-1}(B)) = P_X(g^{-1}(B))$ elde edilir.

İ Hangi fonksiyonlar Borel-ölçülebilirdir?

Pratikte karşılaşılan her fonksiyon bu koşulu sağlar. Sürekli her g Borel-ölçülebilirdir: açık kümenin sürekli fonksiyon altındaki ters görüntüsü açıktır, açık kümeler Borel cebirini ürettiğinden (Tanım 2.4) koşul bütün Borel kümelerine taşınır. Monoton her g Borel-ölçülebilirdir: $g^{-1}((-\infty, a])$ her zaman bir aralık ya da boş kümedir. Sonlu ya da sayılabilir parçada sürekli fonksiyonlar, tam değer fonksiyonu $\lfloor x \rfloor$, bir aralığın gösterge fonksiyonu da bu sınıftadır. Borel-ölçülebilir fonksiyonların toplamı, çarpımı ve bileşkesi yine Borel-ölçülebilirdir. Bu kitapta kullanacağımız her g bu sınıftadır; bundan sonra koşulu her seferinde anmayacağız.

Teoremin ikinci kısmı bu bölümün bütün hesaplarının ortak ilkesidir; ona **ters görüntü ilkesi** diyelim: Y ile ilgili her olasılık, X ile ilgili bir olasılığa çevrilir; Y 'nin B 'ye düşmesi, X 'in $g^{-1}(B)$ 'ye düşmesidir. Bölümün geri kalanı bu tek cümlemin kesikli ve sürekli durumda somutlaştırılmasından ibarettir.

12.2 Kesikli Rastgele Değişkenin Dönüşümü

X kesikli olsun (Tanım 10.1): değer kümesi D_X sayılabilir, $P(X \in D_X) = 1$ ve olasılık fonksiyonu $f_X(x) = P(X = x)$, $x \in D_X$ (Tanım 10.2). Bu durumda g 'nin yalnızca D_X üzerinde tanımlı olması yeter. $Y = g(X)$ 'in y değerini alması, X 'in $g(x) = y$ olan x değerlerinden birini alması demektir; öyleyse Y 'nin olasılık fonksiyonu X 'in olasılık fonksiyonunun "aynı görüntüye giden değerler üzerinden toplanmasıyla" bulunur.

Önerme 12.1 (Kesikli Dönüşümün Olasılık Fonksiyonu). X , değer kümesi D_X ve olasılık fonksiyonu f_X olan kesikli bir rastgele değişken, $g : D_X \rightarrow \mathbb{R}$ herhangi bir fonksiyon olsun. O zaman $Y = g(X)$ kesikli bir rastgele değişkendir; değer kümesi

$$D_Y = g(D_X) = \{g(x) : x \in D_X\}$$

ve olasılık fonksiyonu

$$f_Y(y) = P(Y = y) = \sum_{x \in D_X : g(x)=y} f_X(x), \quad y \in D_Y$$

dir.

İspat.

Önce Y 'nin bir rastgele değişken olduğunu görelim. D_X sayılabilir olduğundan bir Borel kümesidir (Önerme 2.5). g 'yi D_X dışında 0 olarak genişletelim; her $B \in \mathcal{B}(\mathbb{R})$ için $g^{-1}(B)$, ya D_X 'in sayılabilir bir alt kümesidir ya da böyle bir kümeyle $\mathbb{R} \setminus D_X$ 'in birleşimidir; iki durumda da Borel kümesidir. Teorem 12.1 gereği Y rastgele değişkendir.

Şimdi $y \in \mathbb{R}$ sabit olsun ve $A_y = \{x \in D_X : g(x) = y\}$ yazalım. Her ω için $X(\omega) \in D_X$ olduğundan, $Y(\omega) = y$ olması $X(\omega) \in A_y$ olması demektir; yani

$$(Y = y) = (X \in A_y) = \bigcup_{x \in A_y} (X = x).$$

Sağdaki birleşim, $A_y \subseteq D_X$ sayılabilir olduğundan sayılabilir, farklı x 'ler için $(X = x)$ olayları ayrık olduğundan ayrık bir birleşimdir. Sayılabilir toplamsallıkla (Tanım 3.1)

$$P(Y = y) = \sum_{x \in A_y} P(X = x) = \sum_{x \in D_X : g(x)=y} f_X(x).$$

$y \notin D_Y$ ise $A_y = \emptyset$ ve $P(Y = y) = 0$ 'dır. Y 'nin değer kümesi $D_Y = Y(\Omega) = g(X(\Omega)) = g(D_X)$, sayılabilir bir kümenin görüntüsü olarak sayılabilir; öyleyse Y kesiklidir (Tanım 10.1) ve f_Y onun olasılık fonksiyonudur (Tanım 10.2). Sağlama olarak toplamın 1 olduğunu görelim: A_y kümeleri ($y \in D_Y$) D_X 'i parçaladığından

$$\sum_{y \in D_Y} f_Y(y) = \sum_{y \in D_Y} \sum_{x \in A_y} f_X(x) = \sum_{x \in D_X} f_X(x) = 1,$$

yani $P(Y \in D_Y) = 1$; bir olasılık fonksiyonundan beklenen budur. ■

Toplamın anlamı şudur: g , D_X üzerinde birebir ise her $y \in D_Y$ için toplamda tek terim vardır ve $f_Y(g(x)) = f_X(x)$ olur; tablo yalnızca yeniden etiketlenir. g birebir değilse aynı görüntüye giden değerlerin olasılıkları toplanır. İkisini de aynı örnekte görelim.

Örnek 12.1 (Kesikli Bir Değişkenin Karesi, Doğrusal Dönüşümü ve Küpü). X rastgele değişkeninin olasılık fonksiyonu

$$f_X(x) = \frac{1}{5}, \quad x \in \{-2, -1, 0, 1, 2\}$$

olsun. $Y = X^2$, $Z = 3X + 1$ ve $W = X^3$ rastgele değişkenlerinin olasılık fonksiyonlarını bulunuz.

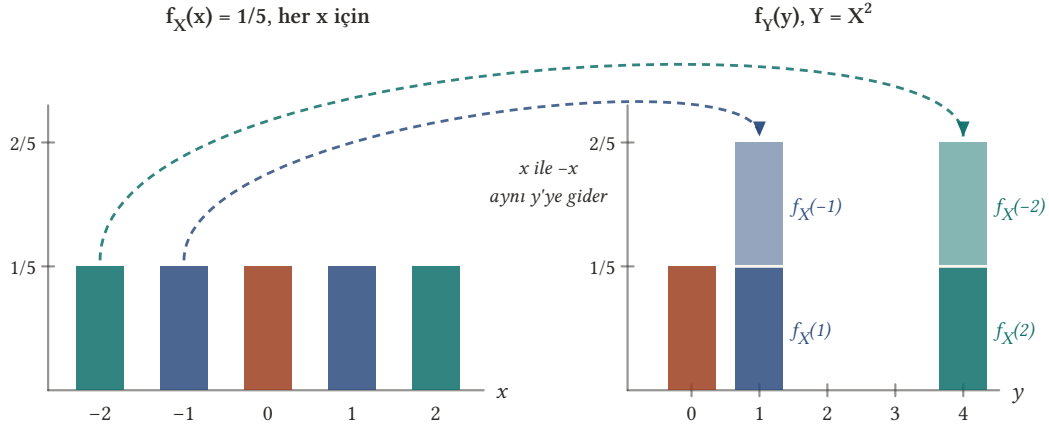
Çözüm.

$Y = X^2$. $g(x) = x^2$ fonksiyonu D_X 'i $\{0, 1, 4\}$ kümesine götürür; $D_Y = \{0, 1, 4\}$. Önerme 12.1 ile

$$f_Y(0) = f_X(0) = \frac{1}{5}, \quad f_Y(1) = f_X(-1) + f_X(1) = \frac{2}{5}, \quad f_Y(4) = f_X(-2) + f_X(2) = \frac{2}{5}.$$

y	0	1	4
$f_Y(y) = P(Y = y)$	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{2}{5}$

Toplam $\frac{1}{5} + \frac{2}{5} + \frac{2}{5} = 1$ 'dir. 1 ve 4 değerlerinin olasılığı ikiye katlanmıştır, çünkü kare fonksiyonu ± 1 'i ve ± 2 'yi aynı noktaya götürür.



Şekil 12.2: $X, \{-2, -1, 0, 1, 2\}$ üzerinde her değeri $1/5$ olasılıkla alırken $Y = X^2$ 'nin olasılık fonksiyonu. Kare fonksiyonu x ile $-x$ 'i aynı noktaya götürdüğünden negatif değerlerin olasılığı pozitif olanların üstüne katlanır: $f_Y(1) = f_X(-1) + f_X(1) = 2/5$ ve $f_Y(4) = 2/5$. Yalnızca 0'ın tek ters görüntüsü vardır; $f_Y(0) = 1/5$ olarak kalır ve toplam yine 1'dir.

$Z = 3X + 1$. $g(x) = 3x + 1$ birebirdir; $D_Z = \{-5, -2, 1, 4, 7\}$ ve her $z \in D_Z$ için $f_Z(z) = f_X\left(\frac{z-1}{3}\right) = \frac{1}{5}$:

z	-5	-2	1	4	7
$f_Z(z) = P(Z = z)$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$

$W = X^3$. $g(x) = x^3$ de D_X üzerinde birebirdir (kare fonksiyonundan farklı olarak işareti korur); $D_W = \{-8, -1, 0, 1, 8\}$ ve her $w \in D_W$ için $f_W(w) = \frac{1}{5}$. Doğrusal dönüşümde olduğu gibi tablo yalnızca yeniden etiketlenmiştir.

■

12.3 Dağılım Fonksiyonu Tekniği

X , olasılık yoğunluk fonksiyonu f_X olan sürekli bir rastgele değişken olsun (Tanım 11.1, Tanım 11.2). Bu durumda her x için $P(X = x) = 0$ 'dır (Önerme 11.1); kesikli durumdaki toplam formülü işe yaramaz, çünkü toplanacak her terim sıfırdır. Yoğunluk fonksiyonu noktaların değil aralıkların olasılığını verir; öyleyse Y 'nin dağılımına aralık olasılıkları üzerinden, yani dağılım fonksiyonu üzerinden ulaşmalıyız. Ters görüntü ilkesi burada şu biçimi alır:

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \in \{x : g(x) \leq y\}).$$

Sağ taraf X 'e ilişkin bir olasılıktır ve f_X 'in $\{x : g(x) \leq y\}$ kümesi üzerindeki integraliyle hesaplanır. F_Y bulunduktan sonra türev alarak f_Y elde edilir (Teorem 11.2).

Tanım 12.1 (Dağılım Fonksiyonu Tekniği). X , olasılık yoğunluk fonksiyonu f_X olan sürekli bir rastgele değişken, $D_X \subseteq A$ ve $g : A \rightarrow \mathbb{R}$ olsun. $Y = g(X)$ 'in dağılımını aşağıdaki üç adımla bulmaya **dağılım fonksiyonu tekniği** denir.

1. **Değer kümesi.** $D_Y = g(D_X)$ belirlenir; D_Y alttan ya da üstten sınırlıysa F_Y , bu kümenin solunda 0, sağında 1 değerini alır.

2. **Dağılım fonksiyonu.** Her $y \in D_Y$ için

$$F_Y(y) = P(g(X) \leq y) = \int_{\{x: g(x) \leq y\}} f_X(x) dx$$

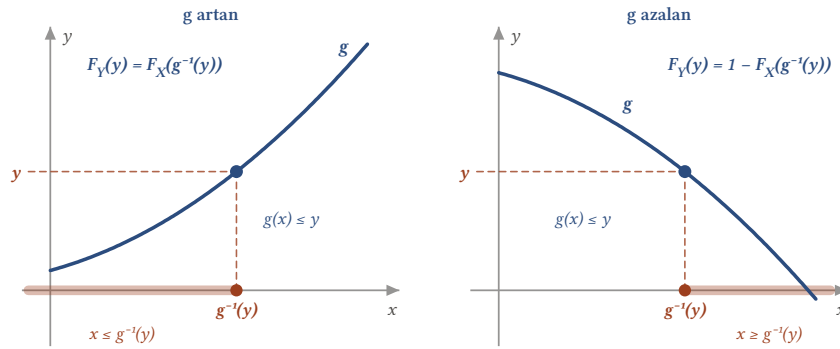
hesaplanır; bunun için $g(x) \leq y$ eşitsizliği x için çözülür.

3. **Türev.** F_Y sürekli ve sonlu sayıda nokta dışında türevlenebilir çıkarsa Y 'nin bir olasılık yoğunluk fonksiyonu vardır ve

$$f_Y(y) = F_Y'(y), \quad y \in D_Y$$

(diğer durumlarda $f_Y(y) = 0$) alınır. F_Y sıçrama yapıyorsa Y 'nin yoğunluğu yoktur; sıçrama noktaları Y 'nin pozitif olasılıklı değerleridir.

Tekniğin özü ikinci adımdaki eşitsizliği çözmektir. g artan ise $g(x) \leq y \Leftrightarrow x \leq g^{-1}(y)$ ve $F_Y(y) = F_X(g^{-1}(y))$; g azalan ise $g(x) \leq y \Leftrightarrow x \geq g^{-1}(y)$ ve $F_Y(y) = 1 - F_X(g^{-1}(y))$ olur (F_X sürekli olduğundan uç noktanın dahil olup olmaması fark etmez). Birebir olmayan g 'de ters görüntü birden çok parçadan oluşur; kare dönüşümünde bunu göreceğiz.



Şekil 12.3: Dağılım fonksiyonu tekniğinin özü $g(x) \leq y$ eşitsizliğini x için çözmektir. g artan ise çözüm kümesi $g^{-1}(y)$ 'nin solundaki yarı doğrudur ve $F_Y(y) = F_X(g^{-1}(y))$; g azalan ise sağdaki yarı doğrudur ve $F_Y(y) = 1 - F_X(g^{-1}(y))$. F_X sürekli olduğundan uç noktanın dahil olup olmaması sonucu değiştirmez.

Örnek 12.2 (Doğrusal Dönüşüm). X rastgele değişkeninin olasılık yoğunluk fonksiyonu

$$f_X(x) = \begin{cases} 2x, & 0 < x < 1 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun. $Y = 3X + 1$ rastgele değişkeninin olasılık yoğunluk fonksiyonunu bulunuz.

Çözüm.

Değer kümesi. $0 < x < 1$ iken $1 < 3x + 1 < 4$; öyleyse $D_Y = (1, 4)$. $y \leq 1$ için $F_Y(y) = 0$, $y \geq 4$ için $F_Y(y) = 1$ 'dir.

Dağılım fonksiyonu. $1 < y < 4$ için $3X + 1 \leq y$ eşitsizliği $X \leq \frac{y-1}{3}$ ile aynıdır ve $0 < \frac{y-1}{3} < 1$ olduğundan

$$F_Y(y) = P(3X + 1 \leq y) = P\left(X \leq \frac{y-1}{3}\right) = \int_0^{\frac{y-1}{3}} 2x dx = \left[x^2\right]_0^{\frac{y-1}{3}} = \frac{(y-1)^2}{9}.$$

Türev. $1 < y < 4$ için $f_Y(y) = F_Y'(y) = \frac{2(y-1)}{9}$. Böylece

$$f_Y(y) = \begin{cases} \frac{2(y-1)}{9}, & 1 < y < 4 \\ 0, & \text{diğer durumlarda.} \end{cases}$$

Sağlama: $\int_1^4 \frac{2(y-1)}{9} dy = \left[\frac{(y-1)^2}{9} \right]_1^4 = 1$.

Aynı hesap genel olarak $a > 0$ için $Y = aX + b$ dönüşümüne $f_Y(y) = \frac{1}{a} f_X\left(\frac{y-b}{a}\right)$ verir: yoğunluğun grafiği a kat genişler, b kadar kayar ve alanı 1'de tutmak için $\frac{1}{a}$ ile çarpılır.

■

Örnek 12.3 (Üstel Dönüşüm). X , önceki örnekteki gibi $f_X(x) = 2x$ ($0 < x < 1$) yoğunluklu rastgele değişken olsun. $Y = e^{-X}$ rastgele değişkeninin olasılık yoğunluk fonksiyonunu bulunuz.

Çözüm.

Değer kümesi. $x \mapsto e^{-x}$ azalandır; $0 < x < 1$ iken $e^{-1} < e^{-x} < 1$. Öyleyse $D_Y = (e^{-1}, 1)$; $y \leq e^{-1}$ için $F_Y(y) = 0$, $y \geq 1$ için $F_Y(y) = 1$.

Dağılım fonksiyonu. $e^{-1} < y < 1$ olsun. Her iki tarafın logaritması alınırsa (logaritma artan olduğundan eşitsizliğin yönü korunur) $e^{-X} \leq y \Leftrightarrow -X \leq \ln y \Leftrightarrow X \geq -\ln y$ bulunur; ayrıca $0 < -\ln y < 1$ 'dir. Buna göre

$$F_Y(y) = P(e^{-X} \leq y) = P(X \geq -\ln y) = \int_{-\ln y}^1 2x dx = \left[x^2 \right]_{-\ln y}^1 = 1 - (\ln y)^2.$$

Uçlarda süreklilik kontrolü: $y \rightarrow e^{-1}$ iken $1 - (-1)^2 = 0$, $y \rightarrow 1$ iken $1 - 0 = 1$; beklendiği gibi.

Türev. $e^{-1} < y < 1$ için

$$f_Y(y) = F_Y'(y) = -2 \ln y \cdot \frac{1}{y} = -\frac{2 \ln y}{y}.$$

Bu aralıkta $\ln y < 0$ olduğundan $f_Y(y) > 0$ 'dır. Böylece

$$f_Y(y) = \begin{cases} -\frac{2 \ln y}{y}, & e^{-1} < y < 1 \\ 0, & \text{diğer durumlarda.} \end{cases}$$

Sağlama: $u = \ln y$ dönüşümüyle $\int_{e^{-1}}^1 -\frac{2 \ln y}{y} dy = \int_{-1}^0 -2u du = \left[-u^2 \right]_{-1}^0 = 1$.

■

Şimdi birebir olmayan bir dönüşüme geçelim. Kare fonksiyonu x ile $-x$ 'i aynı noktaya götürür; $Y = X^2$ 'nin dağılımına ters görüntünün iki dalından, negatif ve pozitif karekök dallarından katkı gelir.

Örnek 12.4 (Simetrik Aralıkta Kare Dönüşümü). X rastgele değişkeninin olasılık yoğunluk fonksiyonu

$$f_X(x) = \begin{cases} \frac{1}{4}, & -2 \leq x \leq 2 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun. $Y = X^2$ rastgele değişkeninin olasılık yoğunluk fonksiyonunu bulunuz.

Çözüm.

Değer kümesi. $-2 \leq x \leq 2$ iken $0 \leq x^2 \leq 4$; $D_Y = [0, 4]$.

Dağılım fonksiyonu. $0 < y < 4$ için $X^2 \leq y \Leftrightarrow -\sqrt{y} \leq X \leq \sqrt{y}$; ters görüntü tek bir aralıktır ama iki daldan gelir. $[-\sqrt{y}, \sqrt{y}] \subseteq [-2, 2]$ olduğundan

$$F_Y(y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{4} dx = \frac{2\sqrt{y}}{4} = \frac{\sqrt{y}}{2}.$$

$y \leq 0$ için $F_Y(y) = 0$, $y \geq 4$ için $F_Y(y) = 1$ 'dir; $y = 4$ 'te $\frac{\sqrt{4}}{2} = 1$ ile uyumludur.

Türev. $0 < y < 4$ için $f_Y(y) = \frac{d}{dy} \frac{\sqrt{y}}{2} = \frac{1}{4\sqrt{y}}$. Böylece

$$f_Y(y) = \begin{cases} \frac{1}{4\sqrt{y}}, & 0 < y < 4 \\ 0, & \text{diğer durumlarda.} \end{cases}$$

Sağlama: $\int_0^4 \frac{1}{4\sqrt{y}} dy = \left[\frac{\sqrt{y}}{2} \right]_0^4 = 1$. Yoğunluk $y \rightarrow 0^+$ iken sınırsız büyür; bir olasılık yoğunluk fonksiyonu sınırlı olmak zorunda değildir, integralinin 1 olması yeter. Sınırsızlığın nedeni, kare fonksiyonunun 0 yakınında x eksenindeki geniş bir aralığı y ekseninde çok dar bir aralığa sıkıştırmasıdır.

Simetrik aralıkta $\pm\sqrt{y}$ dalları eşit katkı verdi. Aralık simetrik değilse büyük y değerleri için dallardan biri tanım bölgesinin dışına düşer ve toplama yalnızca bir dal katkı verir; aşağıdaki gibi düzgün bir yoğunlukta bu, yoğunluğun yarıya inmesi demektir.

Örnek 12.5 (Simetrik Olmayan Aralıkta Kare Dönüşümü). X rastgele değişkeninin olasılık yoğunluk fonksiyonu

$$f_X(x) = \begin{cases} \frac{1}{3}, & -1 \leq x \leq 2 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun. $Y = X^2$ rastgele değişkeninin olasılık yoğunluk fonksiyonunu bulunuz.

Çözüm.

Değer kümesi. $-1 \leq x \leq 2$ iken $0 \leq x^2 \leq 4$; $D_Y = [0, 4]$.

Dağılım fonksiyonu. $0 < y < 4$ için yine $X^2 \leq y \Leftrightarrow -\sqrt{y} \leq X \leq \sqrt{y}$, ama bu aralığın $[-1, 2]$ ile kesişimi y 'ye bağlıdır.

Durum $0 < y \leq 1$: $-1 \leq -\sqrt{y}$ olduğundan $[-\sqrt{y}, \sqrt{y}] \subseteq [-1, 2]$ ve

$$F_Y(y) = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{3} dx = \frac{2\sqrt{y}}{3}.$$

Durum $1 < y < 4$: $-\sqrt{y} < -1$ olduğundan $X \geq -\sqrt{y}$ koşulu kendiliğinden sağlanır; $(X^2 \leq y) = (-1 \leq X \leq \sqrt{y})$ ve

$$F_Y(y) = \int_{-1}^{\sqrt{y}} \frac{1}{3} dx = \frac{\sqrt{y} + 1}{3}.$$

Süreklilik kontrolü: $y = 1$ 'de iki ifade de $\frac{2}{3}$ verir; $y \rightarrow 4$ iken $\frac{2+1}{3} = 1$.

Türev. Parçalara ayrı ayrı türev alınır:

$$f_Y(y) = \begin{cases} \frac{1}{3\sqrt{y}}, & 0 < y < 1 \\ \frac{1}{6\sqrt{y}}, & 1 < y < 4 \\ 0, & \text{diğer durumlarda.} \end{cases}$$

($F_Y, y = 1$ 'de sürekli ama türevlenemez; sol türev $\frac{1}{3}$, sağ türev $\frac{1}{6}$ 'dır. Tek bir noktadaki değer yoğunluğu değiştirmez.) Sağlama:

$$\int_0^1 \frac{1}{3\sqrt{y}} dy + \int_1^4 \frac{1}{6\sqrt{y}} dy = \left[\frac{2\sqrt{y}}{3} \right]_0^1 + \left[\frac{\sqrt{y}}{3} \right]_1^4 = \frac{2}{3} + \frac{2-1}{3} = 1.$$

$1 < y < 4$ için yoğunluğun yarıya inmesinin nedeni açıktır: X hiçbir zaman -1 'in altına düşmediğinden $x = -\sqrt{y}$ dalı katkı vermez, yalnızca $x = \sqrt{y}$ dalı kalır.

12.4 Dönüşüm Kesikli mi, Sürekli mi?

Buraya kadar kesikli X 'in dönüşümü kesikli çıktı; yoğunluğu olan X 'in dönüşümünün de yoğunluğu çıktı. İlki her zaman doğrudur; ikincisi değildir. Bu farkı açıkça kaydedelim. Bunun için kesikli (değer kümesi sayılabilir) ile sürekli (değer kümesi sayılamaz) ayrımını, yoğunluğun var olup olmamasından ayrı tutmak gerekir.

Önerme 12.2 (Dönüşümün Türü). (i) Kesikli bir rastgele değişkenin herhangi bir fonksiyonu yine kesiklidir; kesikli bir değişkenin dönüşümü hiçbir zaman sürekli bir rastgele değişken vermez. (ii) Yoğunluğu olan bir rastgele değişkenin fonksiyonu yoğunluğu olan bir değişken olabilir, kesikli olabilir ya da değer kümesi sayılamaz olduğu hâlde yoğunluğu olmayabilir.

İspat.

(i) **Önerme 12.1**'da gösterildi: $D_Y = g(D_X)$ sayılabilir, dolayısıyla Y kesiklidir. Sürekli bir rastgele değişkenin değer kümesi sayılamaz olduğundan (**Tanım 11.1**) Y sürekli olamaz. Daha fazlası da söylenebilir: Y 'nin yoğunluğu da olamaz. Yoğunluğu olan bir değişken için sayılabilir bir kümenin olasılığı, tek noktali olayların olasılıklarının toplamı olarak 0'dır (**Önerme 11.1**); oysa $P(Y \in D_Y) = 1$ 'dir. (ii) Her üç durum için birer örnek yeter. Önceki örneklerde yoğunluğu olan X 'in dönüşümlerinin yoğunluğu çıktı. Aşağıdaki **Örnek 12.6**'nda yoğunluğu olan X 'in tam değeri kesikli çıkacak. Üçüncü durum için X , $(-1, 1)$ üzerinde $f_X(x) = \frac{1}{2}$ yoğunluklu olsun ve $Y = \max(X, 0)$ alınsın. Y 'nin değer kümesi $[0, 1)$ sayılamaz olduğundan Y sürekli, kesikli değildir. Öte yandan $P(Y = 0) = P(X \leq 0) = \frac{1}{2}$ 'dir; yoğunluğu olan bir değişkende her noktanın olasılığı 0 olduğundan (**Önerme 11.1**) Y 'nin yoğunluğu yoktur. Dağılım fonksiyonu bunu açıkça gösterir: $F_Y(y) = 0$ ($y < 0$), $F_Y(y) = P(X \leq y) = \frac{y+1}{2}$ ($0 \leq y < 1$), $F_Y(y) = 1$ ($y \geq 1$); F_Y , 0'da $\frac{1}{2}$ kadar sıçrar, sonra eğik bir doğru parçası boyunca artar.

▲ Sürekli değişkenin dönüşümü için türü baştan varsaymayın

Dağılım fonksiyonu tekniğinin üçüncü adımı bu yüzden “türev al” değil, “önce F_Y 'ye bak” der. F_Y sürekli ve parçalı türevlenebilir ise Y 'nin yoğunluğu vardır ve türev onu verir; F_Y basamaklı ise Y kesiklidir ve sıçramalar olasılık fonksiyonunu verir; ikisi karışıksa Y kesikli değildir ama yoğunluğu da yoktur. g 'nin bir aralığı tek bir noktaya çökertmesi (sabit kalması) kesikliliğin tipik kaynağıdır.

Örnek 12.6 (Tam Değer Dönüşümü). X rastgele değişkeninin olasılık yoğunluk fonksiyonu

$$f_X(x) = \begin{cases} \frac{1}{5}, & 0 \leq x \leq 5 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun. $Y = \lfloor X \rfloor$ (X 'in tam değeri, yani X 'i aşmayan en büyük tam sayı) rastgele değişkeninin dağılımını bulunuz.

Çözüm.

X , 1 olasılıkla $[0, 5]$ içinde değer aldığından Y 'nin alabileceği değerler 0, 1, 2, 3, 4, 5 tam sayılarıdır; Y kesikli bir rastgele değişkendir. Her k için $\lfloor X \rfloor = k \Leftrightarrow k \leq X < k+1$ olduğundan, $k = 0, 1, 2, 3, 4$ için

$$P(Y = k) = P(k \leq X < k+1) = \int_k^{k+1} \frac{1}{5} dx = \frac{1}{5}.$$

$k = 5$ için ise $(5 \leq X < 6) = (X = 5)$ ve X 'in yoğunluğu olduğundan (**Önerme 11.1**)

$$P(Y = 5) = P(X = 5) = 0.$$

Demek ki 5 değeri olasılık taşımaz; olasılık taşıyan değerler kümesi $\{0, 1, 2, 3, 4\}$ ve olasılık fonksiyonu

$$f_Y(y) = P(Y = y) = \frac{1}{5}, \quad y \in \{0, 1, 2, 3, 4\}$$

dir. Toplam $5 \cdot \frac{1}{5} = 1$ 'dir. Y 'nin dağılım fonksiyonu, 0, 1, 2, 3, 4 noktalarının her birinde $\frac{1}{5}$ sıçrayan bir basamak fonksiyonudur (Önerme 10.2): $F_Y(y) = 0$ ($y < 0$), $F_Y(y) = \frac{k+1}{5}$ ($k \leq y < k+1$, $k = 0, \dots, 3$), $F_Y(y) = 1$ ($y \geq 4$). Sürekli bir değişkenden kesikli bir değişken elde ettik: tam değer fonksiyonu her $[k, k+1)$ aralığını tek bir noktaya çökertir ve o aralığın olasılığı o noktada birikir.

■

12.5 Dönüşüm Teoremi

Dağılım fonksiyonu tekniği her zaman işler ama her seferinde $g(x) \leq y$ eşitsizliğini çözmeyi, integrali kurmayı ve türev almayı gerektirir. Örneklere geri bakınca hesabın hep aynı iskelete oturduğu görülür: g monoton parçalara ayrılır, her parçada ters fonksiyon alınır, zincir kuralıyla türev, $f_X(g_i^{-1}(y))$ ile ters fonksiyonun türevinin çarpımı olarak çıkar. Bu iskeleti bir kez ispatlayıp sonra doğrudan formül olarak kullanmak çok zaman kazandırır. Formülü önce sezgisel olarak görelim: g artan ve türevlenebilir olsun. y ile $y + dy$ arasına düşme olasılığı $f_Y(y) dy$ 'dir; bu olay, X 'in ters görüntü aralığı $[x, x + dx]$ 'e düşmesiyle aynıdır ($x = g^{-1}(y)$, $dx = (g^{-1})'(y) dy$), olasılığı $f_X(x) dx$ 'tir. Öyleyse $f_Y(y) dy = f_X(x) dx$, yani $f_Y(y) = f_X(g^{-1}(y)) \frac{dx}{dy}$. Azalan g 'de dx negatif çıkar, mutlak değer gerekir. Birebir olmayan g 'de her monoton parça kendi katkısını getirir ve katkılar toplanır.

Teorem 12.2 (Dönüşüm Teoremi). X , olasılık yoğunluk fonksiyonu f_X olan sürekli bir rastgele değişken olsun; f_X , (a, b) aralığının dışında sıfır ve (a, b) üzerinde sonlu sayıda noktada sürekli olsun. $(a_1, b_1), (a_2, b_2), \dots, (a_k, b_k)$ ikiye ikiye ayrık açık aralıkları (a, b) 'yi sonlu sayıda nokta dışında örtsün. $g : (a, b) \rightarrow \mathbb{R}$ fonksiyonunun her (a_i, b_i) aralığına kısıtlanması g_i için şunlar sağlansın:

- g_i sürekli ve kesin monotondur (artan ya da azalan); dolayısıyla (a_i, b_i) 'yi bir (c_i, d_i) açık aralığına birebir ve örten biçimde götürür,
- ters fonksiyonu $g_i^{-1} : (c_i, d_i) \rightarrow (a_i, b_i)$ sürekli türevlenebilirdir.

O zaman $Y = g(X)$ 'in bir olasılık yoğunluk fonksiyonu vardır ve bu fonksiyon

$$f_Y(y) = \sum_{i: c_i < y < d_i} f_X(g_i^{-1}(y)) \left| \frac{dg_i^{-1}(y)}{dy} \right|$$

dir; toplam, y 'yi görüntü aralığında bulunduran i 'ler üzerindendir ve hiçbir (c_i, d_i) 'ye düşmeyen y için $f_Y(y) = 0$ 'dir. Özel olarak bütün g_i 'ler aynı (c, d) aralığına örten ise, $c < y < d$ için

$$f_Y(y) = \sum_{i=1}^k f_X(g_i^{-1}(y)) \left| \frac{dg_i^{-1}(y)}{dy} \right|$$

olur.

İspat.

İspat, her parçaya dağılım fonksiyonu tekniğini uygulayıp zincir kuralıyla türev almaktan ibarettir.

Adım 1: F_Y 'yi parçalara ayırma. N , (a, b) 'nin $\bigcup_i (a_i, b_i)$ dışında kalan sonlu sayıdaki noktasının kümesi olsun. X 'in yoğunluğu olduğundan $P(X \in N) = 0$ (Önerme 11.1) ve f_X , (a, b) dışında sıfır olduğundan $P(X \notin (a, b)) = \int_{\mathbb{R} \setminus (a, b)} f_X(x) dx = 0$ 'dir. $(a_i < X < b_i)$ olayları ayrıktır ve olasılıkları toplamı 1'dir. Öyleyse her $y \in \mathbb{R}$ için

$$F_Y(y) = P(g(X) \leq y) = \sum_{i=1}^k P(a_i < X < b_i, g_i(X) \leq y) = \sum_{i=1}^k H_i(y), \quad H_i(y) = P(a_i < X < b_i, g_i(X) \leq y).$$

Adım 2: Her H_i 'nin hesabı. g_i 'nin görüntüsü (c_i, d_i) olduğundan üç durum vardır.

- $y \leq c_i$ ise $g_i(X) \leq y$ olayı imkânsızdır: $H_i(y) = 0$.
- $y \geq d_i$ ise $g_i(X) \leq y$ kesindir: $H_i(y) = P(a_i < X < b_i) = F_X(b_i) - F_X(a_i)$.
- $c_i < y < d_i$ ise ters fonksiyon devreye girer. g_i artan olduğunda $x \in (a_i, b_i)$ için $g_i(x) \leq y \Leftrightarrow x \leq g_i^{-1}(y)$; Teorem 9.2 ile

$$H_i(y) = P(a_i < X \leq g_i^{-1}(y)) = F_X(g_i^{-1}(y)) - F_X(a_i).$$

g_i azalan olduğunda $g_i(x) \leq y \Leftrightarrow x \geq g_i^{-1}(y)$; sürekli değişkende uç noktalar olasılık taşımadığından

$$H_i(y) = P(g_i^{-1}(y) \leq X < b_i) = F_X(b_i) - F_X(g_i^{-1}(y)).$$

Adım 3: F_Y sürekli. X 'in yoğunluğu olduğundan F_X sürekli (Teorem 11.1), g_i^{-1} sürekli ve $y \rightarrow c_i^+$ ya da $y \rightarrow d_i^-$ iken $g_i^{-1}(y)$ ilgili uca yaklaşır (g_i artarsa c_i^+ için a_i 'ye, d_i^- için b_i 'ye; azalansa tersi). Bu yüzden Adım 2'deki üç parça c_i ve d_i 'de birbirine eklenir ve H_i bütün $\mathbb{R}$ üzerinde sürekli. Sonlu toplam olan F_Y de sürekli; demek ki Y 'nin pozitif olasılıklı noktası yoktur.

Adım 4: Zincir kuralı. $c_i < y < d_i$ olsun ve $f_X, g_i^{-1}(y)$ noktasında sürekli olsun (bu koşul sonlu sayıda y dışında sağlanır, çünkü f_X 'in süreksizlik noktaları sonludur ve g_i^{-1} birebirdir). Teorem 11.2 gereği $F_Y'(g_i^{-1}(y)) = f_X(g_i^{-1}(y))$ ve zincir kuralıyla, artan durumda

$$H_i'(y) = f_X(g_i^{-1}(y)) \frac{dg_i^{-1}(y)}{dy}, \quad \text{azalan durumda} \quad H_i'(y) = -f_X(g_i^{-1}(y)) \frac{dg_i^{-1}(y)}{dy}.$$

Artan fonksiyonun tersi artan, azalanın tersi azalan olduğundan ilk durumda $\frac{dg_i^{-1}}{dy} \geq 0$, ikincide $\frac{dg_i^{-1}}{dy} \leq 0$ 'dır. İki durumda da

$$H_i'(y) = f_X(g_i^{-1}(y)) \left| \frac{dg_i^{-1}(y)}{dy} \right|.$$

$y \notin [c_i, d_i]$ için H_i sabittir ve türevi 0'dır.

Adım 5: Toplama. Adım 4'ü i üzerinden toplayınca, sonlu sayıda y dışında

$$F_Y'(y) = \sum_{i: c_i < y < d_i} f_X(g_i^{-1}(y)) \left| \frac{dg_i^{-1}(y)}{dy} \right|$$

bulunur. F_Y sürekli, sonlu sayıda nokta dışında türevlenebilir ve türevi sonlu sayıda nokta dışında sürekli olduğundan (her terim, sürekli fonksiyonların çarpımıdır) $F_Y(y) = \int_{-\infty}^y F_Y'(t) dt$ sağlanır. Ayrıca $F_Y' \geq 0$ 'dır ve $\int_{-\infty}^{\infty} F_Y'(t) dt = \lim_{y \rightarrow \infty} F_Y(y) = 1$ 'dir. Demek ki $f_Y = F_Y', Y$ 'nin bir olasılık yoğunluk fonksiyonudur (Tanım 11.2). Hiçbir (c_i, d_i) 'ye düşmeyen y 'de bütün H_i 'ler sabit olduğundan $f_Y(y) = 0$ 'dır. ■

💡 Monoton dönüşüm: $k = 1$ durumu

$g, D_X = (a, b)$ üzerinde kesin monoton ve g^{-1} sürekli türevlenebilir ise, $y \in g((a, b))$ için

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{dg^{-1}(y)}{dy} \right| = \frac{f_X(x)}{|g'(x)|}, \quad x = g^{-1}(y)$$

olur; ikinci eşitlik $g'(x) \neq 0$ olduğunda ters fonksiyonun türev kuralından gelir. Doğrusal ve üstel dönüşüm örnekleri bu formülle tek satırda çözülür. **Örnek 12.2'**de $g^{-1}(y) = \frac{y-1}{3}$, türevi $\frac{1}{3}$: $f_Y(y) = 2 \cdot \frac{y-1}{3} \cdot \frac{1}{3} = \frac{2(y-1)}{9}$. **Örnek 12.3'**de $g^{-1}(y) = -\ln y$, türevi $-\frac{1}{y}$, mutlak değeri $\frac{1}{y}$: $f_Y(y) = 2(-\ln y) \cdot \frac{1}{y} = -\frac{2 \ln y}{y}$. Her iki sonuç da dağılım fonksiyonu tekniğiyle bulduğumuzla aynıdır.

Formülün asıl gücü birebir olmayan dönüşümlerde ortaya çıkar. Kare dönüşümünü şimdi formülle yeniden ele alalım; simetrik durumda sonucun **Örnek 12.4** ile örtüştüğünü göreceğiz, sonra aynı tekniği daha geniş bir aralığa uygulayacağız.

Örnek 12.7 (Kare Dönüşümü Dönüşüm Teoremiyle). (a) X 'in olasılık yoğunluk fonksiyonu $f_X(x) = \frac{1}{4} (-2 < x < 2; \text{diğer durumlarda } 0)$ olsun. $Y = X^2$ 'nin olasılık yoğunluk fonksiyonunu Teorem 12.2 ile bulunuz.

(b) X 'in olasılık yoğunluk fonksiyonu $f_X(x) = \frac{1}{5} (-2 < x < 3; \text{diğer durumlarda } 0)$ olsun. $Y = X^2$ 'nin olasılık yoğunluk fonksiyonunu bulunuz.

Çözüm.

(a) $g(x) = x^2$, $(-2, 2)$ üzerinde monoton değildir; ama 0 noktası çıkarılınca $(-2, 0)$ üzerinde azalan, $(0, 2)$ üzerinde artandır ($P(X = 0) = 0$ olduğundan noktanın çıkarılması bir şey değiştirmez). Parçaları ve terslerini yazalım:

$$g_1 : (-2, 0) \rightarrow (0, 4), g_1(x) = x^2, \quad g_1^{-1} : (0, 4) \rightarrow (-2, 0), g_1^{-1}(y) = -\sqrt{y}, \quad \frac{dg_1^{-1}(y)}{dy} = -\frac{1}{2\sqrt{y}};$$

$$g_2 : (0, 2) \rightarrow (0, 4), g_2(x) = x^2, \quad g_2^{-1} : (0, 4) \rightarrow (0, 2), g_2^{-1}(y) = \sqrt{y}, \quad \frac{dg_2^{-1}(y)}{dy} = \frac{1}{2\sqrt{y}}.$$

İki parça da aynı $(0, 4)$ aralığına örtendir; $0 < y < 4$ için

$$f_Y(y) = f_X(-\sqrt{y}) \left| -\frac{1}{2\sqrt{y}} \right| + f_X(\sqrt{y}) \left| \frac{1}{2\sqrt{y}} \right| = \frac{1}{4} \cdot \frac{1}{2\sqrt{y}} + \frac{1}{4} \cdot \frac{1}{2\sqrt{y}} = \frac{1}{4\sqrt{y}}.$$

Böylece

$$f_Y(y) = \begin{cases} \frac{1}{4\sqrt{y}}, & 0 < y < 4 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olur; **Örnek 12.4** ile aynıdır, ama integral kurmadan ve türev almadan bulunmuştur.

(b) Şimdi $(-2, 3)$ aralığı 0 noktasıyla $(-2, 0)$ ve $(0, 3)$ parçalarına ayrılır:

$$g_1 : (-2, 0) \rightarrow (0, 4), g_1^{-1}(y) = -\sqrt{y}; \quad g_2 : (0, 3) \rightarrow (0, 9), g_2^{-1}(y) = \sqrt{y}.$$

Görüntü aralıkları farklıdır: $(c_1, d_1) = (0, 4)$, $(c_2, d_2) = (0, 9)$. Formüldeki toplam, y 'yi içeren görüntü aralıkları üzerindendir.

$0 < y < 4$: iki parça da katkı verir:

$$f_Y(y) = \frac{1}{5} \cdot \frac{1}{2\sqrt{y}} + \frac{1}{5} \cdot \frac{1}{2\sqrt{y}} = \frac{1}{5\sqrt{y}}.$$

$4 < y < 9$: yalnızca g_2 katkı verir, çünkü $-\sqrt{y} < -2$ artık X 'in tanım bölgesinin dışındadır:

$$f_Y(y) = \frac{1}{5} \cdot \frac{1}{2\sqrt{y}} = \frac{1}{10\sqrt{y}}.$$

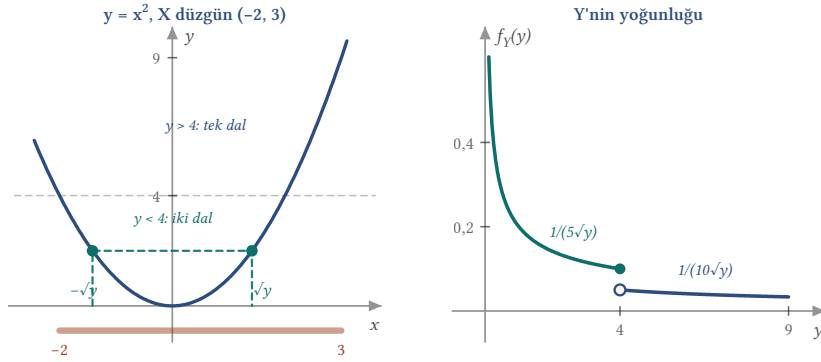
Sonuç:

$$f_Y(y) = \begin{cases} \frac{1}{5\sqrt{y}}, & 0 < y < 4 \\ \frac{1}{10\sqrt{y}}, & 4 < y < 9 \\ 0, & \text{diğer durumlarda.} \end{cases}$$

Sağlama:

$$\int_0^4 \frac{dy}{5\sqrt{y}} + \int_4^9 \frac{dy}{10\sqrt{y}} = \left[\frac{2\sqrt{y}}{5} \right]_0^4 + \left[\frac{\sqrt{y}}{5} \right]_4^9 = \frac{4}{5} + \frac{3-2}{5} = 1.$$

İlk parçanın olasılığı $\frac{4}{5} = P(-2 < X < 2)$, ikincinininki $\frac{1}{5} = P(2 < X < 3)$ 'tür; beklendiği gibi.



Şekil 12.4: $X, (-2, 3)$ üzerinde düzgün dağılırken $Y = X^2$ 'nin yoğunluğu. $0 < y < 4$ için hem $-\sqrt{y}$ hem $\sqrt{y}$ tanım aralığındadır: iki dal katkı verir ve yoğunluk $1/(5\sqrt{y})$ olur. $4 < y < 9$ için yalnızca $\sqrt{y}$ kalır ve yoğunluk yarıya, $1/(10\sqrt{y})$ 'ye düşer; $y = 4$ 'te yoğunluk sıçrar.

■

Son olarak, monoton bir dönüşümün bizi yepyeni bir dağılıma götürdüğü bir örnek görelim. $(-\frac{\pi}{2}, \frac{\pi}{2})$ aralığında düzgün dağılan bir açının tanjantı, bütün gerçel eksene yayılan, kuyrukları çok ağır bir dağılım verir.

Örnek 12.8 (Tanjant Dönüşümü ve Cauchy Dağılımı). X rastgele değişkeninin olasılık yoğunluk fonksiyonu

$$f_X(x) = \begin{cases} \frac{1}{\pi}, & -\frac{\pi}{2} < x < \frac{\pi}{2} \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun. $Y = \tan X$ rastgele değişkeninin olasılık yoğunluk fonksiyonunu bulunuz.

Çözüm.

Dönüşüm teoremiyle. $g(x) = \tan x$, $(-\frac{\pi}{2}, \frac{\pi}{2})$ üzerinde sürekli ve kesin artandır; $x \rightarrow -\frac{\pi}{2}^+$ iken $-\infty$ 'a, $x \rightarrow \frac{\pi}{2}^-$ iken $+\infty$ 'a gittiğinden bu aralığı $\mathbb{R} = (-\infty, \infty)$ 'a birebir ve örten biçimde götürür. Ters fonksiyonu

$$g^{-1}(y) = \arctan y, \quad \frac{dg^{-1}(y)}{dy} = \frac{1}{1+y^2}$$

her yerde sürekli türevlenebilirdir. Tek parça vardır ($k = 1$); her $y \in \mathbb{R}$ için

$$f_Y(y) = f_X(\arctan y) \left| \frac{1}{1+y^2} \right| = \frac{1}{\pi} \cdot \frac{1}{1+y^2} = \frac{1}{\pi(1+y^2)}.$$

Burada $-\frac{\pi}{2} < \arctan y < \frac{\pi}{2}$ olduğundan $f_X(\arctan y) = \frac{1}{\pi}$ 'dir.

Dağılım fonksiyonu tekniğiyle sağlama. tan artan olduğundan $\tan X \leq y \Leftrightarrow X \leq \arctan y$ ve

$$F_Y(y) = P(X \leq \arctan y) = \int_{-\pi/2}^{\arctan y} \frac{1}{\pi} dx = \frac{1}{\pi} \left(\arctan y + \frac{\pi}{2} \right) = \frac{1}{2} + \frac{\arctan y}{\pi}, \quad y \in \mathbb{R}.$$

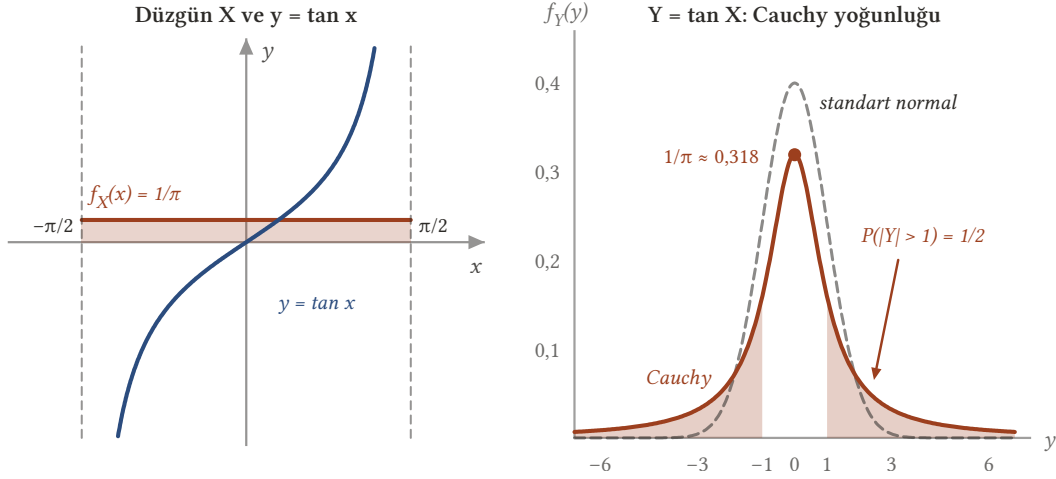
Türevi $\frac{1}{\pi(1+y^2)}$ 'dir; iki yol örtüşür. Toplam olasılık:

$$\int_{-\infty}^{\infty} \frac{dy}{\pi(1+y^2)} = \frac{1}{\pi} [\arctan y]_{-\infty}^{\infty} = \frac{1}{\pi} \left(\frac{\pi}{2} + \frac{\pi}{2} \right) = 1.$$

Bulduğumuz

$$f_Y(y) = \frac{1}{\pi(1+y^2)}, \quad y \in \mathbb{R}, \quad F_Y(y) = \frac{1}{2} + \frac{\arctan y}{\pi}$$

dağılımına **Cauchy dağılımı** denir. Yoğunluk 0 etrafında simetrik ve çan biçimlidir, ama kuyrukları $\frac{1}{\pi y^2}$ gibi, yani çok yavaş söner: $\arctan 1 = \frac{\pi}{4}$ olduğundan $P(Y > 1) = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$ ve simetriyle $P(|Y| > 1) = \frac{1}{2}$; $P(|Y| > 10) = 1 - \frac{2}{\pi} \arctan 10 \approx 0,064$. Geometrik yorumu şudur: $(0, 1)$ noktasındaki bir ışık kaynağı düşeyle rastgele bir X açısı yaparak bir ışın gönderirse, ışın x eksenini $\tan X$ noktasında keser; kesim noktası Cauchy dağılımlıdır. Ağır kuyruklar, kaynağın neredeyse yatay gönderdiği ışınların eksenin çok uzağına düşmesinden gelir.



Şekil 12.5: Tanjant dönüşümü ve Cauchy yoğunluğu. Solda X , $(-\pi/2, \pi/2)$ aralığında düzgün dağılır: yoğunluğu $1/\pi$ yüksekliğindeki dikdörtgendir; $y = \tan x$ eğrisi bu aralığı bütün gerçel eksene gerer, uçlara yaklaştıkça diklesir. Sağda $Y = \tan X$ 'in yoğunluğu $1/(\pi(1+y^2))$ ile kesikli standart normal yoğunluk karşılaştırılmıştır: tepe $1/\pi \approx 0,318$ normalinkinden alçaktır, ama kuyruklar çok daha ağırdır; taralı $|Y| > 1$ bölgesinin olasılığı $1/2$ 'dir.

■

12.6 Bir Yoğunluğun Baştan Sona İncelenmesi

Bölümün araçlarını bir arada kullanan iki örnekle bitirelim. İlkinde bir yoğunluk fonksiyonu doğrulanacak, koşullu olasılık hesaplanacak, doğrusal dönüşümün yoğunluğu bulunacak ve X 'in dağılım fonksiyonu kurulacak. İkincisinde parçalı bir yoğunluktaki bilinmeyen sabit belirlenecek ve dağılım fonksiyonu kurulacak.

Örnek 12.9 (Yoğunluk, Koşullu Olasılık ve Doğrusal Dönüşüm). X rastgele değişkeninin olasılık yoğunluk fonksiyonu

$$f_X(x) = \begin{cases} \frac{x}{2}, & 0 < x < 2 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun.

- f_X 'in bir olasılık yoğunluk fonksiyonu olup olmadığını inceleyiniz.
- $P(X < 1)$ ve $P(\frac{1}{2} < X < \frac{3}{2} \mid X < 1)$ olasılıklarını hesaplayınız.
- $Y = 3X + 2$ rastgele değişkeninin olasılık yoğunluk fonksiyonunu bulunuz.
- X 'in dağılım fonksiyonunu bulunuz.

Çözüm.

(a) Yoğunluk fonksiyonu olmanın iki koşulu vardır (Tanım 11.2): $f_X \geq 0$ ve toplam integral 1. $0 < x < 2$ için $\frac{x}{2} > 0$, diğer durumlarda $f_X = 0$; ilk koşul sağlanır. İkincisi:

$$\int_{-\infty}^{\infty} f_X(x) dx = \int_0^2 \frac{x}{2} dx = \left[\frac{x^2}{4} \right]_0^2 = 1.$$

Demek ki f_X bir olasılık yoğunluk fonksiyonudur.

(b) $P(X < 1) = \int_0^1 \frac{x}{2} dx = \left[\frac{x^2}{4} \right]_0^1 = \frac{1}{4}$.

Koşullu olasılık için (Tanım 6.1) kesişim olayına bakalım: $(\frac{1}{2} < X < \frac{3}{2}) \cap (X < 1) = (\frac{1}{2} < X < 1)$ ve

$$P\left(\frac{1}{2} < X < 1\right) = \int_{1/2}^1 \frac{x}{2} dx = \left[\frac{x^2}{4} \right]_{1/2}^1 = \frac{1}{4} - \frac{1}{16} = \frac{3}{16}.$$

Böylece

$$P\left(\frac{1}{2} < X < \frac{3}{2} \mid X < 1\right) = \frac{P\left(\frac{1}{2} < X < 1\right)}{P(X < 1)} = \frac{3/16}{1/4} = \frac{3}{4}.$$

(c) $0 < x < 2$ iken $2 < 3x + 2 < 8$; $D_Y = (2, 8)$. $g(x) = 3x + 2$ artan, $g^{-1}(y) = \frac{y-2}{3}$, türevi $\frac{1}{3}$. Dönüşüm teoremiyle ($k = 1$), $2 < y < 8$ için

$$f_Y(y) = f_X\left(\frac{y-2}{3}\right) \cdot \frac{1}{3} = \frac{1}{2} \cdot \frac{y-2}{3} \cdot \frac{1}{3} = \frac{y-2}{18}.$$

Dağılım fonksiyonu tekniğiyle de aynı sonuç çıkar: $F_Y(y) = P\left(X \leq \frac{y-2}{3}\right) = \frac{1}{4} \left(\frac{y-2}{3}\right)^2 = \frac{(y-2)^2}{36}$, türevi $\frac{y-2}{18}$. Demek ki

$$f_Y(y) = \begin{cases} \frac{y-2}{18}, & 2 < y < 8 \\ 0, & \text{diğer durumlarda.} \end{cases}$$

Sağlama: $\int_2^8 \frac{y-2}{18} dy = \left[\frac{(y-2)^2}{36} \right]_2^8 = \frac{36}{36} = 1$.

(d) $F_X(x) = \int_{-\infty}^x f_X(t) dt$ (Tanım 9.1). $x < 0$ için integral 0'dır. $0 \leq x < 2$ için $F_X(x) = \int_0^x \frac{t}{2} dt = \frac{x^2}{4}$. $x \geq 2$ için $F_X(x) = 1$. Böylece

$$F_X(x) = \begin{cases} 0, & x < 0 \\ \frac{x^2}{4}, & 0 \leq x < 2 \\ 1, & x \geq 2. \end{cases}$$

F_X süreklidir, 0'dan 1'e artar ve türevi f_X 'tir (Teorem 9.1, Teorem 11.2). (b)'deki $P(X < 1) = F_X(1) = \frac{1}{4}$ değeri buradan da okunur.

■

Örnek 12.10 (Parçalı Yoğunlukta Sabitin Belirlenmesi ve Dağılım Fonksiyonu). X rastgele değişkeninin olasılık yoğunluk fonksiyonu

$$f_X(x) = \begin{cases} x+1, & -1 < x < 0 \\ 2x-6, & 3 < x < c \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun.

(a) c sabitini bulunuz.

(b) $P(X = 2)$ olasılığını hesaplayınız.

(c) F_X dağılım fonksiyonunu oluşturunuz.

Çözüm.

(a) Önce işaret koşulu: $-1 < x < 0$ için $x + 1 > 0$; $c > 3$ olmak üzere $3 < x < c$ için $2x - 6 > 0$. Toplam integral 1 olmalıdır:

$$\int_{-1}^0 (x+1) dx = \left[\frac{(x+1)^2}{2} \right]_{-1}^0 = \frac{1}{2}, \quad \int_3^c (2x-6) dx = \left[x^2 - 6x \right]_3^c = c^2 - 6c + 9 = (c-3)^2.$$

$\frac{1}{2} + (c-3)^2 = 1$ denkleminde $(c-3)^2 = \frac{1}{2}$, $c > 3$ olduğundan

$$c = 3 + \frac{1}{\sqrt{2}} = 3 + \frac{\sqrt{2}}{2} \approx 3,71.$$

(b) X 'in yoğunluğu vardır; **Önerme 11.1** gereği $P(X=2) = 0$. (Burada ayrıca $f_X(2) = 0$ 'dır, ama sonucun nedeni bu değildir: $P(X=x) = 0$ eşitliği $f_X(x) > 0$ olan noktalarda da geçerlidir.)

(c) $F_X(x) = \int_{-\infty}^x f_X(t) dt$ 'yi parça parça hesaplayalım.

- $x < -1$: $F_X(x) = 0$.
- $-1 \leq x < 0$: $F_X(x) = \int_{-1}^x (t+1) dt = \frac{(x+1)^2}{2}$.
- $0 \leq x < 3$: bu aralıkta $f_X = 0$ olduğundan F_X sabittir; $F_X(x) = F_X(0) = \frac{1}{2}$.
- $3 \leq x < c$: $F_X(x) = \frac{1}{2} + \int_3^x (2t-6) dt = \frac{1}{2} + (x-3)^2$.
- $x \geq c$: $F_X(x) = 1$.

Yani

$$F_X(x) = \begin{cases} 0, & x < -1 \\ \frac{(x+1)^2}{2}, & -1 \leq x < 0 \\ \frac{1}{2}, & 0 \leq x < 3 \\ \frac{1}{2} + (x-3)^2, & 3 \leq x < 3 + \frac{\sqrt{2}}{2} \\ 1, & x \geq 3 + \frac{\sqrt{2}}{2}. \end{cases}$$

Süreklilik kontrolü: $x = 0$ 'da $\frac{1}{2}$, $x = 3$ 'te $\frac{1}{2}$, $x = c$ 'de $\frac{1}{2} + \frac{1}{2} = 1$; parçalar birbirine eklenir. $[0, 3]$ üzerindeki yatay basamak, $P(0 \leq X \leq 3) = 0$ olduğunu söyler: yoğunluğun sıfır olduğu bölgede dağılım fonksiyonu yerinde sayar.

■

12.7 Alıştırmalar

Alıştırma 12.1 (Dönüşümler). (a) X , hilesiz bir zarın gösterdiği sayı olsun. $Y = (X-3)^2$ rastgele değişkeninin olasılık fonksiyonunu bulunuz.

(b) X 'in olasılık yoğunluk fonksiyonu $f_X(x) = 2x$ ($0 < x < 1$) olsun. $Y = X^2$ 'nin $(0, 1)$ üzerinde düzgün dağıldığını, yani $f_Y(y) = 1$ ($0 < y < 1$) olduğunu gösteriniz.

(c) X 'in olasılık yoğunluk fonksiyonu $f_X(x) = e^{-x}$ ($x > 0$) olsun. $Y = 1 - e^{-X}$ 'in $(0, 1)$ üzerinde düzgün dağıldığını gösteriniz.

(d) X 'in olasılık yoğunluk fonksiyonu $f_X(x) = \frac{1}{2}$ ($-1 < x < 1$) olsun. $Y = |X|$ 'in olasılık yoğunluk fonksiyonunu dönüşüm teoremiyle bulunuz.

(e) X , $(0, 1)$ üzerinde düzgün dağılsın: $f_X(x) = 1$ ($0 < x < 1$). $Y = -2 \ln X$ 'in olasılık yoğunluk fonksiyonunu bulunuz.

(f) X 'in olasılık yoğunluk fonksiyonu

$$f_X(x) = \begin{cases} x+1, & -1 < x < 0 \\ 3x, & c < x \leq 3 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun. c sabitini ve F_X dağılım fonksiyonunu bulunuz.

(g) X , $(0, 1)$ üzerinde düzgün dağılsın. $Y = \lfloor 4X \rfloor$ rastgele değişkeni kesikli midir? Olasılık fonksiyonunu bulunuz.

Çözüm.

(a) $D_X = \{1, 2, 3, 4, 5, 6\}$, her değer $\frac{1}{6}$ olasılıklı. $g(x) = (x-3)^2$ değerleri: $1 \mapsto 4, 2 \mapsto 1, 3 \mapsto 0, 4 \mapsto 1, 5 \mapsto 4, 6 \mapsto 9$. $D_Y = \{0, 1, 4, 9\}$ ve **Önerme 12.1** ile

y	0	1	4	9
$f_Y(y)$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$

Toplam $\frac{1}{6} + \frac{2}{6} + \frac{2}{6} + \frac{1}{6} = 1$.

(b) $D_Y = (0, 1)$. $X > 0$ olduğundan $X^2 \leq y \Leftrightarrow X \leq \sqrt{y}$ ve $0 < y < 1$ için

$$F_Y(y) = P(X \leq \sqrt{y}) = \int_0^{\sqrt{y}} 2x \, dx = y.$$

$f_Y(y) = F'_Y(y) = 1$ ($0 < y < 1$). Dönüşüm teoremiyle de: $g^{-1}(y) = \sqrt{y}$, türevi $\frac{1}{2\sqrt{y}}$, $f_Y(y) = 2\sqrt{y} \cdot \frac{1}{2\sqrt{y}} = 1$.

(c) $X > 0$ iken $0 < e^{-X} < 1$, dolayısıyla $0 < 1 - e^{-X} < 1$; $D_Y = (0, 1)$. $0 < y < 1$ için

$$1 - e^{-X} \leq y \Leftrightarrow e^{-X} \geq 1 - y \Leftrightarrow -X \geq \ln(1 - y) \Leftrightarrow X \leq -\ln(1 - y)$$

ve $-\ln(1 - y) > 0$ olduğundan

$$F_Y(y) = \int_0^{-\ln(1-y)} e^{-x} \, dx = 1 - e^{\ln(1-y)} = 1 - (1 - y) = y.$$

Öyleyse $f_Y(y) = 1$ ($0 < y < 1$): Y düzgün dağılır. Burada $g(x) = 1 - e^{-x}$ tam olarak X 'in dağılım fonksiyonu F_X 'tir; dağılım fonksiyonu sürekli olan her X için $F_X(X)$ 'in $(0, 1)$ üzerinde düzgün dağılması genel bir olgudur.

(d) $g(x) = |x|$, $(-1, 0)$ üzerinde azalan ($g_1^{-1}(y) = -y$, türevi -1), $(0, 1)$ üzerinde artandır ($g_2^{-1}(y) = y$, türevi 1); her ikisi de $(0, 1)$ 'e örtendir. $0 < y < 1$ için

$$f_Y(y) = f_X(-y) \cdot |-1| + f_X(y) \cdot |1| = \frac{1}{2} + \frac{1}{2} = 1.$$

$Y = |X|$, $(0, 1)$ üzerinde düzgün dağılır.

(e) $g(x) = -2 \ln x$, $(0, 1)$ üzerinde kesin azalandır ve $(0, \infty)$ 'a örtendir; $g^{-1}(y) = e^{-y/2}$, türevi $-\frac{1}{2}e^{-y/2}$. $y > 0$ için

$$f_Y(y) = f_X(e^{-y/2}) \cdot \left| -\frac{1}{2}e^{-y/2} \right| = \frac{1}{2}e^{-y/2}.$$

Sağlama: $F_Y(y) = P(-2 \ln X \leq y) = P(X \geq e^{-y/2}) = 1 - e^{-y/2}$, türevi $\frac{1}{2}e^{-y/2}$. Bu, parametresi $\frac{1}{2}$ olan üstel dağılımdır.

(f) İşaret koşulu için $3x \geq 0$, yani $c \geq 0$ gerekir. İlk parçanın integrali $\int_{-1}^0 (x+1) \, dx = \frac{1}{2}$, ikincinin $\int_c^3 3x \, dx = \frac{3}{2}(9 - c^2)$. $\frac{1}{2} + \frac{3}{2}(9 - c^2) = 1$ denkleminde $9 - c^2 = \frac{1}{3}$, $c^2 = \frac{26}{3}$ ve $c \geq 0$ ile

$$c = \sqrt{\frac{26}{3}} = \frac{\sqrt{78}}{3} \approx 2,94.$$

Dağılım fonksiyonu: $x < -1$ için 0 ; $-1 \leq x < 0$ için $\frac{(x+1)^2}{2}$; $0 \leq x \leq c$ için $\frac{1}{2}$; $c < x \leq 3$ için

$$F_X(x) = \frac{1}{2} + \int_c^x 3t \, dt = \frac{1}{2} + \frac{3}{2} \left(x^2 - \frac{26}{3} \right) = \frac{3x^2 - 25}{2};$$

$x > 3$ için 1 . Kontrol: $F_X(c) = \frac{26-25}{2} = \frac{1}{2}$, $F_X(3) = \frac{27-25}{2} = 1$.

(g) $0 < X < 1$ iken $0 < 4X < 4$, dolayısıyla $[4X] \in \{0, 1, 2, 3\}$; Y kesiklidir. $k = 0, 1, 2, 3$ için

$$P(Y = k) = P(k \leq 4X < k+1) = P\left(\frac{k}{4} \leq X < \frac{k+1}{4}\right) = \frac{1}{4}.$$

Y , $\{0, 1, 2, 3\}$ üzerinde dűzgűn dađılan kesikli bir rastgele deđiřkendir; sűrekli bir deđiřkenden kesikli bir deđiřken elde edilmiřtir.

■

Buraya kadar hep tek bir rastgele deđiřkenle alıřtıktık. Bundan sonra aynı deneyde birden ok rastgele deđiřkenin birlikte davranıřını, yani ortak dađılımlarını inceleyeceđiz: [Rastgele Vektörler ve Ortak Dađılımlar](#).

13. Rastgele Vektörler ve Ortak Dağılımlar

Rastgele Değişkenin Fonksiyonunun Dağılımı bölümüyle birlikte tek bir rastgele değişkenin incelenmesini tamamladık: bir rastgele değişkenin dağılım fonksiyonunu, kesikli durumda olasılık fonksiyonunu, sürekli durumda yoğunluk fonksiyonunu tanıdık ve bu araçlarla X 'e bağlı her olayın olasılığını hesaplamayı öğrendik. Ne var ki bir rastgele deneyde çoğu zaman tek bir sayı değil, birden çok sayı gözlenir. Bir hastanın boyu ile kilosunu, bir parçanın çapı ile ağırlığı, üç para atışında gelen toplam tura sayısı ile ilk iki atıştaki tura sayısı aynı deneyin sonucuna bağlı büyüklüklerdir. Bu büyüklükler birbirini etkileyebildiğinden her birini ayrı ayrı incelemek yeterli değildir; **birlikte** davranışlarını anlatan bir araca ihtiyaç vardır.

Bu bölümde bu aracı kuruyoruz. Aynı örnek uzayda tanımlı n tane rastgele değişkeni yan yana yazıp elde edilen $(X_1, X_2, \dots, X_n)$ nesnesine **rastgele vektör** diyeceğiz. Tek değişkenli kuramda yaptığımız her şeyin bir karşılığı burada da vardır: rastgele vektörün bir olasılık dağılımı, bir ortak dağılım fonksiyonu, kesikli durumda bir ortak olasılık fonksiyonu, sürekli durumda bir ortak yoğunluk fonksiyonu vardır. Yeni olan iki şey var: tek boyutta karşılığı bulunmayan **dikdörtgen formülü** ve ortak dağılımdan tek bir bileşenin dağılımına inmeyi sağlayan **marjinal dağılım**.

Kesikli durumda ortak olasılık fonksiyonunun tablo gösterimini, sürekli durumda ise ortak yoğunluğun tanım bölgesi üzerinde integral almayı iki ayrıntılı örnekle işleyeceğiz. Bağımsızlık ve koşullu dağılım kavramları bir sonraki bölümün konusudur.

13.1 Rastgele Vektör ve Olasılık Dağılımı

Bir rastgele değişken, her $a \in \mathbb{R}$ için $(X \leq a)$ kümesini olay yapan bir $X : \Omega \rightarrow \mathbb{R}$ fonksiyonuydu (**Tanım 8.1**). Bunun n boyutlu karşılığında Ω 'dan $\mathbb{R}^n$ 'e giden bir fonksiyon alınır ve her bileşen için aynı anda konan eşitsizliklerin olay tanımlaması istenir.

Tanım 13.1 (Rastgele Vektör). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı ve $X_1, X_2, \dots, X_n : \Omega \rightarrow \mathbb{R}$ fonksiyonlar olsun.

$$(X_1, X_2, \dots, X_n) : \Omega \rightarrow \mathbb{R}^n, \quad \omega \mapsto (X_1(\omega), X_2(\omega), \dots, X_n(\omega))$$

fonksiyonu, her $(a_1, a_2, \dots, a_n) \in \mathbb{R}^n$ için

$$\{\omega \in \Omega : X_1(\omega) \leq a_1, X_2(\omega) \leq a_2, \dots, X_n(\omega) \leq a_n\} \in \mathcal{U}$$

koşulunu sağlıyorsa $(X_1, X_2, \dots, X_n)$ fonksiyonuna **n -boyutlu rastgele vektör** (n -boyutlu rastgele değişken) denir. Rastgele vektörün aldığı değerlerin kümesi

$$D_{(X_1, \dots, X_n)} = \{(X_1(\omega), \dots, X_n(\omega)) : \omega \in \Omega\} \subseteq \mathbb{R}^n$$

rastgele vektörün **değer kümesidir**. Yukarıdaki olay kısaca $(X_1 \leq a_1, \dots, X_n \leq a_n)$ ile gösterilir.

Tanımdaki koşul ile “her bileşenin bir rastgele değişken olması” koşulu aynı şeydir. Bu, uygulamada rastgele vektör kurmanın yolunu açar: aynı deney üzerinde tanımlı rastgele değişkenleri yan yana yazmak yeter.

Lemma 13.1 (Rastgele Vektör ve Bileşenleri). $X_1, \dots, X_n : \Omega \rightarrow \mathbb{R}$ fonksiyonlar olsun. $(X_1, \dots, X_n)$ 'in bir rastgele vektör olması için gerek ve yeter koşul, her $i = 1, \dots, n$ için X_i 'nin bir rastgele değişken olmasıdır.

İspat.

Her X_i rastgele değişken olsun. $(a_1, \dots, a_n) \in \mathbb{R}^n$ verildiğinde

$$(X_1 \leq a_1, \dots, X_n \leq a_n) = \bigcap_{i=1}^n (X_i \leq a_i)$$

yazılır. Sağdaki her küme $\mathcal{U}$ 'dadır ve $\mathcal{U}$ bir σ -cebiri olduğundan sonlu kesişimlere kapalıdır; öyleyse soldaki küme de $\mathcal{U}$ 'dadır. Demek ki $(X_1, \dots, X_n)$ bir rastgele vektördür.

Tersine $(X_1, \dots, X_n)$ bir rastgele vektör olsun; X_1 'in rastgele değişken olduğunu gösterelim (öteki bileşenler için akıl yürütme aynıdır). $a_1 \in \mathbb{R}$ verilsin. Her ω için $X_2(\omega), \dots, X_n(\omega)$ birer gerçel sayı olduğundan yeterince büyük bir $k \in \mathbb{N}$ için hiçbir k 'yı aşmaz; bu yüzden

$$(X_1 \leq a_1) = \bigcup_{k=1}^{\infty} (X_1 \leq a_1, X_2 \leq k, \dots, X_n \leq k)$$

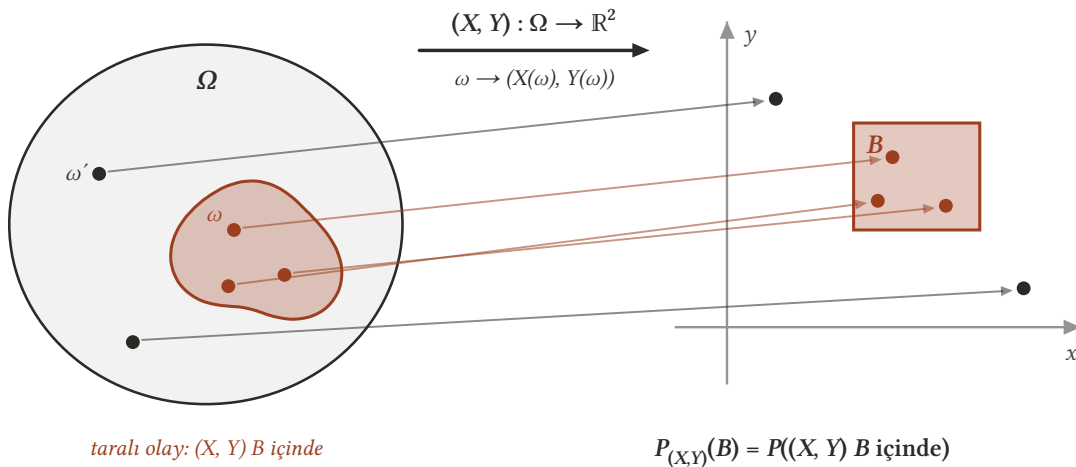
olur. Sağdaki kümelerin her biri rastgele vektör tanımı gereği $\mathcal{U}$ 'dadır; $\mathcal{U}$ sayılabilir birleşimlere kapalı olduğundan $(X_1 \leq a_1) \in \mathcal{U}$ bulunur. Yani X_1 bir rastgele değişkendir. ■

Tek boyutta rastgele değişken, örnek uzaydaki olasılığı gerçel sayılara taşıyordu: $P_X(B) = P(X \in B)$ bir olasılık ölçüsüydü (Teorem 8.1). Rastgele vektör de aynı işi $\mathbb{R}^n$ üzerinde yapar. Bu sonuç, vektörle ilgili her olasılığın $\mathbb{R}^n$ 'in Borel kümeleri üzerinde tanımlı tek bir ölçüden okunabileceğini söyler; bölümün geri kalanı bu ölçüyü elle tutulur biçimlerde (fonksiyon, tablo, integral) yazma çabasıdır.

Teorem 13.1 (Rastgele Vektörün Olasılık Dağılımı). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı ve $(X_1, \dots, X_n)$ bir n -boyutlu rastgele vektör olsun. Her $B \in \mathcal{B}(\mathbb{R}^n)$ için $\{\omega \in \Omega : (X_1(\omega), \dots, X_n(\omega)) \in B\}$ kümesi bir olaydır ve

$$P_{(X_1, \dots, X_n)} : \mathcal{B}(\mathbb{R}^n) \rightarrow [0, 1], \quad P_{(X_1, \dots, X_n)}(B) = P(\{\omega \in \Omega : (X_1(\omega), \dots, X_n(\omega)) \in B\})$$

fonksiyonu $\mathcal{B}(\mathbb{R}^n)$ üzerinde bir olasılık ölçüsüdür. Bu ölçüye $(X_1, \dots, X_n)$ rastgele vektörünün **olasılık dağılımı** denir.



Şekil 13.1: Rastgele vektör, örnek uzayın her sonucuna bir sayı çifti (bir düzlem noktası) eşleyen fonksiyondur. Düzlemde bir B Borel kümesi (burada bir dikdörtgen) alındığında, görüntüsü B 'ye düşen sonuçların kümesi Ω 'da bir olaydır: $(X, Y) \in B$ önima. Vektörün dağılımı $P_{(X,Y)}$, her B 'ye bu önimaın olasılığını verir; böylece Ω üzerindeki olasılık düzleme taşınır.

İspat.

Kısalık için $\mathbf{X} = (X_1, \dots, X_n)$ yazalım; $B \subseteq \mathbb{R}^n$ için $\mathbf{X}^{-1}(B) = \{\omega : \mathbf{X}(\omega) \in B\}$ ters görüntü kümesidir.

Adım 1: $\mathbf{X}^{-1}(B)$ her Borel kümesi için bir olaydır. Ters görüntüsü olay olan kümelerin sınıfını

$$\mathcal{C} = \{B \subseteq \mathbb{R}^n : \mathbf{X}^{-1}(B) \in \mathcal{U}\}$$

ile gösterelim. Ters görüntü alma işlemi tümleyen ve birleşimle uyumludur:

$$\mathbf{X}^{-1}(B^c) = (\mathbf{X}^{-1}(B))^c, \quad \mathbf{X}^{-1}\left(\bigcup_{k=1}^{\infty} B_k\right) = \bigcup_{k=1}^{\infty} \mathbf{X}^{-1}(B_k).$$

$\mathcal{U}$ bir σ -cebiri olduğundan bu eşitlikler, $\mathcal{C}$ 'nin tümleyene ve sayılabilir birleşime kapalı olduğunu söyler; $\mathbf{X}^{-1}(\mathbb{R}^n) = \Omega \in \mathcal{U}$ olduğundan $\mathbb{R}^n \in \mathcal{C}$ 'dir. Yani $\mathcal{C}$, $\mathbb{R}^n$ üzerinde bir σ -cebiri. Öte yandan $\mathbf{a} = (a_1, \dots, a_n) \in \mathbb{R}^n$ için $K(\mathbf{a}) = (-\infty, a_1] \times \dots \times (-\infty, a_n]$ yazarsak, rastgele vektör tanımı bu biçimdeki her "köşe kümesinin" $\mathcal{C}$ 'de olduğunu söyler. Köşe kümelerinden σ -cebiri işlemleriyle açık dikdörtgenlere ulaşılır. $a_i < b_i$ olmak üzere yarı açık dikdörtgen

$$(a_1, b_1] \times \dots \times (a_n, b_n] = K(\mathbf{b}) \setminus \bigcup_{i=1}^n K(b_1, \dots, b_{i-1}, a_i, b_{i+1}, \dots, b_n)$$

biçiminde yazılır: bir nokta $K(\mathbf{b})$ 'de olup sağdaki birleşimin dışında kalıyorsa her i için $a_i < x_i \leq b_i$ 'dir. Açık dikdörtgen ise yarı açık dikdörtgenlerin sayılabilir birleşimidir:

$$(a_1, b_1) \times \dots \times (a_n, b_n) = \bigcup_{k \geq k_0} \left(a_1, b_1 - \frac{1}{k} \right] \times \dots \times \left(a_n, b_n - \frac{1}{k} \right];$$

burada k_0 , her i için $b_i - \frac{1}{k_0} > a_i$ olacak kadar büyük seçilir. Demek ki her açık dikdörtgen $\mathcal{C}$ 'dedir. $\mathcal{B}(\mathbb{R}^n)$, açık dikdörtgenlerin ürettiği en küçük σ -cebiri olduğundan (Tanım 2.5) $\mathcal{B}(\mathbb{R}^n) \subseteq \mathcal{C}$ elde edilir. Bu, $P_{\mathbf{X}}(B)$ 'nin her Borel kümesi için tanımlı olduğunu gösterir.

Adım 2: Kolmogorov aksiyomları. P bir olasılık ölçüsü olduğundan her B için $P_{\mathbf{X}}(B) = P(\mathbf{X}^{-1}(B)) \geq 0$ 'dır ve $P_{\mathbf{X}}(\mathbb{R}^n) = P(\Omega) = 1$ 'dir. $B_1, B_2, \dots \in \mathcal{B}(\mathbb{R}^n)$ ikiye ikiye ayrık olsun. Bir ω aynı anda iki farklı B_k 'ye düşmeyeceğinden $\mathbf{X}^{-1}(B_1), \mathbf{X}^{-1}(B_2), \dots$ olayları da ikiye ikiye ayrıktır. P 'nin sayılabilir toplamsallığından

$$P_{\mathbf{X}}\left(\bigcup_{k=1}^{\infty} B_k\right) = P\left(\bigcup_{k=1}^{\infty} \mathbf{X}^{-1}(B_k)\right) = \sum_{k=1}^{\infty} P(\mathbf{X}^{-1}(B_k)) = \sum_{k=1}^{\infty} P_{\mathbf{X}}(B_k)$$

bulunur. Üç aksiyom da sağlandığından $P_{\mathbf{X}}$ bir olasılık ölçüsüdür. ■

Bundan böyle $P_{(X_1, \dots, X_n)}(B)$ yerine daha okunaklı olan $P((X_1, \dots, X_n) \in B)$ yazacağız. Bölümün geri kalanında iki boyutlu rastgele vektörleri, yani (X, Y) ya da (X_1, X_2) çiftlerini ele alacağız. Yapılan işlem, indis sayısını artırmaktan başka bir değişiklik gerektirmeden n -boyutlu rastgele vektörlere genişler.

13.2 Ortak Dağılım Fonksiyonu

Tek boyutta P_X ölçüsünün bütün bilgisini $F_X(x) = P(X \leq x)$ dağılım fonksiyonu taşıyordu (Tanım 9.1). İki boyutta aynı rolü, köşesi (x, y) noktasında olan sol-alt çeyrek düzlemin olasılığı oynar.

Tanım 13.2 (Ortak Dağılım Fonksiyonu). (X, Y) bir rastgele vektör olsun.

$$F_{X,Y} : \mathbb{R}^2 \rightarrow [0, 1], \quad F_{X,Y}(x, y) = P(X \leq x, Y \leq y) = P((X, Y) \in (-\infty, x] \times (-\infty, y])$$

funksiyonuna (X, Y) rastgele vektörünün **dağılım fonksiyonu** ya da X ile Y rastgele değişkenlerinin **ortak dağılım fonksiyonu** (ortak birikimli dağılım fonksiyonu) denir. Burada $(X \leq x, Y \leq y)$ yazımı $(X \leq x) \cap (Y \leq y)$ olayının kısaltmasıdır. X ve Y 'nin kendi dağılım fonksiyonları F_X ve F_Y 'ye bu bağlamda **marjinal dağılım fonksiyonları** denir.

Tek değişkenli dağılım fonksiyonunun temel özellikleri (monotonluk, sınırlar, sağdan süreklilik) burada da geçerlidir; ayrıca ortak dağılım fonksiyonundan marjinal dağılım fonksiyonları okunur. Yeni olan, bir dikdörtgenin olasılığını veren formüldür: bu formül, tek boyuttaki " $F(b) - F(a)$ " işleminin yerini alır ve dört köşe değerinin işaretli toplamı olarak çıkar.

Önerme 13.1 (Ortak Dağılım Fonksiyonunun Özellikleri). $F = F_{X,Y}$, (X, Y) rastgele vektörünün ortak dağılım fonksiyonu olsun.

(i) F her değişkeninde azalmayıp: $x_1 \leq x_2$ ise $F(x_1, y) \leq F(x_2, y)$; $y_1 \leq y_2$ ise $F(x, y_1) \leq F(x, y_2)$.

(ii) Her sabit y için $\lim_{x \rightarrow -\infty} F(x, y) = 0$; her sabit x için $\lim_{y \rightarrow -\infty} F(x, y) = 0$; ayrıca $x \rightarrow \infty$ ve $y \rightarrow \infty$ iken $F(x, y) \rightarrow 1$.

(iii) Marjinal dağılım fonksiyonları F' 'den okunur:

$$F_X(x) = \lim_{y \rightarrow \infty} F(x, y), \quad F_Y(y) = \lim_{x \rightarrow \infty} F(x, y).$$

(iv) F her değişkeninde sağdan süreklidir.

(v) **(Dikdörtgen formülü)** $a < b$ ve $c < d$ için

$$P(a < X \leq b, c < Y \leq d) = F(b, d) - F(a, d) - F(b, c) + F(a, c).$$

Özel olarak sağ taraf her zaman ≥ 0 'dır.

İspat.

Bütün ispatlar olasılık ölçüsünün iki özelliğine dayanır: monotonluk ve fark kuralı (Teorem 3.1) ile artan ya da azalan olay dizileri için olasılığın sürekliliği (Teorem 3.5); yani $A_1 \subseteq A_2 \subseteq \dots$ ise $P(A_k) \rightarrow P(\bigcup_k A_k)$ ve $A_1 \supseteq A_2 \supseteq \dots$ ise $P(A_k) \rightarrow P(\bigcap_k A_k)$.

(i) $x_1 \leq x_2$ ise $(X \leq x_1, Y \leq y) \subseteq (X \leq x_2, Y \leq y)$ olduğundan monotonluk gereği $F(x_1, y) \leq F(x_2, y)$ 'dir. İkinci değişken için akıl yürütme aynıdır.

(ii) y sabit olsun ve $A_k = (X \leq -k, Y \leq y)$, $k \in \mathbb{N}$, olaylarını alalım. Bu dizi azalandır ve kesişimi boşdur: $X(\omega)$ bir gerçel sayı olduğundan yeterince büyük k için $X(\omega) \leq -k$ sağlanamaz. Süreklilikten $F(-k, y) = P(A_k) \rightarrow P(\emptyset) = 0$ olur. $F(\cdot, y)$ azalmayan olduğundan $x \rightarrow -\infty$ iken limit vardır ve bu dizisel limite eşittir; yani $\lim_{x \rightarrow -\infty} F(x, y) = 0$. İkinci limit aynı biçimde görülür. Son olarak $B_k = (X \leq k, Y \leq k)$ olayları artandır ve birleşimleri Ω 'dır (her ω için $X(\omega)$ ve $Y(\omega)$ sonlu sayılardır); süreklilikten $F(k, k) = P(B_k) \rightarrow P(\Omega) = 1$. $x \geq k$ ve $y \geq k$ olduğunda (i) gereği $F(k, k) \leq F(x, y) \leq 1$ olduğundan $x, y \rightarrow \infty$ iken $F(x, y) \rightarrow 1$ bulunur.

(iii) x sabit olsun. $C_k = (X \leq x, Y \leq k)$ olayları artandır ve $Y(\omega)$ sonlu olduğundan birleşimleri $(X \leq x)$ 'tir. Süreklilikten $F(x, k) = P(C_k) \rightarrow P(X \leq x) = F_X(x)$. $F(x, \cdot)$ azalmayan olduğundan bu, $\lim_{y \rightarrow \infty} F(x, y) = F_X(x)$ demektir. F_Y için aynı yol izlenir.

(iv) $D_k = (X \leq x + \frac{1}{k}, Y \leq y)$ olayları azalandır ve kesişimleri $(X \leq x, Y \leq y)$ 'dir; süreklilikten $F(x + \frac{1}{k}, y) \rightarrow F(x, y)$. $F(\cdot, y)$ azalmayan olduğundan bu, $F(x^+, y) = F(x, y)$ demektir. İkinci değişken için aynı.

(v) $E = (X \leq b, Y \leq d)$, $E_1 = (X \leq a, Y \leq d)$, $E_2 = (X \leq b, Y \leq c)$ olaylarını alalım. $a < b$ ve $c < d$ olduğundan $E_1 \subseteq E$ ve $E_2 \subseteq E$ 'dir. İstenen olay, E 'den $E_1 \cup E_2$ 'nin çıkarılmasıyla elde edilir:

$$(a < X \leq b, c < Y \leq d) = E \setminus (E_1 \cup E_2).$$

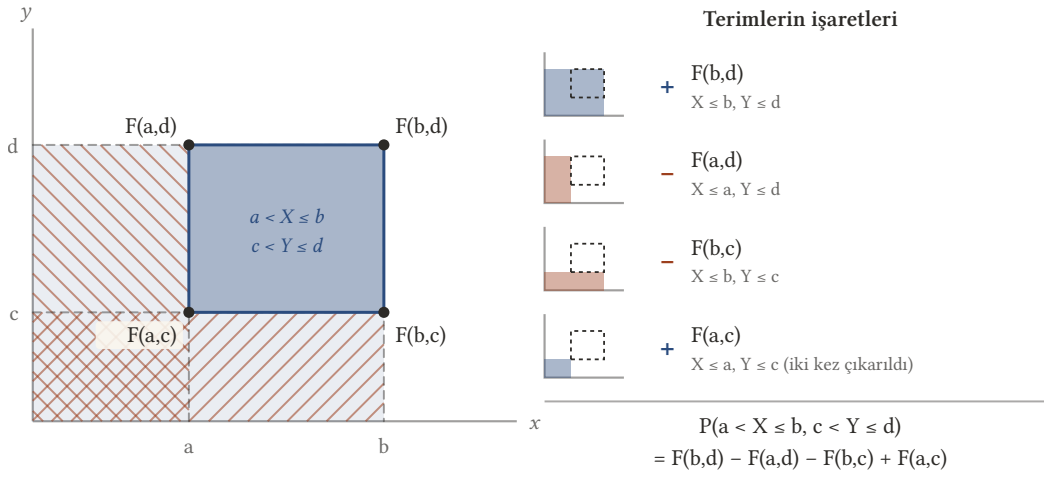
Gerçekten $\omega \in E$ için " $\omega \notin E_1$ " demek $X(\omega) > a$ demek, " $\omega \notin E_2$ " demek $Y(\omega) > c$ demektir. $E_1 \cup E_2 \subseteq E$ olduğundan fark kuralı uygulanır:

$$P(a < X \leq b, c < Y \leq d) = P(E) - P(E_1 \cup E_2).$$

$E_1 \cap E_2 = (X \leq a, Y \leq c)$ olduğundan toplama kuralı (Teorem 3.2) ile $P(E_1 \cup E_2) = F(a, d) + F(b, c) - F(a, c)$ bulunur. Yerine yazınca

$$P(a < X \leq b, c < Y \leq d) = F(b, d) - F(a, d) - F(b, c) + F(a, c)$$

elde edilir. Sol taraf bir olasılık olduğundan sağ taraf negatif olamaz.



Şekil 13.2: Dikdörtgen formülünün içirme-dışarma yorumu. Solda koyu taralı dikdörtgen $(a, b] \times (c, d]$ olayıdır; $F(b,d)$ bu dikdörtgeni içeren sol-alt çeyrek düzlemin olasılığıdır. Eğik taramalı şeritler $F(a,d)$ ve $F(b,c)$ bölgeleridir; ortak parçaları $F(a,c)$ çapraz taralıdır ve iki kez çıkarıldığı için bir kez geri eklenir. Sağdaki tablo her terimin işaretini gösterir.

⚠ Dikdörtgen koşulu tek boyuttan gelen özelliklerden çıkmaz

Tek boyutta azalmayan, sağdan sürekli ve sınırları 0 ile 1 olan her fonksiyon bir dağılım fonksiyonudur. İki boyutta (i)-(iv) yetmez; (v)'deki işaretli toplamın negatif olmaması ayrıca istenmelidir. Örneğin

$$G(x, y) = \begin{cases} 1, & x \geq 0, y \geq 0, x + y \geq 1 \\ 0, & \text{diğer durumlarda} \end{cases}$$

fonksiyonu her değişkeninde azalmayıp, sağdan sürekli ve $-\infty$ 'da 0, $+\infty$ 'da 1 sınırlarına sahiptir. Ama $(0, 1] \times (0, 1]$ dikdörtgeni için

$$G(1, 1) - G(0, 1) - G(1, 0) + G(0, 0) = 1 - 1 - 1 + 0 = -1 < 0$$

olduğundan G hiçbir rastgele vektörün ortak dağılım fonksiyonu olamaz.

Tek boyutta olduğu gibi burada da iki ana sınıf vardır: değer kümesi sayılabilir olan **kesikli** vektörler ve olasılığı bir yoğunluk fonksiyonunun integraliyle verilen **sürekli** vektörler. Önce kesikli durumu ele alıyoruz.

13.3 Kesikli Rastgele Vektörler

Tanım 13.3 (Kesikli Rastgele Vektör). (X_1, X_2) rastgele vektörünün değer kümesi $D_{(X_1, X_2)}$ sonlu ya da sayılabilir sonsuz ise (X_1, X_2) 'ye **kesikli rastgele vektör** denir.

Bu koşul, bileşenlerin kesikli olmasıyla aynıdır. Gerçekten D_{X_1} ve D_{X_2} , $D_{(X_1, X_2)}$ kümesinin eksenlere izdüşümleridir; sayılabilir bir kümenin izdüşümü sayılabilir olduğundan kesikli bir vektörün bileşenleri kesikli rastgele değişkenlerdir (Tanım 10.1). Tersine X_1 ve X_2 kesikli ise $D_{(X_1, X_2)} \subseteq D_{X_1} \times D_{X_2}$ ve iki sayılabilir kümenin çarpımı sayılabilir olduğundan (X_1, X_2) kesiklidir.

Tek değişkenli kesikli durumda P_X ölçüsü, $f_X(x) = P(X = x)$ olasılık fonksiyonuyla (Tanım 10.2) tümüyle belirleniyordu. İki boyutta aynı işi, her değer çiftinin olasılığını veren fonksiyon görür.

Tanım 13.4 (Ortak Olasılık Fonksiyonu). (X_1, X_2) kesikli bir rastgele vektör olsun.

$$f_{X_1, X_2}(x_1, x_2) = P(X_1 = x_1, X_2 = x_2), \quad (x_1, x_2) \in D_{(X_1, X_2)}$$

fonksiyonuna (X_1, X_2) rastgele vektörünün **olasılık fonksiyonu** ya da X_1 ile X_2 rastgele değişkenlerinin **ortak olasılık fonksiyonu** (ortak olasılık kütle fonksiyonu) denir. $D_{(X_1, X_2)}$ dışındaki $(x_1, x_2) \in \mathbb{R}^2$ noktalarında $f_{X_1, X_2}(x_1, x_2) = 0$ alınır.

Aşağıdaki sonuç, ortak olasılık fonksiyonunun tek değişkenli olasılık fonksiyonunun iki koşulunu sağladığını ve tıpkı onun gibi bütün olasılık dağılımını belirlediğini söyler.

Önerme 13.2 (Ortak Olasılık Fonksiyonunun Özellikleri). (X_1, X_2) kesikli bir rastgele vektör ve $f = f_{X_1, X_2}$ onun ortak olasılık fonksiyonu olsun.

(i) Her $(x_1, x_2) \in D_{(X_1, X_2)}$ için $f(x_1, x_2) \geq 0$ 'dır.

(ii) $\sum_{(x_1, x_2) \in D_{(X_1, X_2)}} f(x_1, x_2) = 1$ 'dir.

(iii) Her $B \in \mathcal{B}(\mathbb{R}^2)$ için

$$P((X_1, X_2) \in B) = \sum_{(x_1, x_2) \in B \cap D_{(X_1, X_2)}} f(x_1, x_2).$$

İspat.

$D = D_{(X_1, X_2)}$ yazalım. (i) Her olasılık gibi $f(x_1, x_2) = P(X_1 = x_1, X_2 = x_2) \geq 0$ 'dır.

(iii) B bir Borel kümesi olsun. Her ω için $(X_1(\omega), X_2(\omega)) \in D$ olduğundan $((X_1, X_2) \in B)$ olayı, $B \cap D$ 'deki noktalara göre parçalanır:

$$((X_1, X_2) \in B) = \bigcup_{(x_1, x_2) \in B \cap D} (X_1 = x_1, X_2 = x_2).$$

Sağdaki olaylar ikiye ikiye ayrıktır (bir ω , iki farklı noktaya aynı anda eşlenemez) ve D sayılabilir olduğundan birleşim sayılabilir. Sayılabilir toplamsallıkla

$$P((X_1, X_2) \in B) = \sum_{(x_1, x_2) \in B \cap D} P(X_1 = x_1, X_2 = x_2) = \sum_{(x_1, x_2) \in B \cap D} f(x_1, x_2)$$

bulunur.

(ii) (iii)'te $B = \mathbb{R}^2$ alınırsa sol taraf $P(\Omega) = 1$, sağ taraf D üzerindeki toplamdır.

■

(iii) maddesi, kesikli bir vektörle ilgili her olasılığın, olayın içine düşen değer çiftlerinin olasılıklarını toplayarak hesaplandığını söyler. Özel olarak ortak dağılım fonksiyonu

$$F_{X_1, X_2}(x_1, x_2) = \sum_{\substack{(u, v) \in D \\ u \leq x_1, v \leq x_2}} f_{X_1, X_2}(u, v)$$

toplamıdır.

13.3.1 Marjinal Olasılık Fonksiyonları

Ortak olasılık fonksiyonu elimizdeyken tek bir bileşenle ilgili olasılıklar nasıl bulunur? Örneğin $P(X_1 = x_1)$ için X_2 'nin hangi değeri aldığı önemsizdir; X_2 'nin alabileceği bütün değerler üzerinden toplamak gerekir. Bu toplama işlemine **marjinalleştirme** denir.

Tanım 13.5 (Marjinal Olasılık Fonksiyonları). (X, Y) kesikli bir rastgele vektör ve $f_{X, Y}$ ortak olasılık fonksiyonu olsun.

$$f_X(x) = \sum_{y \in D_Y} f_{X, Y}(x, y), \quad x \in D_X; \quad f_Y(y) = \sum_{x \in D_X} f_{X, Y}(x, y), \quad y \in D_Y$$

fonksiyonlarına sırasıyla X 'in ve Y 'nin **marjinal olasılık fonksiyonları** denir.

Adlandırma, marjinal fonksiyonların gerçekten X ve Y 'nin olasılık fonksiyonları olmasını gerektirir; bunu ispatlayalım.

Önerme 13.3 (Marjinal Olasılık Fonksiyonları Birer Olasılık Fonksiyonudur). (X, Y) kesikli bir rastgele vektör olsun. Her $x \in D_X$ ve $y \in D_Y$ için

$$f_X(x) = P(X = x), \quad f_Y(y) = P(Y = y)$$

olur; yani f_X ve f_Y , sırasıyla X ve Y kesikli rastgele değişkenlerinin olasılık fonksiyonlarıdır. Özel olarak $f_X \geq 0$, $f_Y \geq 0$ ve

$$\sum_{x \in D_X} f_X(x) = 1, \quad \sum_{y \in D_Y} f_Y(y) = 1.$$

İspat.

$x \in D_X$ sabit olsun. Her ω için $Y(\omega) \in D_Y$ olduğundan $(X = x)$ olayı, Y 'nin değerlerine göre parçalanır:

$$(X = x) = \bigcup_{y \in D_Y} (X = x, Y = y).$$

Farklı y 'lere karşılık gelen olaylar ayrıktır ve D_Y sayılabilir. Sayılabilir toplamsallıkla

$$P(X = x) = \sum_{y \in D_Y} P(X = x, Y = y) = \sum_{y \in D_Y} f_{X,Y}(x, y) = f_X(x).$$

Burada $(x, y) \notin D_{(X,Y)}$ olan çiftler için $f_{X,Y}(x, y) = 0$ olduğundan toplamı D_Y 'nin tamamı üzerinden yazmak sakıncasızdır. X kesikli bir rastgele değişkendir ([Lemma 13.1](#) ve D_X 'in sayılabilirliği); olasılık fonksiyonu tanım gereği $P(X = x)$ 'tir, dolayısıyla f_X bu olasılık fonksiyonudur ve [Tanım 10.2](#)'ndeki iki koşulu sağlar. Toplamın 1 olduğu doğrudan da görülür: negatif olmayan terimlerin toplama sırası değiştirilebildiğinden

$$\sum_{x \in D_X} f_X(x) = \sum_{x \in D_X} \sum_{y \in D_Y} f_{X,Y}(x, y) = \sum_{(x,y) \in D_{(X,Y)}} f_{X,Y}(x, y) = 1$$

([Önerme 13.2](#)). f_Y için ispat, X ile Y 'nin rollerini değiştirerek aynen yürür.

■

13.3.2 Tablo Gösterimi

Değer kümeleri sonlu olduğunda ortak olasılık fonksiyonu en rahat bir tablo ile verilir: satırlar X 'in, sütunlar Y 'nin değerlerini taşır; hücrelere $f_{X,Y}(x_i, y_j)$ yazılır. Marjinaller ise satır ve sütun toplamlarıdır; tablonun kenarında ("marjında") yer aldıkları için bu adı almışlardır.

Önerme 13.4 (Ortak Dağılım Tablosu). (X, Y) kesikli bir rastgele vektör, $D_X = \{x_1, \dots, x_m\}$ ve $D_Y = \{y_1, \dots, y_n\}$ sonlu olsun. $f = f_{X,Y}$ değerleri aşağıdaki tabloya yerleştirildiğinde her satırın toplamı o satırdaki x_i için $f_X(x_i)$ 'ye, her sütunun toplamı o sütundaki y_j için $f_Y(y_j)$ 'ye ve bütün hücrelerin toplamı 1'e eşittir.

$X \downarrow / Y \rightarrow$	y_1	y_2	$\dots$	y_n	$f_X(x)$
x_1	$f(x_1, y_1)$	$f(x_1, y_2)$	$\dots$	$f(x_1, y_n)$	$f_X(x_1)$
x_2	$f(x_2, y_1)$	$f(x_2, y_2)$	$\dots$	$f(x_2, y_n)$	$f_X(x_2)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_m	$f(x_m, y_1)$	$f(x_m, y_2)$	$\dots$	$f(x_m, y_n)$	$f_X(x_m)$
$f_Y(y)$	$f_Y(y_1)$	$f_Y(y_2)$	$\dots$	$f_Y(y_n)$	1

Son sütun, X 'in değerleriyle birlikte X 'in marjinal olasılık fonksiyonunu; son satır, Y 'nin değerleriyle birlikte Y 'nin marjinal olasılık fonksiyonunu verir.

İspat.

i -inci satırın toplamı $\sum_{j=1}^n f(x_i, y_j) = \sum_{y \in D_Y} f_{X,Y}(x_i, y) = f_X(x_i)$ 'dir (Tanım 13.5); j -inci sütunun toplamı da aynı biçimde $f_Y(y_j)$ 'dir. Bütün hücrelerin toplamı, $D_{(X,Y)} \subseteq D_X \times D_Y$ ve $D_{(X,Y)}$ dışında $f = 0$ olduğundan $\sum_{(x,y) \in D_{(X,Y)}} f_{X,Y}(x, y) = 1$ 'dir (Önerme 13.2). Son sütunun (ya da son satırın) toplamının 1 olması Önerme 13.3'nun sonucudur. ■

i Değer kümesi çarpım kümesi olmak zorunda değildir

Tablo, $D_X \times D_Y$ çarpım kümesinin bütün hücrelerini içerir; ama $D_{(X,Y)}$ bu çarpımın yalnızca bir alt kümesidir. Rastgele vektörün alamadığı çiftlere karşılık gelen hücrelere 0 yazılır. Aşağıdaki örnekte $D_{X_1} \times D_{X_2}$ 'nin 12 elemanı varken $D_{(X_1,X_2)}$ 'nin yalnızca 6 elemanı vardır. Tabloda sıfırların bulunması, iki bileşen arasında sıkı bir bağ olduğunun ilk işaretidir; bu gözlem bir sonraki bölümde bağımsızlık kavramıyla kesinleşecektir.

Örnek 13.1 (Üç Para Atışı: Toplam Tura ve İlk İki Atıştaki Tura). Düzgün bir para üç kez atılıyor. X_1 üç atışta gelen tura sayısı, X_2 ilk iki atışta gelen tura sayısı olsun.

- (X_1, X_2) rastgele vektörünün ortak olasılık fonksiyonunu tablo olarak oluşturunuz.
- X_1 ve X_2 'nin marjinal olasılık fonksiyonlarını bulunuz.
- $P(X_1 = 1)$ olasılığını hem doğrudan örnek uzaydan hem de ortak olasılık fonksiyonundan hesaplayınız.
- Üç atışta gelen tura sayısının ilk iki atıştaki tura sayısına eşit olması olasılığını bulunuz.

Çözüm.

a) Y yazıyı, T turayı gösterebilir. Örnek uzay

$$\Omega = \{YYY, YYT, YTY, YTT, TYY, TYT, TTY, TTT\}$$

sekiz elemanlıdır ve para düzgün olduğundan her sonucun olasılığı $\frac{1}{8}$ 'dir. Her sonuç için X_1 (toplam tura) ve X_2 (ilk iki atıştaki tura) değerlerini yazalım:

ω	YYY	YYT	YTY	YTT	TYY	TYT	TTY	TTT
$X_1(\omega)$	0	1	1	2	1	2	2	3
$X_2(\omega)$	0	0	1	1	1	1	2	2

Buradan $D_{X_1} = \{0, 1, 2, 3\}$, $D_{X_2} = \{0, 1, 2\}$ ve

$$D_{(X_1,X_2)} = \{(0,0), (1,0), (1,1), (2,1), (2,2), (3,2)\}$$

okunur. Her çiftin olasılığı, o çifti veren sonuçların sayısının $\frac{1}{8}$ katıdır. Örneğin

$$f_{X_1,X_2}(0,0) = P(X_1 = 0, X_2 = 0) = P(\{YYY\} \cap \{YYY, YYT\}) = P(\{YYY\}) = \frac{1}{8},$$

$$f_{X_1,X_2}(1,1) = P(X_1 = 1, X_2 = 1) = P(\{YTY, TYY\}) = \frac{2}{8},$$

$$f_{X_1,X_2}(2,1) = P(\{YTT, TYT\}) = \frac{2}{8}, \quad f_{X_1,X_2}(2,2) = P(\{TTY\}) = \frac{1}{8}.$$

$D_{(X_1,X_2)}$ dışındaki çiftler için $f_{X_1,X_2} = 0$ 'dır; örneğin $X_1 = 0$ iken $X_2 = 1$ olamaz. Ortak olasılık fonksiyonunun tablosu (satır ve sütun toplamlarıyla birlikte) şöyledir:

$X_1 \downarrow / X_2 \rightarrow$	0	1	2	$f_{X_1}(x_1)$
0	$\frac{1}{8}$	0	0	$\frac{1}{8}$
1	$\frac{1}{8}$	$\frac{2}{8}$	0	$\frac{3}{8}$
2	0	$\frac{2}{8}$	$\frac{1}{8}$	$\frac{3}{8}$
3	0	0	$\frac{1}{8}$	$\frac{1}{8}$
$f_{X_2}(x_2)$	$\frac{2}{8}$	$\frac{4}{8}$	$\frac{2}{8}$	1

Hücrelerin toplamı $\frac{1+1+2+2+1+1}{8} = 1$ 'dir; **Önerme 13.2** sağlar.

b) Satır toplamları X_1 'in marjinal olasılık fonksiyonunu verir:

x_1	0	1	2	3
$f_{X_1}(x_1)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Örneğin $f_{X_1}(2) = f_{X_1, X_2}(2, 0) + f_{X_1, X_2}(2, 1) + f_{X_1, X_2}(2, 2) = 0 + \frac{2}{8} + \frac{1}{8} = \frac{3}{8}$. Bu, **Örnek 8.2**'nde bulunan üç atıştaki tura sayısı dağılımıdır: $f_{X_1}(x_1) = \binom{3}{x_1} \frac{1}{8}$.

Sütun toplamaları X_2 'nin marjinal olasılık fonksiyonunu verir:

x_2	0	1	2
$f_{X_2}(x_2)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Bu da iki atışta gelen tura sayısının dağılımıdır: $f_{X_2}(x_2) = \binom{2}{x_2} \frac{1}{4}$. İki marjinalin toplamı da 1'dir.

c) Doğrudan: $(X_1 = 1) = \{YYT, YTY, TYY\}$ olduğundan $P(X_1 = 1) = \frac{3}{8}$.

Ortak olasılık fonksiyonundan: X_2 'nin bütün değerleri üzerinden toplayarak (**Tanım 13.5**)

$$P(X_1 = 1) = \sum_{x_2 \in D_{X_2}} f_{X_1, X_2}(1, x_2) = f_{X_1, X_2}(1, 0) + f_{X_1, X_2}(1, 1) + f_{X_1, X_2}(1, 2) = \frac{1}{8} + \frac{2}{8} + 0 = \frac{3}{8}.$$

Bu, tablodaki $x_1 = 1$ satırının toplamı, yani $f_{X_1}(1) = \frac{3}{8}$ 'dir; iki yol aynı sonucu verir.

d) İstenen olasılık $P(X_1 = X_2)$ 'dir. $(X_1 = X_2)$ olayı, $B = \{(u, v) \in \mathbb{R}^2 : u = v\}$ köşegeni için $((X_1, X_2) \in B)$ olayıdır; **Önerme 13.2** gereği $B \cap D_{(X_1, X_2)} = \{(0, 0), (1, 1), (2, 2)\}$ üzerinden toplanır:

$$P(X_1 = X_2) = f_{X_1, X_2}(0, 0) + f_{X_1, X_2}(1, 1) + f_{X_1, X_2}(2, 2) = \frac{1}{8} + \frac{2}{8} + \frac{1}{8} = \frac{1}{2}.$$

Sonuç sağduyuyla uyumludur: toplam tura sayısının ilk iki atıştaki tura sayısına eşit olması, üçüncü atışın yazı gelmesi demektir; bunun olasılığı $\frac{1}{2}$ 'dir.

■

13.4 Sürekli Rastgele Vektörler

Tek boyutta sürekli rastgele değişken, değer kümesi sayılamayan değişkendi (**Tanım 11.1**); olasılık hesabını, dağılım fonksiyonu bir yoğunluğun integrali olarak yazılabilen mutlak sürekli değişkenler için yapmıştık (**Tanım 11.2**). İki boyutta yoğunluk, düzlem üzerinde tanımlı ve toplam kütlesi 1 olan negatif olmayan bir fonksiyondur; olasılıklar bu fonksiyonun bölgeler üzerindeki çift katlı integralleridir.

Tanım 13.6 (İki Değişkenli Olasılık Yoğunluk Fonksiyonu). $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ fonksiyonu

(i) her $(x_1, x_2) \in \mathbb{R}^2$ için $f(x_1, x_2) \geq 0$,

(ii) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2) dx_1 dx_2 = 1$

koşullarını sağlıyorsa f 'ye **iki değişkenli olasılık yoğunluk fonksiyonu** denir.

Tanım 13.7 (Sürekli Rastgele Vektör ve Ortak Yoğunluk). (X_1, X_2) bir rastgele vektör olsun. Ortak dağılım fonksiyonu, iki değişkenli bir olasılık yoğunluk fonksiyonu f yardımıyla her $(x_1, x_2) \in \mathbb{R}^2$ için

$$F_{X_1, X_2}(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2) = \int_{-\infty}^{x_2} \int_{-\infty}^{x_1} f(u, v) du dv$$

biçiminde yazılabiliyorsa (X_1, X_2) 'ye **sürekli rastgele vektör** (mutlak sürekli rastgele vektör) ve f 'ye (X_1, X_2) rastgele vektörünün **olasılık yoğunluk fonksiyonu** ya da X_1 ile X_2 'nin **ortak olasılık yoğunluk fonksiyonu** denir; f_{X_1, X_2} ile gösterilir. Yoğunluğun sıfırdan farklı olduğu $\{(x_1, x_2) : f_{X_1, X_2}(x_1, x_2) > 0\}$ kümesine yoğunluğun **tanım bölgesi** (desteği) denir; bu küme düzlemde sayılamayan bir kümedir.

Sürekli vektörlerde olasılık hesabının bütün pratiği tek bir ilkeye dayanır: **bir bölgenin olasılığı, yoğunluğun o bölge üzerindeki integralidir.** Aşağıdaki önerme bunu ve doğrudan sonuçlarını toplar.

Önerme 13.5 (Sürekli Rastgele Vektörde Olasılık Hesabı). (X, Y) sürekli bir rastgele vektör ve $f = f_{X, Y}$ ortak yoğunluğu olsun.

(i) $a < b, c < d$ için $P(a < X \leq b, c < Y \leq d) = \int_c^d \int_a^b f(x, y) dx dy$.

(ii) Her $B \in \mathcal{B}(\mathbb{R}^2)$ için $P((X, Y) \in B) = \iint_B f(x, y) dx dy$.

(iii) Her (x_0, y_0) için $P(X = x_0, Y = y_0) = 0$; daha genel olarak her doğru parçasının olasılığı 0'dır. Bu yüzden (i)'deki eşitsizliklerde $<$ ile $\leq$ birbirinin yerine kullanılabilir.

(iv) f 'nin sürekli olduğu her (x, y) noktasında $\frac{\partial^2 F_{X, Y}}{\partial x \partial y}(x, y) = f(x, y)$.

İspat.

(i) Dikdörtgen formülü (Önerme 13.1) ve $F_{X, Y}$ 'nin integral gösterimi kullanılır. $F(b, d) - F(a, d)$ farkı, $\int_{-\infty}^d (\int_{-\infty}^b f dx - \int_{-\infty}^a f dx) dy = \int_{-\infty}^d \int_a^b f dx dy$ 'dir. Aynı biçimde $F(b, c) - F(a, c) = \int_{-\infty}^c \int_a^b f dx dy$ olur. İkisinin farkı

$$F(b, d) - F(a, d) - F(b, c) + F(a, c) = \int_c^d \int_a^b f(x, y) dx dy$$

verir.

(ii) (i), iddiayı yarı açık dikdörtgenler için ispatlar. Dikdörtgenlerin sonlu ayrık birleşimleri için sonuç, olasılığın ve integralin toplamsallığından çıkar. Genel bir Borel kümesine geçiş, her Borel kümesine bu tür birleşimlerle istenildiği kadar iyi yaklaşılabilmesine ve iki tarafın da sayılabilir toplamsal ölçüler olmasına dayanır; iki olasılık ölçüsü $\mathcal{B}(\mathbb{R}^2)$ 'yi üreten dikdörtgenler sınıfı üzerinde çakışıyorsa bütün Borel kümelerinde çakışır. Bu ölçü teorisi sonucunu ispatsız kullanıyoruz.

(iii) $\varepsilon > 0$ için (i) ve monotonluk gereği

$$P(X = x_0, Y = y_0) \leq P(x_0 - \varepsilon < X \leq x_0, y_0 - \varepsilon < Y \leq y_0) = \int_{y_0 - \varepsilon}^{y_0} \int_{x_0 - \varepsilon}^{x_0} f dx dy$$

ve $\varepsilon \rightarrow 0$ iken sağ taraf 0'a gider (integrallenebilir bir fonksiyonun, alanı sıfıra giden bölgeler üzerindeki integrali sıfıra gider). Yatay bir doğru parçası $\{(x, y_0) : a < x \leq b\}$ için de aynı sıkıştırma $\int_{y_0 - \varepsilon}^{y_0} \int_a^b f dx dy \rightarrow 0$ verir; dikey parça için benzerdir. Eğik bir doğru parçası, toplam alanı istenildiği kadar küçük olan sonlu sayıda dikdörtgenle örtülebildiğinden onun olasılığı da 0'dır. Bir dikdörtgenin kenarlarını dahil etmek ya da çıkarmak yalnızca sıfır olasılıklı doğru parçaları ekleyip çıkardığından olasılık değişmez.

(iv) $F_{X, Y}(x, y) = \int_{-\infty}^y (\int_{-\infty}^x f(u, v) du) dv$ yazımında önce y 'ye göre türev alınırsa analizin temel teoremi $\frac{\partial F}{\partial y}(x, y) = \int_{-\infty}^x f(u, y) du$ verir; ardından x 'e göre türev alınırsa aynı teoreme $f(x, y)$ elde edilir. Ne var ki bu ardışık türev alma, iç integralin ikinci değişkende sürekli olmasını gerektirir; bu, f 'nin yalnızca (x, y) 'nin bir komşuluğunda sürekli olmasından çıkmaz. Bu yüzden hesabı doğrudan dikdörtgen formülü üzerinden yürütelim: $h, k > 0$ için Önerme 13.1 (v) ve (i) gereği

$$\frac{F(x+h, y+k) - F(x, y+k) - F(x+h, y) + F(x, y)}{hk} = \frac{1}{hk} \int_y^{y+k} \int_x^{x+h} f(u, v) du dv$$

olur ve f 'nin (x, y) 'deki sürekliliği gereği sağ taraf $h, k \rightarrow 0$ iken $f(x, y)$ 'ye gider. Bu, tek değişkenli $F'_X = f_X$ ilişkisinin (Teorem 11.2) iki boyutlu karşılığıdır.

■

13.4.1 Marjinal Yoğunluk Fonksiyonları

Kesikli durumda X 'in olasılık fonksiyonu, ortak olasılık fonksiyonunu y üzerinden toplayarak bulunuyordu. Sürekli durumda toplamın yerini integral alır.

Tanım 13.8 (Marjinal Olasılık Yoğunluk Fonksiyonları). (X, Y) sürekli bir rastgele vektör ve $f_{X,Y}$ ortak yoğunluğu olsun.

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy, \quad x \in \mathbb{R}; \quad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx, \quad y \in \mathbb{R}$$

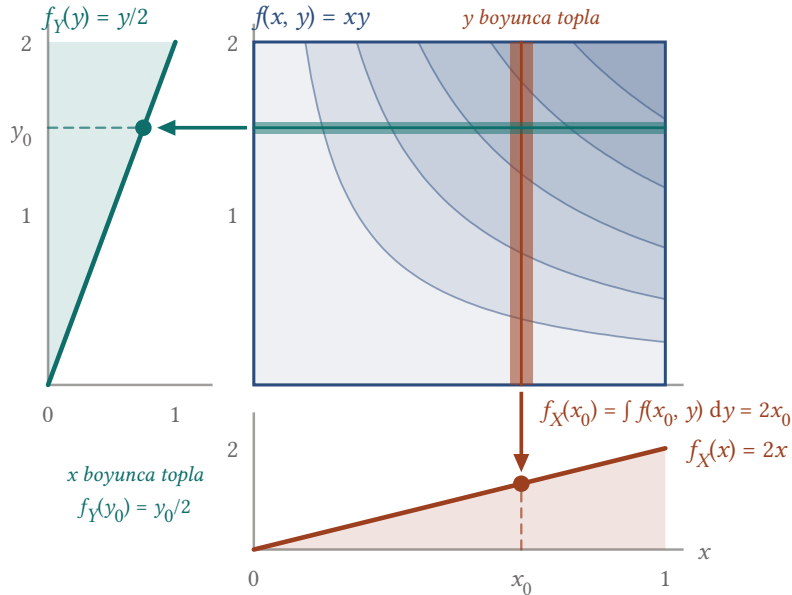
fonskiyonlarına sırasıyla X 'in ve Y 'nin **marjinal olasılık yoğunluk fonksiyonları** denir.

Ortak yoğunluk, tanım bölgesi dışında sıfır olduğundan bu integraller pratikte yalnızca tanım bölgesinin ilgili kesiti üzerinden alınır: $f_X(x)$ için x sabitlenir ve (x, y) 'nin bölgede kaldığı y değerleri üzerinden integral alınır. Adlandırmanın haklı olduğunu, yani marjinal yoğunlukların gerçekten X ve Y 'nin yoğunluk fonksiyonları olduğunu gösterelim.

Önerme 13.6 (Marjinal Yoğunluklar Birer Yoğunluk Fonksiyonudur). (X, Y) sürekli bir rastgele vektör olsun. X ve Y yoğunluğa sahip (mutlak sürekli) rastgele değişkenlerdir ve yoğunluk fonksiyonları sırasıyla f_X ile f_Y marjinal yoğunluklarıdır; yani her $x, y \in \mathbb{R}$ için

$$F_X(x) = \int_{-\infty}^x f_X(u) du, \quad F_Y(y) = \int_{-\infty}^y f_Y(v) dv.$$

Özel olarak $f_X \geq 0$, $f_Y \geq 0$ ve $\int_{-\infty}^{\infty} f_X(x) dx = \int_{-\infty}^{\infty} f_Y(y) dy = 1$ 'dir.



Şekil 13.3: Marjinal yoğunluk, öteki değişkeni integrale yok ederek bulunur. Ortadaki dikdörtgen $f(x, y) = xy$ ortak yoğunluğunun tanım bölgesidir; köşeye doğru koyulaşan gölge ve $xy = \text{sabit}$ düzey eğrileri yoğunluğun büyüdüğü yönü gösterir. $x = x_0$ düşey şeridi üzerinde y 'ye göre integral alınınca alttaki eğrinin bir noktası, $f_X(x_0) = 2x_0$, çıkar; $y = y_0$ yatay şeridi üzerinde x 'e göre integral almak da soldaki $f_Y(y) = y/2$ eğrisinin bir noktasını verir.

İspat.

$f = f_{X,Y}$ yazalım. $f \geq 0$ olduğundan $f_X(x) = \int_{-\infty}^{\infty} f(x,y) dy \geq 0$ 'dır. **Önerme 13.1** (iii) gereği $F_X(x) = \lim_{y \rightarrow \infty} F_{X,Y}(x,y)$ 'dir. Yoğunluk gösterimini yazıp negatif olmayan integrandlar için integral sırasını değiştirirsek

$$F_{X,Y}(x,y) = \int_{-\infty}^y \int_{-\infty}^x f(u,v) du dv = \int_{-\infty}^x \left(\int_{-\infty}^y f(u,v) dv \right) du$$

olur. $y \rightarrow \infty$ iken iç integral artarak $\int_{-\infty}^{\infty} f(u,v) dv = f_X(u)$ 'ya gider ve limit dış integralin içine alınabilir (artan negatif olmayan fonksiyon dizileri için limit ile integral yer değiştirir). Böylece

$$F_X(x) = \lim_{y \rightarrow \infty} F_{X,Y}(x,y) = \int_{-\infty}^x f_X(u) du.$$

Bu, X 'in dağılım fonksiyonunun negatif olmayan bir fonksiyonun integrali olarak yazılabildiğini söyler; yani X mutlak sürekli bir rastgele değişkendir ve yoğunluğu f_X 'tir (**Tanım 11.2**). $x \rightarrow \infty$ alınırsa $\int_{-\infty}^{\infty} f_X(u) du = \lim_{x \rightarrow \infty} F_X(x) = 1$ bulunur; bu aynı zamanda $\iint f = 1$ eşitliğinin integral sırası değiştirilerek okunmasıdır. Y için ispat, değişkenlerin rolleri değiştirilerek aynen yürür.

■

 Tanım bölgesini çizmeden integral sınırı yazmayın

Sürekli vektörlerde yapılan hataların hemen hepsi integral sınırlarından kaynaklanır. Bir olasılık hesaplamadan önce ortak yoğunluğun tanım bölgesini düzlemde çizin, istenen olayın bölgesini üzerine tarayın ve integrali yalnızca **iki bölgenin kesişimi** üzerinden alın. Marjinal yoğunluk hesaplarırken de x 'i sabitleyip bölgeyi dikey bir doğruyla kesin; y 'nin sınırları bu kesitten okunur. Bölge bir dikdörtgense sınırlar sabittir; değilse en az bir sınır öteki değişkene bağlıdır.

Örnek 13.2 (Dikdörtgen Üzerinde Sabit Ortak Yoğunluk). X ve Y rastgele değişkenlerinin ortak olasılık yoğunluk fonksiyonu

$$f_{X,Y}(x,y) = \begin{cases} c, & 0 < x < 2, 2 < y < 4 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun.

- $f_{X,Y}$ 'nin bir olasılık yoğunluk fonksiyonu olması için c ne olmalıdır?
- X ve Y 'nin marjinal olasılık yoğunluk fonksiyonlarını bulunuz.
- $P(X \leq 1, Y \leq 3)$ olasılığını hesaplayınız ve ortak dağılım fonksiyonunu $0 < x < 2, 2 < y < 4$ için yazınız.
- $P(X + Y \leq 3)$ olasılığını hesaplayınız.

Çözüm.

Yoğunluğun tanım bölgesi $D = (0, 2) \times (2, 4)$ dikdörtgenidir; alanı $2 \cdot 2 = 4$ 'tür.

a) $f_{X,Y} \geq 0$ olması için $c \geq 0$ gerekir. Toplam integralin 1 olması koşulundan (**Tanım 13.6**)

$$1 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx dy = \int_2^4 \int_0^2 c dx dy = c \int_2^4 2 dy = 4c$$

bulunur; öyleyse $c = \frac{1}{4}$. Bu, dikdörtgenin alanının tersidir: (X, Y) , D dikdörtgeni üzerinde **düzgün dağılmıştır** ve D içindeki bir bölgenin olasılığı, o bölgenin alanının D 'nin alanına oranıdır. Bu, daha önce gördüğümüz geometrik olasılık modelidir (**Tanım 4.5**).

b) $0 < x < 2$ sabit olsun; $(x, y) \in D$ olması için $2 < y < 4$ gerekir. **Tanım 13.8** gereği

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy = \int_2^4 \frac{1}{4} dy = \frac{1}{2}, \quad 0 < x < 2,$$

ve $x \notin (0, 2)$ için $f_X(x) = 0$ 'dır. Benzer biçimde $2 < y < 4$ için

$$f_Y(y) = \int_0^2 \frac{1}{4} dx = \frac{1}{2}, \quad 2 < y < 4,$$

ve $y \notin (2, 4)$ için $f_Y(y) = 0$. Yani

$$f_X(x) = \begin{cases} \frac{1}{2}, & 0 < x < 2 \\ 0, & \text{diğer durumlarda,} \end{cases} \quad f_Y(y) = \begin{cases} \frac{1}{2}, & 2 < y < 4 \\ 0, & \text{diğer durumlarda.} \end{cases}$$

X , $(0, 2)$ üzerinde; Y , $(2, 4)$ üzerinde düzgün dağılmıştır. Her iki marjinalin integrali $2 \cdot \frac{1}{2} = 1$ 'dir (Önerme 13.6).

c) $(X \leq 1, Y \leq 3)$ olayının bölgesi $\{x \leq 1, y \leq 3\}$ çeyrek düzlemdir; D ile kesişimi $(0, 1] \times (2, 3]$ dikdörtgenidir. Önerme 13.5 (ii) gereği

$$P(X \leq 1, Y \leq 3) = \int_2^3 \int_0^1 \frac{1}{4} dx dy = \frac{1}{4} \cdot 1 \cdot 1 = \frac{1}{4}.$$

Alan yorumuyla: kesişimin alanı 1, D 'nin alanı 4; oran $\frac{1}{4}$. Bu değer $F_{X,Y}(1, 3)$ 'tür. Genel olarak $0 < x < 2, 2 < y < 4$ için

$$F_{X,Y}(x, y) = \int_2^y \int_0^x \frac{1}{4} du dv = \frac{x(y-2)}{4}$$

bulunur; $x = 1, y = 3$ için gerçekten $\frac{1}{4}$ çıkar. Doğrulama olarak $\frac{\partial^2}{\partial x \partial y} \frac{x(y-2)}{4} = \frac{1}{4} = f_{X,Y}(x, y)$ 'dir (Önerme 13.5 (iv)).

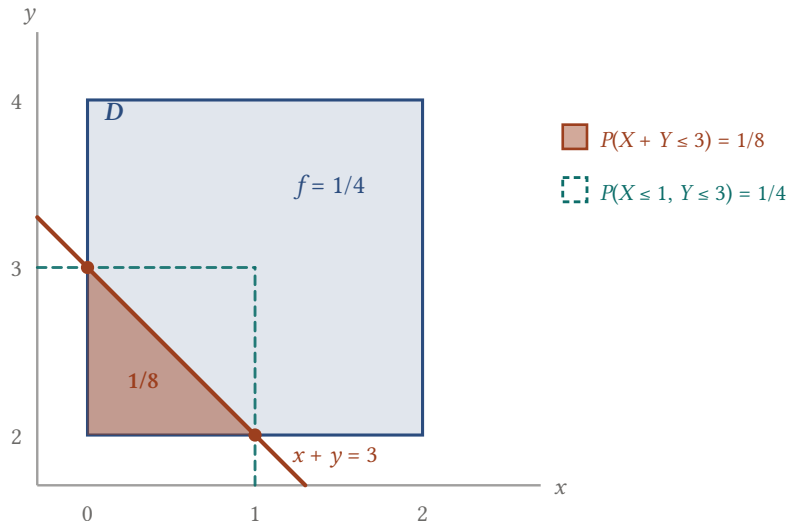
d) $(X + Y \leq 3)$ olayının bölgesi $x + y \leq 3$ yarı düzlemdir. D ile kesişimini bulalım: $y > 2$ ve $y \leq 3 - x$ olması için $3 - x > 2$, yani $x < 1$ gerekir. Demek ki kesişim

$$\{(x, y) : 0 < x < 1, 2 < y \leq 3 - x\}$$

üçgenidir; köşeleri $(0, 2)$, $(1, 2)$ ve $(0, 3)$ 'tür. İntegral

$$P(X + Y \leq 3) = \int_0^1 \int_2^{3-x} \frac{1}{4} dy dx = \frac{1}{4} \int_0^1 (3 - x - 2) dx = \frac{1}{4} \int_0^1 (1 - x) dx = \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8}.$$

Alan yorumuyla: dik kenarları 1 olan üçgenin alanı $\frac{1}{2}$, D 'nin alanı 4; oran $\frac{1}{8}$.



Şekil 13.4: Dikdörtgen üzerinde sabit ortak yoğunluk. Açık taralı $D = (0, 2) \times (2, 4)$ dikdörtgeni yoğunluğun $1/4$ değerini aldığı bölgedir; D içindeki bir bölgenin olasılığı, alanının 4 'e oranıdır. $x + y = 3$ doğrusunun altında kalan koyu üçgenin alanı $1/2$ olduğundan $P(X + Y \leq 3) = 1/8$; kesikli çizgilerle sınırlanan $(0, 1] \times (2, 3]$ karesinin alanı 1 olduğundan $P(X \leq 1, Y \leq 3) = 1/4$.

Kesikli örnekte tabloda beliren sıfırlar, iki bileşen arasında sıkı bir bağ olduğunu gösteriyordu; bu örnekte ise tersi bir durum gözlenir: ortak yoğunluk, marjinal yoğunlukların çarpımına eşittir ($\frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2}$). Bu eşitliğin tesadüf olmadığı ve “bağımsızlık” anlamına geldiği bir sonraki bölümde gösterilecektir.

13.5 Alıştırmalar

- Alıştırma 13.1 (Rastgele Vektörler ve Ortak Dağılımlar Üzerine).** a) Düzgün bir zar iki kez atılıyor. X gelen sayıların en küçüğü, Y en büyüğü olsun. (X, Y) 'nin ortak olasılık fonksiyonunu formülle yazınız, marjinal olasılık fonksiyonlarını bulunuz ve $P(X = Y)$ ile $P(Y - X \leq 1)$ olasılıklarını hesaplayınız.
- b) X ve Y rastgele değişkenlerinin ortak olasılık fonksiyonu $x \in \{1, 2, 3\}$, $y \in \{1, 2\}$ için $f_{X,Y}(x, y) = kxy$ olsun. k sabitini, marjinal olasılık fonksiyonlarını ve $P(X + Y \leq 3)$ olasılığını bulunuz.
- c) Örnek 13.1'teki (X_1, X_2) vektörü için $F_{X_1, X_2}(1, 1)$ ve $F_{X_1, X_2}(2, 0)$ değerlerini bulunuz. Ardından $P(0 < X_1 \leq 2, 0 < X_2 \leq 1)$ olasılığını hem dikdörtgen formülüyle hem de tablodan doğrudan hesaplayıp karşılaştırınız.
- d) (X, Y) rastgele vektörünün ortak dağılım fonksiyonu $x \geq 0, y \geq 0$ için $F_{X,Y}(x, y) = (1 - e^{-x})(1 - e^{-y})$, diğer durumlarda 0 olsun. Marjinal dağılım fonksiyonlarını, ortak yoğunluk fonksiyonunu ve $P(1 < X \leq 2, 0 < Y \leq 1)$ olasılığını bulunuz.
- e) X ve Y 'nin ortak yoğunluğu $0 < x < 1, 0 < y < 2$ için $f_{X,Y}(x, y) = kxy$, diğer durumlarda 0 olsun. k 'yı, marjinal yoğunlukları, $P(X \leq \frac{1}{2}, Y \leq 1)$ ve $P(X + Y \leq 1)$ olasılıklarını bulunuz.
- f) X ve Y 'nin ortak yoğunluğu $0 < x < y < 1$ için $f_{X,Y}(x, y) = 2$, diğer durumlarda 0 olsun. Bunun bir yoğunluk fonksiyonu olduğunu gösteriniz, marjinal yoğunlukları bulunuz ve $P(Y > 2X)$ olasılığını hesaplayınız.

Çözüm.

- a) Örnek uzay $\{1, \dots, 6\}^2$ 'nin 36 elemanı eşit olasılıklıdır. $x < y$ olan bir (x, y) çifti iki sonuçtan $((x, y)$ ve $(y, x))$ gelir; $x = y$ çifti tek sonuçtan gelir; $x > y$ olamaz. Öyleyse $x, y \in \{1, \dots, 6\}$ için

$$f_{X,Y}(x,y) = \begin{cases} \frac{2}{36}, & x < y \\ \frac{1}{36}, & x = y \\ 0, & x > y. \end{cases}$$

Marjinaler: x sabitken $y \in \{x, x+1, \dots, 6\}$ olduğundan

$$f_X(x) = \frac{1}{36} + (6-x)\frac{2}{36} = \frac{13-2x}{36}, \quad x = 1, \dots, 6;$$

y sabitken $x \in \{1, \dots, y\}$ olduğundan

$$f_Y(y) = \frac{1}{36} + (y-1)\frac{2}{36} = \frac{2y-1}{36}, \quad y = 1, \dots, 6.$$

Sağlama: $\sum_{x=1}^6 (13-2x) = 78 - 42 = 36$ ve $\sum_{y=1}^6 (2y-1) = 36$. Olasılıklar: $P(X=Y) = 6 \cdot \frac{1}{36} = \frac{1}{6}$. $(Y-X \leq 1)$ olayı $x=y$ olan 6 çift ile $y=x+1$ olan 5 çiftten oluşur: $P(Y-X \leq 1) = \frac{6}{36} + 5 \cdot \frac{2}{36} = \frac{16}{36} = \frac{4}{9}$.

b) $\sum_{x=1}^6 \sum_{y=1}^2 kxy = k(1+2+3)(1+2) = 18k = 1$ olduğundan $k = \frac{1}{18}$. Marjinaler:

$$f_X(x) = \frac{x}{18}(1+2) = \frac{x}{6}, \quad x = 1, 2, 3; \quad f_Y(y) = \frac{y}{18}(1+2+3) = \frac{y}{3}, \quad y = 1, 2.$$

$X+Y \leq 3$ koşulunu sağlayan çiftler $(1,1)$, $(1,2)$, $(2,1)$ 'dir:

$$P(X+Y \leq 3) = \frac{1+2+2}{18} = \frac{5}{18}.$$

c) Tabloyu kullanalım. $F_{X_1, X_2}(1,1) = P(X_1 \leq 1, X_2 \leq 1)$ değeri, $x_1 \in \{0,1\}$ ve $x_2 \in \{0,1\}$ hücrelerinin toplamıdır: $\frac{1}{8} + 0 + \frac{1}{8} + \frac{2}{8} = \frac{1}{2}$. $F_{X_1, X_2}(2,0)$, $x_1 \in \{0,1,2\}$, $x_2 = 0$ hücrelerinin toplamıdır: $\frac{1}{8} + \frac{1}{8} + 0 = \frac{1}{4}$.

Dikdörtgen formülü için ayrıca $F(2,1) = \frac{1}{8} + 0 + \frac{1}{8} + \frac{2}{8} + 0 + \frac{2}{8} = \frac{3}{4}$, $F(0,1) = \frac{1}{8} + 0 = \frac{1}{8}$ ve $F(0,0) = \frac{1}{8}$ gerekir. **Önerme 13.1** (v) ile

$$P(0 < X_1 \leq 2, 0 < X_2 \leq 1) = F(2,1) - F(0,1) - F(2,0) + F(0,0) = \frac{3}{4} - \frac{1}{8} - \frac{1}{4} + \frac{1}{8} = \frac{1}{2}.$$

Doğrudan: olay $x_1 \in \{1,2\}$, $x_2 = 1$ hücrelerinden oluşur; $f(1,1) + f(2,1) = \frac{2}{8} + \frac{2}{8} = \frac{1}{2}$. İki yol uyuyor.

d) **Önerme 13.1** (iii) ile $x \geq 0$ için $F_X(x) = \lim_{y \rightarrow \infty} (1 - e^{-x})(1 - e^{-y}) = 1 - e^{-x}$, $x < 0$ için 0; simetriyle $F_Y(y) = 1 - e^{-y}$ ($y \geq 0$). Ortak yoğunluk karma türevdir (**Önerme 13.5** (iv)): $x, y > 0$ için

$$f_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} [(1 - e^{-x})(1 - e^{-y})] = e^{-x}e^{-y} = e^{-(x+y)},$$

diğer durumlarda 0. Bunun negatif olmadığı ve $\int_0^\infty \int_0^\infty e^{-(x+y)} dx dy = 1 \cdot 1 = 1$ olduğu görülür. **Dikdörtgen formülüyle**

$$P(1 < X \leq 2, 0 < Y \leq 1) = F(2,1) - F(1,1) - F(2,0) + F(1,0) = (1 - e^{-2})(1 - e^{-1}) - (1 - e^{-1})^2 - 0 + 0.$$

$(1 - e^{-1})$ ortak çarpanı alınırsa sonuç şu değere iner:

$$(1 - e^{-1}) [(1 - e^{-2}) - (1 - e^{-1})] = (1 - e^{-1})(e^{-1} - e^{-2}) = e^{-1}(1 - e^{-1})^2 \approx 0,147.$$

Aynı sonuç $\int_0^1 \int_1^2 e^{-(x+y)} dx dy = (e^{-1} - e^{-2})(1 - e^{-1})$ integralinden de çıkar.

e) $\int_0^2 \int_0^1 kxy dx dy = k \cdot \frac{1}{2} \cdot 2 = k$ olduğundan $k = 1$. Marjinaler: $0 < x < 1$ için $f_X(x) = \int_0^2 xy dy = 2x$; $0 < y < 2$ için $f_Y(y) = \int_0^1 xy dx = \frac{y}{2}$; bölge dışında 0. Sağlama: $\int_0^1 2x dx = 1$, $\int_0^2 \frac{y}{2} dy = 1$.

$$P\left(X \leq \frac{1}{2}, Y \leq 1\right) = \int_0^1 \int_0^{1/2} xy \, dx \, dy = \frac{1}{8} \cdot \frac{1}{2} = \frac{1}{16}.$$

$(X + Y \leq 1)$ olayının bölge ile kesişimi $0 < x < 1, 0 < y \leq 1 - x$ üçgenidir:

$$P(X + Y \leq 1) = \int_0^1 \int_0^{1-x} xy \, dy \, dx = \int_0^1 \frac{x(1-x)^2}{2} \, dx = \frac{1}{2} \int_0^1 (x - 2x^2 + x^3) \, dx = \frac{1}{2} \left(\frac{1}{2} - \frac{2}{3} + \frac{1}{4} \right) = \frac{1}{24}.$$

f) Tanım bölgesi, köşeleri $(0, 0)$, $(0, 1)$, $(1, 1)$ olan üçgendir; alanı $\frac{1}{2}$. $f \geq 0$ ve $\iint f = 2 \cdot \frac{1}{2} = 1$ olduğundan f bir yoğunluk fonksiyonudur. Marjinaler: $0 < x < 1$ sabitken $y \in (x, 1)$ olduğundan $f_X(x) = \int_x^1 2 \, dy = 2(1 - x)$; $0 < y < 1$ sabitken $x \in (0, y)$ olduğundan $f_Y(y) = \int_0^y 2 \, dx = 2y$; bölge dışında 0. Sağlama: $\int_0^1 2(1 - x) \, dx = 1$, $\int_0^1 2y \, dy = 1$.

$(Y > 2X)$ olayı için bölgede $2x < y < 1$ olmalı; bu, $x < \frac{1}{2}$ gerektirir ($y > 2x$ zaten $y > x$ 'i sağlar). Öyleyse

$$P(Y > 2X) = \int_0^{1/2} \int_{2x}^1 2 \, dy \, dx = \int_0^{1/2} 2(1 - 2x) \, dx = 2 \left[x - x^2 \right]_0^{1/2} = 2 \left(\frac{1}{2} - \frac{1}{4} \right) = \frac{1}{2}.$$

Alan yorumuyla: $y = 2x$ doğrusu üçgeni, alanları $\frac{1}{4}$ olan iki parçaya böler; istenen olasılık $2 \cdot \frac{1}{4} = \frac{1}{2}$ 'dir. ■

Ortak dağılımı bilmek, iki rastgele değişkenin **birlikte** davranışını bilmek demektir; bir sonraki bölümde bu bilgiyi iki soruya yöneltiliyoruz: ortak dağılım ne zaman marjinalerin çarpımıdır ve biri gözlemlendiğinde ötekinin dağılımı nasıl güncellenir: [Bağımsız Rastgele Değişkenler ve Koşullu Dağılımlar](#).

14. Bağımsız Rastgele Değişkenler ve Koşullu Dağılımlar

Rastgele Vektörler ve Ortak Dağılımlar bölümünde iki rastgele değişkeni bir arada ele almayı öğrendik: ortak olasılık (yoğunluk) fonksiyonu vektörün taşıdığı bütün olasılık bilgisini içerir; marjinal fonksiyonlar ise bu bilgidan tek bir bileşene ait olanı süzer. Ortaktan marjinala geçiş her zaman mümkündür. Ters yön ise genel olarak mümkün değildir: aynı marjinallere sahip pek çok farklı ortak dağılım vardır, dolayısıyla marjinaller bilinince ortak fonksiyon geri kurulamaz.

Bu bölümde bu boşluğu dolduran iki kavramı inceliyoruz. Birincisi **bağımsızlık**: ortak fonksiyonun marjinallerin çarpımına eşit olduğu, yani bir bileşenin aldığı değer diğer bileşen hakkında hiçbir bilgi taşımadığı özel durum. İkincisi **koşullu dağılım**: bir bileşenin değeri bilindiğinde diğer bileşenin dağılımının nasıl güncellendiğini anlatan nesne. Bu ikisi birbirinin aynası gibidir: koşullu dağılım marjinala eşitse bağımsızlık vardır; bağımsızlık varsa koşullu dağılım marjinala eşittir.

Kesikli durumda koşullu olasılık fonksiyonu, koşullu olasılık tanımının doğrudan sonucudur; sürekli durumda $P(X = x) = 0$ olduğundan koşullu yoğunluk önce bir limit düşüncesiyle gerekçelendirilir, sonra tanım olarak konur. Bölümün ikinci yarısı bu araçları kesikli ve sürekli örneklerde kullanır.

14.1 Bağımsız Rastgele Değişkenler

İki olayın bağımsızlığını (Tanım 7.1) $P(A \cap B) = P(A)P(B)$ eşitliğiyle tanımlamıştık. Rastgele değişkenler için aynı fikri kullanacağız: X_1 'in değeriyle ilgili her olay, X_2 'nin değeriyle ilgili her olaydan bağımsız olsun. Bu isteğin ortak fonksiyon diliyle karşılığı, ortak fonksiyonun marjinallerin çarpımı olmasıdır.

Tanım 14.1 (Bağımsız Rastgele Değişkenler). (X_1, X_2) bir rastgele vektör olsun.

(1) **Sürekli durum.** (X_1, X_2) sürekli bir rastgele vektör (Tanım 13.7), ortak olasılık yoğunluk fonksiyonu f_{X_1, X_2} ve marjinal olasılık yoğunluk fonksiyonları f_{X_1}, f_{X_2} olsun. Her $(x_1, x_2) \in \mathbb{R}^2$ için

$$f_{X_1, X_2}(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2)$$

ise X_1 ve X_2 sürekli rastgele değişkenlerine **bağımsız** denir.

(2) **Kesikli durum.** (X_1, X_2) kesikli bir rastgele vektör (Tanım 13.3) olsun. Her $(x_1, x_2) \in \mathbb{R}^2$ için

$$P(X_1 = x_1, X_2 = x_2) = P(X_1 = x_1)P(X_2 = x_2), \quad \text{yani} \quad f_{X_1, X_2}(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2)$$

ise X_1 ve X_2 kesikli rastgele değişkenlerine **bağımsız** denir.

Bağımsız olmayan rastgele değişkenlere **bağımlı** denir.

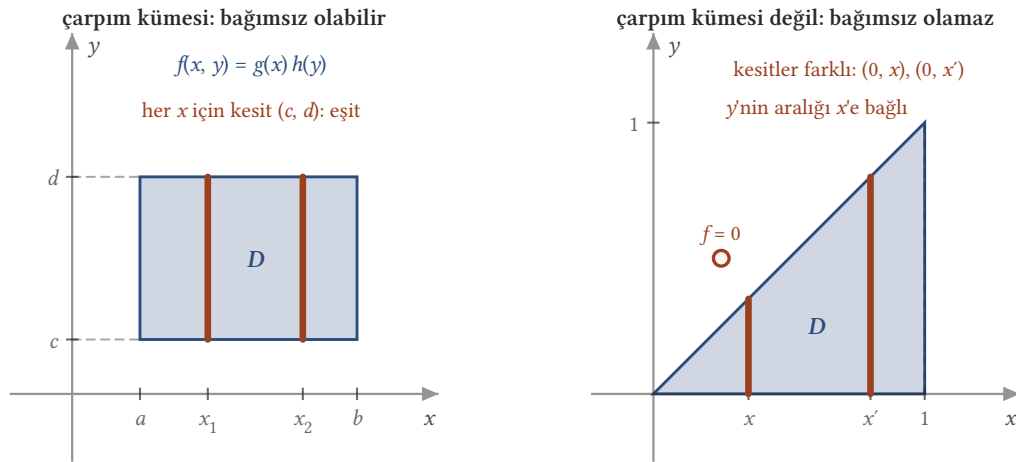
Tanımdaki eşitlik **her** noktada istenir; yalnızca ortak fonksiyonun pozitif olduğu noktalarda değil. Ortak fonksiyon bir noktada sıfır olduğu hâlde marjinallerin çarpımı orada pozitif olabilir; böyle bir nokta bağımsızlığı tek başına bozar. Aşağıdaki iki gözlem, bağımsızlığı çoğu örnekte hesap yapmadan görmeyi sağlar.

💡 Bağımsızlığı hızlı görmek: çarpanlara ayrılma ve destek

$f = f_{X,Y}$ ortak olasılık (yoğunluk) fonksiyonu olsun; $D = \{(x, y) : f(x, y) > 0\}$ kümesine ortak fonksiyonun **desteği** diyelim.

Çarpanlara ayrılma. Her $(x, y) \in \mathbb{R}^2$ için $f(x, y) = g(x)h(y)$ olacak biçimde negatif olmayan g ve h fonksiyonları varsa X ve Y bağımsızdır. Gerçekten, $c_g = \int g$ ve $c_h = \int h$ dersek $f_X(x) = g(x)c_h$, $f_Y(y) = h(y)c_g$ ve $1 = \iint f = c_g c_h$ olur; dolayısıyla $f_X(x)f_Y(y) = g(x)h(y) = f(x, y)$ (kesikli durumda integraller yerine toplamlar). g ve h 'nin marjinallerin kendisi olması gerekmez.

Destek çarpım biçiminde olmalı. X ve Y bağımsızsa $D = D_X \times D_Y$ olmak zorundadır; çünkü $f = f_X f_Y$ ise $f(x, y) > 0$ olması tam olarak $f_X(x) > 0$ ve $f_Y(y) > 0$ olması demektir. Öyleyse desteği çarpım kümesi olmayan, örneğin üçgen bir bölgede pozitif olan ortak yoğunluğa sahip değişkenler **bağımsız olamaz**: üçgenin dışında ama marjinallerin ikisinin de pozitif olduğu bir (x, y) noktasında $f(x, y) = 0 \neq f_X(x)f_Y(y)$ olur.



Şekil 14.1: Destek ölçütü. Solda destek $D = [a, b] \times [c, d]$ bir çarpım kümesidir: x ne olursa olsun dikey kesit aynı (c, d) aralığıdır; yoğunluk $g(x)h(y)$ biçiminde çarpanlara ayrılabilir, yani X ve Y bağımsız olabilir. Sağda destek üçgendir: kesit $(0, x)$, x büyüdükçe uzar; y 'nin aralığı x 'e bağlı. Köşegenin üstündeki içi boş noktada $f = 0$ iken iki marjinal de pozitifdir, dolayısıyla $f \neq f_X f_Y$; değişkenler bağımsız olamaz.

Bağımsızlık tanımı ortak fonksiyon üzerinden verildi. Aynı özellik ortak dağılım fonksiyonuyla (Tanım 13.2) da yakalanabilir; bu, kesikli ve sürekli durumu tek bir ifadede birleştirir.

Önerme 14.1 (Bağımsızlığın Dağılım Fonksiyonuyla Karakterizasyonu). (X, Y) kesikli ya da sürekli bir rastgele vektör; $F_{X,Y}$ ortak dağılım fonksiyonu, F_X ve F_Y marjinal dağılım fonksiyonları olsun. X ve Y 'nin bağımsız olması için gerek ve yeter koşul, her $(x, y) \in \mathbb{R}^2$ için

$$F_{X,Y}(x, y) = F_X(x) F_Y(y)$$

olmasıdır.

İspat.

Gerekliklik. X ve Y bağımsız olsun. Sürekli durumda ortak dağılım fonksiyonunun tanımı ve bağımsızlık gereği

$$F_{X,Y}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_X(s) f_Y(t) dt ds = \left(\int_{-\infty}^x f_X(s) ds \right) \left(\int_{-\infty}^y f_Y(t) dt \right) = F_X(x) F_Y(y)$$

olur. Kesikli durumda integraller yerine $s \leq x$, $t \leq y$ üzerinden toplamlar yazılır; çarpanlara ayırma aynıdır.

Yeterlilik. Her (x, y) için $F_{X,Y}(x, y) = F_X(x)F_Y(y)$ olsun.

Kesikli durum. $x \in D_X$, $y \in D_Y$ ve $h > 0$ olsun. $(x - h, x] \times (y - h, y]$ dikdörtgenine düşme olasılığı, ortak dağılım fonksiyonunun köşe değerlerinin işaretli toplamıdır:

$$P(x - h < X \leq x, y - h < Y \leq y) = F_{X,Y}(x, y) - F_{X,Y}(x - h, y) - F_{X,Y}(x, y - h) + F_{X,Y}(x - h, y - h).$$

Çarpım biçimini yerine yazıp düzenlersek

$$P(x - h < X \leq x, y - h < Y \leq y) = [F_X(x) - F_X(x - h)] [F_Y(y) - F_Y(y - h)] = P(x - h < X \leq x) P(y - h < Y \leq y)$$

elde edilir. $h \rightarrow 0^+$ iken dikdörtgenler azalarak $\{(x, y)\}$ noktasına iner; olasılık ölçüsünün azalan dizilerdeki sürekliliği (Teorem 3.5) gereği sol taraf $P(X = x, Y = y)$ 'ye, sağdaki çarpanlar $P(X = x)$ ve $P(Y = y)$ 'ye yakınsar. Böylece $P(X = x, Y = y) = P(X = x) P(Y = y)$ olur.

Sürekli durum. Ortak yoğunluk, ortak dağılım fonksiyonunun karma ikinci türevidir: $f_{X,Y}$ 'nin sürekli olduğu her noktada $f_{X,Y}(x, y) = \frac{\partial^2 F_{X,Y}}{\partial x \partial y}(x, y)$ olur. Çarpım biçimini türevlersek

$$\frac{\partial^2}{\partial x \partial y} [F_X(x) F_Y(y)] = F_X'(x) F_Y'(y) = f_X(x) f_Y(y)$$

bulunur (Teorem 11.2). Yoğunluğun süreksiz olduğu noktalar olasılık hesaplarını etkilemez; ortak yoğunluk bu noktalarda da marjinalerin çarpımı olarak seçilebilir. Demek ki $f_{X,Y} = f_X f_Y$, yani X ve Y bağımsızdır.

■

Bağımsız X ve Y alalım. Her $A, B \in \mathcal{B}(\mathbb{R})$ için $P(X \in A, Y \in B) = P(X \in A) P(Y \in B)$ olur; bu, çarpım biçimindeki fonksiyonun $A \times B$ üzerindeki integralinin (toplamının) çarpanlara ayrılmasıdır. Bağımsızlığın ilk uygulaması, tek tek bilinen dağılımlardan ortak dağılımı kurmaktır: bağımsız değişkenlerin ortak yoğunluğu marjinalerin çarpımıdır.

Örnek 14.1 (Bağımsız Parçaların Ömürleri). Belli bir parçanın tükeninceye kadar geçen ömrü (yıl olarak),

$$f(x) = \begin{cases} \frac{1}{3} e^{-x/3}, & x > 0 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olasılık yoğunluk fonksiyonuna sahip olsun. Birbirinden bağımsız çalışan böyle iki parçanın ömürleri sırasıyla X ve Y rastgele değişkenleriyle gösterilsin. $P(X \geq 1, Y \geq 1)$ ve $P(X + Y \geq 2)$ olasılıklarını hesaplayınız.

Çözüm.

Ortak yoğunluk. X ve Y bağımsız ve her ikisi de f yoğunluğuna sahip olduğundan (Tanım 14.1)

$$f_{X,Y}(x, y) = f(x) f(y) = \begin{cases} \frac{1}{9} e^{-(x+y)/3}, & x > 0, y > 0 \\ 0, & \text{diğer durumlarda.} \end{cases}$$

Birinci olasılık. $(X \geq 1)$ ve $(Y \geq 1)$ olayları bağımsız olduğundan

$$P(X \geq 1, Y \geq 1) = P(X \geq 1) P(Y \geq 1) = \left(\int_1^\infty \frac{1}{3} e^{-x/3} dx \right)^2 = (e^{-1/3})^2 = e^{-2/3} \approx 0,5134.$$

İkinci olasılık. Tümleneyen olayı hesaplamak daha kolaydır. $(X + Y < 2)$ olayı $x > 0, y > 0, x + y < 2$ üçgen bölgesine karşılık gelir:

$$P(X + Y < 2) = \int_0^2 \int_0^{2-x} \frac{1}{9} e^{-(x+y)/3} dy dx.$$

İç integral

$$\int_0^{2-x} \frac{1}{9} e^{-x/3} e^{-y/3} dy = \frac{1}{9} e^{-x/3} \cdot 3(1 - e^{-(2-x)/3}) = \frac{1}{3} (e^{-x/3} - e^{-2/3})$$

olur. Dış integral

$$P(X + Y < 2) = \int_0^2 \frac{1}{3} e^{-x/3} dx - \frac{2}{3} e^{-2/3} = (1 - e^{-2/3}) - \frac{2}{3} e^{-2/3} = 1 - \frac{5}{3} e^{-2/3}.$$

Böylece

$$P(X + Y \geq 2) = 1 - P(X + Y < 2) = \frac{5}{3} e^{-2/3} \approx 0,8557.$$

■

14.2 Koşullu Dağılımlar

Koşullu olasılık (Tanım 6.1), bir B olayının gerçekleştiği bilindiğinde olasılık ölçüsünün nasıl güncellendiğini söyler: $P(A | B) = P(A \cap B)/P(B)$. Aynı fikri bir rastgele vektörün dağılımına uygulayacağız: (X_1, X_2) vektörünün bir A bölgesine düştüğü biliniyorsa vektörün dağılımı ne olur?

Tanım 14.2 (Rastgele Vektörün Koşullu Dağılımı). $(\Omega, \mathcal{U}, P)$ olasılık uzayında tanımlı (X_1, X_2) iki boyutlu rastgele vektörünün olasılık dağılımı $P_{(X_1, X_2)}$ olsun (Teorem 13.1). $A \in \mathcal{B}(\mathbb{R}^2)$ ve $P_{(X_1, X_2)}(A) > 0$ olsun. $B \in \mathcal{B}(\mathbb{R}^2)$ için

$$P_{(X_1, X_2)}(B | A) = \frac{P_{(X_1, X_2)}(A \cap B)}{P_{(X_1, X_2)}(A)} = \frac{P((X_1, X_2) \in A \cap B)}{P((X_1, X_2) \in A)}$$

ile tanımlanan $B \mapsto P_{(X_1, X_2)}(B | A)$ olasılık ölçüsüne, $(X_1, X_2) \in A$ verilmişken (X_1, X_2) rastgele vektörünün **koşullu dağılımı** denir.

$P_{(X_1, X_2)}$ bir olasılık ölçüsü olduğundan, A 'ya göre koşullu olasılığı da $\mathcal{B}(\mathbb{R}^2)$ üzerinde bir olasılık ölçüsüdür. En çok kullanılan özel durum $A = \mathbb{R} \times \{x_2\}$, yani " $X_2 = x_2$ verilmişken" koşuludur. Bu koşulun olasılığı $P(X_2 = x_2)$ kesikli durumda pozitif olabilir; sürekli durumda ise her zaman sıfırdır. İki durumu ayrı ayrı ele alıyoruz.

14.2.1 Kesikli Durum: Koşullu Olasılık Fonksiyonu

Kesikli bir rastgele vektörde $X_2 = x_2$ olayının olasılığı $f_{X_2}(x_2)$ marjinal değeridir ve $x_2 \in D_{X_2}$ için pozitiftir. Öyleyse koşullu dağılım tanımı doğrudan uygulanır ve sonuç tek değişkenli bir olasılık fonksiyonudur.

Teorem 14.1 (Koşullu Olasılık Fonksiyonu). (X_1, X_2) kesikli bir rastgele vektör ve ortak olasılık fonksiyonu f_{X_1, X_2} olsun. $P(X_2 = x_2) \neq 0$ olan bir x_2 için, $X_2 = x_2$ verilmişken X_1 rastgele değişkeninin koşullu dağılımının olasılık fonksiyonu (kısaca **koşullu olasılık fonksiyonu**)

$$f_{X_1|X_2=x_2}(x_1) = P(X_1 = x_1 | X_2 = x_2) = \frac{P(X_1 = x_1, X_2 = x_2)}{P(X_2 = x_2)} = \frac{f_{X_1, X_2}(x_1, x_2)}{f_{X_2}(x_2)}, \quad x_1 \in D_{X_1|X_2=x_2}$$

dir; burada $D_{X_1|X_2=x_2} = \{x_1 \in D_{X_1} : f_{X_1, X_2}(x_1, x_2) > 0\}$ koşullu dağılımın değer kümesidir. Benzer biçimde, $P(X_1 = x_1) \neq 0$ olan bir x_1 için $X_1 = x_1$ verilmişken X_2 'nin koşullu olasılık fonksiyonu

$$f_{X_2|X_1=x_1}(x_2) = P(X_2 = x_2 | X_1 = x_1) = \frac{f_{X_1, X_2}(x_1, x_2)}{f_{X_1}(x_1)}, \quad x_2 \in D_{X_2|X_1=x_1}$$

dir.

İspat.

$A = \mathbb{R} \times \{x_2\}$ alalım; bu küme kapalı olduğundan $\mathcal{B}(\mathbb{R}^2)$ 'ye aittir (Tanım 2.5) ve

$$P_{(X_1, X_2)}(A) = P((X_1, X_2) \in \mathbb{R} \times \{x_2\}) = P(X_2 = x_2) = f_{X_2}(x_2) > 0$$

olduğundan Tanım 14.2 uygulanabilir. X_1 'in x_1 değerini alması olayı $B = \{x_1\} \times \mathbb{R}$ kümesine karşılık gelir ve $A \cap B = \{(x_1, x_2)\}$ tek noktadır. Öyleyse

$$P_{(X_1, X_2)}(B | A) = \frac{P(X_1 = x_1, X_2 = x_2)}{P(X_2 = x_2)} = \frac{f_{X_1, X_2}(x_1, x_2)}{f_{X_2}(x_2)}.$$

Sol taraf, olaylar için koşullu olasılık tanımı gereği tam olarak $P(X_1 = x_1 | X_2 = x_2)$ 'dir. Bu değer $x_1 \notin D_{X_1|X_2=x_2}$ için sıfırdır; pozitif olduğu noktalar tam olarak $D_{X_1|X_2=x_2}$ kümesidir. Fonksiyonun bir olasılık fonksiyonu olduğu aşağıda Önerme 14.2'nde gösterilecektir; dolayısıyla koşullu dağılım, olasılık fonksiyonu bu olan kesikli bir dağılımdır. İkinci formül, rollerin değiştirilmesiyle aynı biçimde elde edilir. ■

Pratikte ortak dağılım tablosunda x_2 sütunu alınır ve her sayı sütun toplamına bölünür. X_1 ile ilgili koşullu olasılıklar bu fonksiyondan toplamla hesaplanır:

$$P(X_1 \in B | X_2 = x_2) = \sum_{x_1 \in B \cap D_{X_1|X_2=x_2}} f_{X_1|X_2=x_2}(x_1).$$

14.2.2 Sürekli Durum: Koşullu Yoğunluk Fonksiyonu

Sürekli bir rastgele vektörde $P(X_2 = x_2) = 0$ olduğundan Tanım 14.2 doğrudan uygulanamaz. Yine de “ $X_2 = x_2$ verilmişken X_1 'in dağılımı” sorusunun anlamlı bir yanıtı vardır; ona bir limitle ulaşılır. $f_{X_2}(x_2) > 0$ ve h küçükken $(x_2 < X_2 \leq x_2 + h)$ olayının olasılığı pozitifdir ve

$$P(X_1 \leq x_1 | x_2 < X_2 \leq x_2 + h) = \frac{F_{X_1, X_2}(x_1, x_2 + h) - F_{X_1, X_2}(x_1, x_2)}{F_{X_2}(x_2 + h) - F_{X_2}(x_2)}$$

olur. Pay ve payda h 'ye bölünüp $h \rightarrow 0^+$ alınırsa pay $\frac{\partial F_{X_1, X_2}}{\partial x_2}(x_1, x_2) = \int_{-\infty}^{x_1} f_{X_1, X_2}(t, x_2) dt$ 'ye, payda $f_{X_2}(x_2)$ 'ye gider. Limit

$$\int_{-\infty}^{x_1} \frac{f_{X_1, X_2}(t, x_2)}{f_{X_2}(x_2)} dt$$

olur; bu, x_1 'in fonksiyonu olarak bir dağılım fonksiyonudur ve yoğunluğu integralin içindeki ifadedir. Bu gözlem aşağıdaki tanımı gerektirir.

Tanım 14.3 (Koşullu Olasılık Yoğunluk Fonksiyonu). (X_1, X_2) sürekli bir rastgele vektör, ortak olasılık yoğunluk fonksiyonu f_{X_1, X_2} , bu fonksiyonun pozitif olduğu bölge D ve marjinal yoğunluklar f_{X_1}, f_{X_2} olsun. $f_{X_2}(x_2) \neq 0$ olan bir x_2 için

$$f_{X_1|X_2=x_2}(x_1) = \begin{cases} \frac{f_{X_1, X_2}(x_1, x_2)}{f_{X_2}(x_2)}, & x_1 \in D_{X_1|X_2=x_2} \\ 0, & \text{diğer durumlarda} \end{cases}$$

fonksiyonuna $X_2 = x_2$ verilmişken X_1 rastgele değişkeninin **koşullu olasılık yoğunluk fonksiyonu** (koşullu marjinal olasılık yoğunluk fonksiyonu) denir. Burada

$$D_{X_1|X_2=x_2} = \{x_1 \in \mathbb{R} : (x_1, x_2) \in D\}$$

kümesi, D bölgesinin x_2 yüksekliğindeki yatay kesitidir. Benzer biçimde, $f_{X_1}(x_1) \neq 0$ olan bir x_1 için

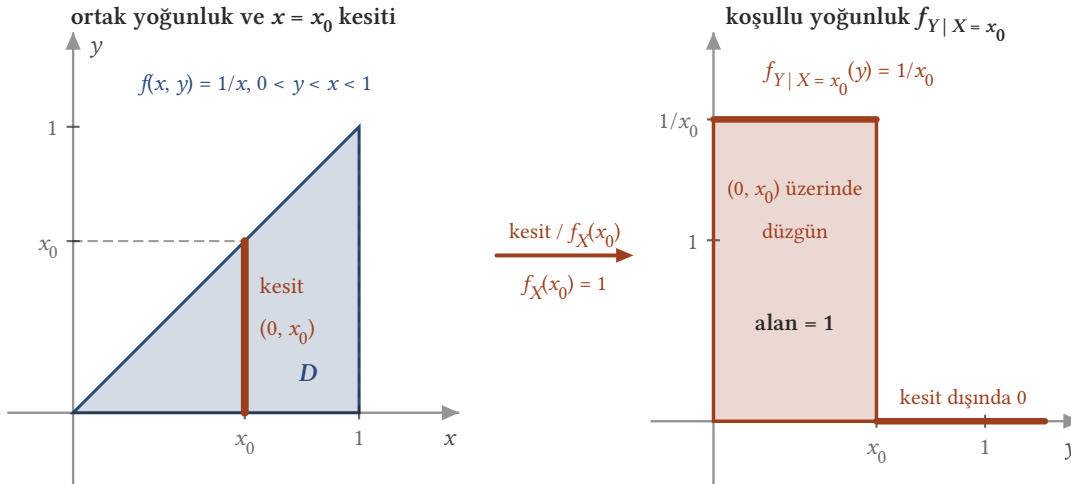
$$f_{X_2|X_1=x_1}(x_2) = \begin{cases} \frac{f_{X_1, X_2}(x_1, x_2)}{f_{X_1}(x_1)}, & x_2 \in D_{X_2|X_1=x_1} \\ 0, & \text{diğer durumlarda} \end{cases}$$

fonksiyonuna $X_1 = x_1$ verilmişken X_2 'nin **koşullu olasılık yoğunluk fonksiyonu** denir; $D_{X_2|X_1=x_1} = \{x_2 \in \mathbb{R} : (x_1, x_2) \in D\}$, D 'nin x_1 apsisindeki düşey kesitidir.

Sürekli durumda X_1 ile ilgili koşullu olasılıklar, koşullu yoğunluğun integraliyle **tanımlanır**:

$$P(X_1 \in B | X_2 = x_2) = \int_B f_{X_1|X_2=x_2}(x_1) dx_1.$$

En sık yapılan yanlış, integral sınırlarını koşullu değer kümesine göre kısıtlamamaktır. $X_2 = x_2$ verildiğinde X_1 yalnızca D 'nin o kesitindeki değerleri alabilir; kesitin dışında koşullu yoğunluk sıfırdır.



Şekil 14.2: Koşullu yoğunluk, ortak yoğunluğun bir kesitidir. Solda $f(x, y) = 1/x$ yoğunluğunun üçgen desteği ve $x = x_0$ apsisindeki kalın düşey kesit $(0, x_0)$ görülür. Bu kesit üzerindeki değerler $f_X(x_0) = 1$ marjinaline bölünmüştür sağdaki koşullu yoğunluk çıkar: $(0, x_0)$ üzerinde sabit $1/x_0$, kesit dışında 0. Dikdörtgenin alanı 1'dir; $X = x_0$ verilmişken Y , $(0, x_0)$ üzerinde düzgün dağılır.

Kesikli ve sürekli tanımlar aynı kalıba uyar: koşullu fonksiyon, ortak fonksiyonun sabit tutulan değişkene ait marjinaline bölünmesidir. Doğal soru, bölme sonucunda gerçekten bir olasılık (yoğunluk) fonksiyonu elde edilip edilmediğidir.

Önerme 14.2 (Koşullu Fonksiyonlar Olasılık (Yoğunluk) Fonksiyonudur). (a) (X_1, X_2) kesikli bir rastgele vektör ve $f_{X_2}(x_2) > 0$ olsun. $f_{X_1|X_2=x_2}$, değer kümesi $D_{X_1|X_2=x_2}$ olan bir olasılık fonksiyonudur: negatif değildir ve

$$\sum_{x_1 \in D_{X_1|X_2=x_2}} f_{X_1|X_2=x_2}(x_1) = 1.$$

(b) (X_1, X_2) sürekli bir rastgele vektör ve $f_{X_2}(x_2) > 0$ olsun. $f_{X_1|X_2=x_2}$ bir olasılık yoğunluk fonksiyonudur: negatif değildir ve

$$\int_{-\infty}^{\infty} f_{X_1|X_2=x_2}(x_1) dx_1 = 1.$$

Aynı sonuçlar $f_{X_2|X_1=x_1}$ için de geçerlidir.

İspat.

(a) Pay $f_{X_1, X_2}(x_1, x_2) \geq 0$ ve payda $f_{X_2}(x_2) > 0$ olduğundan bölüm negatif değildir. Marjinal olasılık fonksiyonunun tanımı (Tanım 13.5) gereği $f_{X_2}(x_2) = \sum_{x_1 \in D_{X_1}} f_{X_1, X_2}(x_1, x_2)$ olur ve toplama yalnızca $f_{X_1, X_2}(x_1, x_2) > 0$ olan terimler, yani $D_{X_1|X_2=x_2}$ 'dekiler katkı verir. Öyleyse

$$\sum_{x_1 \in D_{X_1|X_2=x_2}} f_{X_1|X_2=x_2}(x_1) = \frac{1}{f_{X_2}(x_2)} \sum_{x_1 \in D_{X_1|X_2=x_2}} f_{X_1, X_2}(x_1, x_2) = \frac{f_{X_2}(x_2)}{f_{X_2}(x_2)} = 1.$$

(b) Negatif olmama aynı nedenle açıktır. Marjinal yoğunluğun tanımı (Tanım 13.8) gereği $f_{X_2}(x_2) = \int_{-\infty}^{\infty} f_{X_1, X_2}(x_1, x_2) dx_1$ olur; integrale yalnızca $D_{X_1|X_2=x_2}$ kesiti katkı verir. Öyleyse

$$\int_{-\infty}^{\infty} f_{X_1|X_2=x_2}(x_1) dx_1 = \frac{1}{f_{X_2}(x_2)} \int_{D_{X_1|X_2=x_2}} f_{X_1, X_2}(x_1, x_2) dx_1 = \frac{f_{X_2}(x_2)}{f_{X_2}(x_2)} = 1.$$

$f_{X_2|X_1=x_1}$ için ispat, rollerin değiştirilmesiyle aynı biçimde yapılır.

■

i Çarpım kuralı: koşullu ile marjinalden ortağa

Koşullu fonksiyonun tanımı ters çevrilirse ortak fonksiyon geri elde edilir:

$$f_{X_1, X_2}(x_1, x_2) = f_{X_1}(x_1) f_{X_2|X_1=x_1}(x_2) = f_{X_2}(x_2) f_{X_1|X_2=x_2}(x_1).$$

Bu, olaylar için çarpım kuralının ($P(A \cap B) = P(B) P(A | B)$) rastgele değişkenlerdeki karşılığıdır. İki aşamalı deneylerde ortak dağılım çoğu zaman bu yoldan kurulur: önce X_1 'in dağılımı, sonra $X_1 = x_1$ verilmişken X_2 'nin dağılımı yazılır ve ikisi çarpılır. Girişte sözü edilen “marjinallerden ortağa dönüş” sorununun eksik olan bilgi, koşullu dağılımdır.

14.3 Bağımsızlık ile Koşullu Dağılım Arasındaki İlişki

Bağımsızlığın sezgisel anlamı, “ X ’in değerini bilmek Y hakkında bilgi vermez” idi. Koşullu dağılım ise tam olarak “ X ’in değeri bilindiğinde Y ’nin dağılımı”dır. Öyleyse bağımsızlık, koşullu dağılımın koşuldan etkilenmemesi, yani marjinala eşit kalması demek olmalıdır.

Teorem 14.2 (Bağımsızlık ve Koşullu Dağılımın Marjinal Eşitliği). (X, Y) kesikli ya da sürekli bir rastgele vektör olsun. Aşağıdaki üç ifade birbirine denktir:

- (i) X ve Y bağımsızdır.
 - (ii) $f_X(x) > 0$ olan her x ve her y için $f_{Y|X=x}(y) = f_Y(y)$ olur; yani $X = x$ verilmişken Y ’nin koşullu dağılımı Y ’nin marjinal dağılımına eşittir.
 - (iii) $X = x$ verilmişken Y ’nin koşullu dağılımı x ’e bağlı değildir; yani $f_X(x) > 0$ olan her x için $f_{Y|X=x}(y) = g(y)$ olacak biçimde yalnız y ’ye bağlı bir g fonksiyonu vardır.
- Aynı denklikler X ile Y ’nin rolleri değiştirilerek de geçerlidir: bağımsızlık, $f_{X|Y=y} = f_X$ olmasına denktir.

İspat.

(i) $\Rightarrow$ (ii). X ve Y bağımsız olsun; $f_X(x) > 0$ olan bir x alalım. Koşullu fonksiyonun tanımı ve bağımsızlık gereği

$$f_{Y|X=x}(y) = \frac{f_{X, Y}(x, y)}{f_X(x)} = \frac{f_X(x) f_Y(y)}{f_X(x)} = f_Y(y).$$

(ii) $\Rightarrow$ (iii). $g = f_Y$ alınrsa (iii) sağlanır.

(iii) $\Rightarrow$ (i). $f_X(x) > 0$ olan her x için $f_{Y|X=x}(y) = g(y)$ olsun. Koşullu fonksiyonun tanımından, bu x ’ler için $f_{X, Y}(x, y) = f_X(x) g(y)$ olur. $f_X(x) = 0$ olan x ’lerde ise $f_{X, Y}(x, y) = 0$ alınabilir: kesikli durumda $x \notin D_X$ olduğundan bu kesindir; sürekli durumda $\int_{-\infty}^{\infty} f_{X, Y}(x, y) dy = 0$ ve $f_{X, Y} \geq 0$ ’dır, yoğunluğun sıfır olasılıklı bir kümede değiştirilmesi de olasılık hesaplarını etkilemez. Şimdi g ’yi bulalım; kesikli durumda

$$f_Y(y) = \sum_{x \in D_X} f_{X, Y}(x, y) = g(y) \sum_{x \in D_X} f_X(x) = g(y),$$

sürekli durumda aynı hesap integralle yapılır: $f_Y(y) = \int_{-\infty}^{\infty} f_X(x) g(y) dx = g(y)$. Demek ki $g = f_Y$ ve her (x, y) için $f_{X,Y}(x, y) = f_X(x) f_Y(y)$: $f_X(x) > 0$ ise bu, $g = f_Y$ ile elde edilen eşitliktir; $f_X(x) = 0$ ise iki taraf da sıfırdır. **Tanım 14.1** gereği X ve Y bağımsızdır.

Pratik sonuç: bir örnekte koşullu fonksiyon koşula bağlı çıkıyorsa değişkenler bağımlıdır; bağımsız olduğu bilinen değişkenlerde ise koşullu olasılıklar doğrudan marjinalle hesaplanır.

14.4 Kesikli Örnekler

Tablodan koşullu olasılık fonksiyonu elde etmek için ilgili satır ya da sütun kendi toplamına bölünür; bağımsızlık, her hücrenin satır ve sütun toplamının çarpımına eşit olup olmadığına bakılarak sınanır.

Örnek 14.2 (Üç Para Atışında Koşullu Dağılım). Düzgün bir para üç kez atılıyor. X_1 üç atışta gelen tura sayısını, X_2 ilk iki atışta gelen tura sayısını gösterebilir. (X_1, X_2) rastgele vektörünün ortak olasılık fonksiyonu tablosu **Örnek 13.1**'ta örnek uzayın sekiz eş olasılıklı sonucu sayılarak kurulmuştur:

$x_1 \setminus x_2$	0	1	2	$f_{X_1}(x_1)$
0	$\frac{1}{8}$	0	0	$\frac{1}{8}$
1	$\frac{1}{8}$	$\frac{2}{8}$	0	$\frac{3}{8}$
2	0	$\frac{2}{8}$	$\frac{1}{8}$	$\frac{3}{8}$
3	0	0	$\frac{1}{8}$	$\frac{1}{8}$
$f_{X_2}(x_2)$	$\frac{2}{8}$	$\frac{4}{8}$	$\frac{2}{8}$	1

- $P(X_1 = 2 \mid X_2 = 1)$ koşullu olasılığını hesaplayınız.
- $P(1 < X_1 \leq 2 \mid X_2 = 1)$ koşullu olasılığını hesaplayınız.
- $X_2 = 1$ verilmişken X_1 'in koşullu olasılık fonksiyonunu bulunuz.
- $X_1 = 1$ verilmişken X_2 'nin koşullu olasılık fonksiyonunu bulunuz.
- X_1 ve X_2 bağımsız mıdır?

Çözüm.

(a) **Teorem 14.1** gereği

$$P(X_1 = 2 \mid X_2 = 1) = \frac{f_{X_1, X_2}(2, 1)}{f_{X_2}(1)} = \frac{2/8}{4/8} = \frac{1}{2}.$$

Sezgisel açıklama: ilk iki atışta bir tura varsa toplamın 2 olması üçüncü atışın tura gelmesi demektir.

(b) X_1 tam sayı değerli olduğundan $(1 < X_1 \leq 2) = (X_1 = 2)$ olur:

$$P(1 < X_1 \leq 2 \mid X_2 = 1) = P(X_1 = 2 \mid X_2 = 1) = \frac{1}{2}.$$

(c) $x_2 = 1$ sütununda pozitif hücreler $x_1 = 1$ ve $x_1 = 2$ satırlarındadır; koşullu değer kümesi $D_{X_1|X_2=1} = \{1, 2\}$ 'dir. Sütundaki sayıları sütun toplamı $f_{X_2}(1) = \frac{4}{8}$ 'e bölersek

$$f_{X_1|X_2=1}(1) = \frac{2/8}{4/8} = \frac{1}{2}, \quad f_{X_1|X_2=1}(2) = \frac{2/8}{4/8} = \frac{1}{2}$$

bulunur. Toplam 1'dir (**Önerme 14.2**).

(d) $x_1 = 1$ satırında pozitif hücreler $x_2 = 0$ ve $x_2 = 1$ sütunlarındadır; satır toplamı $f_{X_1}(1) = \frac{3}{8}$ 'dir:

$$f_{X_2|X_1=1}(0) = \frac{1/8}{3/8} = \frac{1}{3}, \quad f_{X_2|X_1=1}(1) = \frac{2/8}{3/8} = \frac{2}{3}.$$

Üç atıştaki tek tura, $\frac{1}{8}$ olasılıkla üçüncü atıştır.

(e) Bağımsızlık için her hücrenin, satır ve sütun toplamalarının çarpımına eşit olması gerekir. Tek bir hücre yeter:

$$f_{X_1, X_2}(0, 0) = \frac{1}{8}, \quad f_{X_1}(0) f_{X_2}(0) = \frac{1}{8} \cdot \frac{2}{8} = \frac{1}{32}.$$

Eşit olmadıklarından X_1 ve X_2 **bağımsız değildir**. Destek de bunu gösterir: $f_{X_1, X_2}(0, 2) = 0$ olduğu hâlde $f_{X_1}(0) > 0$ ve $f_{X_2}(2) > 0$ 'dır. **Teorem 14.2** ile de görülür: (c)'deki koşullu fonksiyon, X_1 'in marjinalinden $(\frac{1}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{8})$ farklıdır. ■

Örnek 14.3 (Üç Toptan İki: Toplam ve Fark). Bir kutudaki üç top 1, 2, 3 rakamlarıyla numaralanmıştır. Kutudan iadesiz olarak iki top çekiliyor. X çekilen toplardaki sayıların toplamını, Y ise bu sayıların farkının mutlak değerini gösterebilir. X ve Y 'nin ortak olasılık fonksiyonunu oluşturunuz; koşullu olasılık fonksiyonlarını bulunuz ve X ile Y 'nin bağımsız olup olmadığını belirleyiniz.

Çözüm.

Örnek uzay. Çekiliş sırası gözlemlenirse örnek uzay

$$\Omega = \{(1, 2), (2, 1), (1, 3), (3, 1), (2, 3), (3, 2)\}$$

altı eş olasılıklı sonuçtan oluşur. $(1, 2)$ ve $(2, 1)$ için $X = 3, Y = 1$; $(1, 3)$ ve $(3, 1)$ için $X = 4, Y = 2$; $(2, 3)$ ve $(3, 2)$ için $X = 5, Y = 1$ 'dir. Değer kümeleri $D_X = \{3, 4, 5\}$ ve $D_Y = \{1, 2\}$ 'dir.

Ortak olasılık fonksiyonu. Her (x, y) çifti için o çifti veren sonuçlar sayılır:

$x \setminus y$	1	2	$f_X(x)$
3	$\frac{2}{6}$	0	$\frac{2}{6}$
4	0	$\frac{2}{6}$	$\frac{2}{6}$
5	$\frac{2}{6}$	0	$\frac{2}{6}$
$f_Y(y)$	$\frac{4}{6}$	$\frac{2}{6}$	1

Ortak fonksiyonun desteği $\{(3, 1), (4, 2), (5, 1)\}$ kümesidir.

Koşullu olasılık fonksiyonları. $Y = 1$ sütununu toplamı $\frac{4}{6}$ 'ya bölersek

$$f_{X|Y=1}(3) = \frac{2/6}{4/6} = \frac{1}{2}, \quad f_{X|Y=1}(5) = \frac{1}{2};$$

$Y = 2$ sütununda tek pozitif hücre vardır: $f_{X|Y=2}(4) = 1$. Satırlara bakarsak

$$f_{Y|X=3}(1) = 1, \quad f_{Y|X=4}(2) = 1, \quad f_{Y|X=5}(1) = 1.$$

X bilindiğinde Y kesin olarak belirlenir; gerçekten Y, X 'in bir fonksiyonudur: $Y = 2$ ancak ve ancak $X = 4$. Bu, bağımsızlığın tam karşısı olan durumdur.

Bağımsızlık. $f_{X,Y}(3, 2) = 0$ olduğu hâlde $f_X(3) f_Y(2) = \frac{2}{6} \cdot \frac{2}{6} = \frac{1}{9} \neq 0$ 'dır. Öyleyse X ve Y **bağımsız değildir**; koşullu fonksiyonların koşula bağlı olması da aynı sonucu verir. ■

Örnek 14.4 (İki Zar: İlk Sayı ve Farkın Mutlak Değeri). Düzgün bir zar iki kez atılıyor. X ilk atışta gelen sayıyı, Y ise ikinci atışta gelen sayı ile ilk atışta gelen sayı arasındaki farkın mutlak değerini gösterebilir. X ve Y 'nin ortak olasılık fonksiyonunu oluşturunuz; marjinaleri, $Y = 5$ verilmişken X 'in ve $X = 1$ verilmişken Y 'nin koşullu olasılık fonksiyonlarını bulunuz; X ile Y 'nin bağımsız olup olmadığını belirleyiniz.

Çözüm.

Kurulum. Örnek uzay, i ilk ve j ikinci atışın sonucu olmak üzere 36 eş olasılıklı (i, j) çiftinden oluşur; $X = i$ ve $Y = |j - i|$ 'dir. $D_X = \{1, \dots, 6\}$, $D_Y = \{0, 1, \dots, 5\}$. Sabit bir x için $|j - x| = y$ denkleminin $j \in \{1, \dots, 6\}$ içindeki çözüm sayısı hücreyi belirler: $y = 0$ için tek çözüm ($j = x$) vardır; $y \geq 1$ için adaylar $j = x - y$ ve $j = x + y$ 'dir ve bunlardan 1 ile 6 arasında kalanlar sayılır.

Ortak olasılık fonksiyonu tablosu.

$x \setminus y$	0	1	2	3	4	5	$f_X(x)$
1	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{6}{36}$
2	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	0	$\frac{6}{36}$
3	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	0	0	$\frac{6}{36}$
4	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	0	0	$\frac{6}{36}$
5	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	0	$\frac{6}{36}$
6	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{6}{36}$
$f_Y(y)$	$\frac{6}{36}$	$\frac{10}{36}$	$\frac{8}{36}$	$\frac{6}{36}$	$\frac{4}{36}$	$\frac{2}{36}$	1

Örneğin $x = 3$, $y = 2$ için adaylar $j = 1$ ve $j = 5$ 'tir, ikisi de geçerlidir: hücre $\frac{2}{36}$. $x = 2$, $y = 5$ için adaylar $j = -3$ ve $j = 7$ 'dir, hiçbiri geçerli değildir: hücre 0. Her satırın toplamı $\frac{1}{6}$ 'dır (X düzgün zarın kendisidir); sütun toplamaları Y 'nin marjinalini verir.

Koşullu fonksiyonlar. $Y = 5$ sütununda pozitif hücreler $x = 1$ ve $x = 6$ 'dadır; sütun toplamı $\frac{2}{36}$ olduğundan

$$f_{X|Y=5}(1) = \frac{1/36}{2/36} = \frac{1}{2}, \quad f_{X|Y=5}(6) = \frac{1}{2}.$$

$X = 1$ satırının bütün hücreleri $\frac{1}{36}$ ve satır toplamı $\frac{6}{36}$ olduğundan

$$f_{Y|X=1}(y) = \frac{1/36}{6/36} = \frac{1}{6}, \quad y = 0, 1, \dots, 5.$$

Buna karşılık $X = 3$ satırı için $f_{Y|X=3}(1) = \frac{2/36}{6/36} = \frac{1}{3}$ ve $f_{Y|X=3}(4) = 0$ olur: koşullu dağılım x 'e bağlıdır.

Bağımsızlık. $f_{X,Y}(1, 5) = \frac{1}{36}$ iken $f_X(1) f_Y(5) = \frac{1}{6} \cdot \frac{2}{36} = \frac{1}{108}$ 'dir. Eşit olmadıklarından X ve Y **bağımsız değildir**. Tablodaki sıfır hücreler de bunu gösterir: destek, $D_X \times D_Y$ çarpım kümesinin tamamı değildir.

■

Örnek 14.5 (Çarpım Biçimindeki Ortak Olasılık Fonksiyonu). X ve Y rastgele değişkenlerinin ortak olasılık fonksiyonu

$$f(x, y) = P(X = x, Y = y) = \left(\frac{2}{3}\right)^{x+y} \left(\frac{1}{3}\right)^{2-x-y}, \quad x \in \{0, 1\}, y \in \{0, 1\}$$

olsun.

(a) f 'nin bir ortak olasılık fonksiyonu olma koşullarını sağladığını gösteriniz.

(b) X ve Y 'nin marjinal olasılık fonksiyonlarını bulunuz.

(c) X ve Y bağımsız mıdır?

(d) $P(X = x | Y = y)$ koşullu olasılık fonksiyonunu oluşturunuz ve $P(1 \leq X \leq 3 | Y = 1)$ olasılığını hesaplayınız.

Çözüm.

(a) Dört değer şunlardır:

$$f(0,0) = \left(\frac{1}{3}\right)^2 = \frac{1}{9}, \quad f(1,0) = f(0,1) = \frac{2}{3} \cdot \frac{1}{3} = \frac{2}{9}, \quad f(1,1) = \left(\frac{2}{3}\right)^2 = \frac{4}{9}.$$

Hepsi negatif değildir ve toplamları $\frac{1}{9} + \frac{2}{9} + \frac{2}{9} + \frac{4}{9} = 1$ 'dir. Öyleyse f bir ortak olasılık fonksiyonudur (Tanım 13.4).

(b) X 'in marjinali için y üzerinden toplanır:

$$f_X(x) = \sum_{y=0}^1 \left(\frac{2}{3}\right)^{x+y} \left(\frac{1}{3}\right)^{2-x-y} = \left(\frac{2}{3}\right)^x \left(\frac{1}{3}\right)^{2-x} + \left(\frac{2}{3}\right)^{x+1} \left(\frac{1}{3}\right)^{1-x}.$$

Ortak çarpan $\left(\frac{2}{3}\right)^x \left(\frac{1}{3}\right)^{1-x}$ dışarı alınır parantez içinde $\frac{1}{3} + \frac{2}{3} = 1$ kalır:

$$f_X(x) = \left(\frac{2}{3}\right)^x \left(\frac{1}{3}\right)^{1-x}, \quad x \in \{0, 1\};$$

yani $f_X(0) = \frac{1}{3}$, $f_X(1) = \frac{2}{3}$. f , x ile y 'ye göre simetrik olduğundan aynı hesap Y için de geçerlidir:

$$f_Y(y) = \left(\frac{2}{3}\right)^y \left(\frac{1}{3}\right)^{1-y}, \quad y \in \{0, 1\}.$$

(c) Üsleri ayırarak yazalım:

$$f(x, y) = \left[\left(\frac{2}{3}\right)^x \left(\frac{1}{3}\right)^{1-x} \right] \left[\left(\frac{2}{3}\right)^y \left(\frac{1}{3}\right)^{1-y} \right] = f_X(x) f_Y(y).$$

Bu eşitlik $D_X \times D_Y$ 'nin dört noktasının **hepsinde** sağlanır; tek bir noktada doğrulamak yetmez, tanım her noktayı ister. Öyleyse X ve Y **bağımsızdır**: tura olasılığı $\frac{2}{3}$ olan bir paranın iki bağımsız atışındaki "tura" göstergeleridir.

(d) X ve Y bağımsız olduğundan Teorem 14.2 gereği koşullu fonksiyon marjinaline eşittir:

$$P(X = x | Y = y) = f_{X|Y=y}(x) = f_X(x) = \left(\frac{2}{3}\right)^x \left(\frac{1}{3}\right)^{1-x}, \quad x \in \{0, 1\},$$

$y \in \{0, 1\}$ hangisi olursa olsun. Doğrudan tanımla da görülür: $f_{X|Y=1}(1) = \frac{f(1,1)}{f_Y(1)} = \frac{4/9}{2/3} = \frac{2}{3} = f_X(1)$. X yalnızca 0 ve 1 değerlerini aldığından $(1 \leq X \leq 3) = (X = 1)$ 'dir:

$$P(1 \leq X \leq 3 | Y = 1) = P(X = 1 | Y = 1) = f_X(1) = \frac{2}{3}.$$

■

Önceki örnekte destek bir çarpım kümesiydi ve ortak fonksiyon çarpanlara ayrıldı. Şimdi desteği yine çarpım kümesi olan ama çarpanlara ayrılmayan bir ortak fonksiyon görelim: destek koşulu bağımsızlık için gereklidir, yeterli değildir.

Örnek 14.6 (Toplam Biçimindeki Ortak Olasılık Fonksiyonu). X ve Y rastgele değişkenlerinin ortak olasılık fonksiyonu

$$f(x, y) = \begin{cases} \frac{x+y}{21}, & x \in \{1, 2, 3\}, y \in \{1, 2\} \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun.

(a) X ve Y 'nin marjinal olasılık fonksiyonlarını bulunuz.

(b) X ve Y bağımsız mıdır?

(c) $P(X = x | Y = y)$ koşullu olasılık fonksiyonunu oluşturunuz; $P(1 \leq X \leq 3 | Y = 1)$ ve $P(X \leq 2 | Y = 1)$ olasılıklarını hesaplayınız.

(d) $P(Y = y | X = x)$ koşullu olasılık fonksiyonunu oluşturunuz.

Çözüm.

Önce f 'nin bir olasılık fonksiyonu olduğunu görelim: değerler pozitifdir ve

$$\sum_{x=1}^3 \sum_{y=1}^2 \frac{x+y}{21} = \frac{1}{21} \sum_{x=1}^3 (2x+3) = \frac{5+7+9}{21} = 1.$$

(a) X 'in marjinali:

$$f_X(x) = \sum_{y=1}^2 \frac{x+y}{21} = \frac{2x+3}{21}, \quad x \in \{1, 2, 3\};$$

yani $f_X(1) = \frac{5}{21}$, $f_X(2) = \frac{7}{21} = \frac{1}{3}$, $f_X(3) = \frac{9}{21} = \frac{3}{7}$. Y 'nin marjinali:

$$f_Y(y) = \sum_{x=1}^3 \frac{x+y}{21} = \frac{6+3y}{21} = \frac{2+y}{7}, \quad y \in \{1, 2\};$$

yani $f_Y(1) = \frac{3}{7}$, $f_Y(2) = \frac{4}{7}$. Ortak dağılım tablosu:

$x \setminus y$	1	2	$f_X(x)$
1	$\frac{2}{21}$	$\frac{3}{21}$	$\frac{5}{21}$
2	$\frac{3}{21}$	$\frac{4}{21}$	$\frac{7}{21}$
3	$\frac{4}{21}$	$\frac{5}{21}$	$\frac{9}{21}$
$f_Y(y)$	$\frac{9}{21}$	$\frac{12}{21}$	1

(b) Destek $\{1, 2, 3\} \times \{1, 2\}$ çarpım kümesidir; destek ölçütü karar vermez, hücreleri sınamak gerekir:

$$f(1, 1) = \frac{2}{21} = \frac{14}{147}, \quad f_X(1) f_Y(1) = \frac{5}{21} \cdot \frac{3}{7} = \frac{15}{147}.$$

Eşit değildir; X ve Y **bağımsız değildir**. Nedeni yapısalıdır: $x + y$ ifadesi $g(x)h(y)$ biçiminde çarpanlara ayrılamaz.

(c) Tanım gereği

$$f_{X|Y=y}(x) = \frac{f(x, y)}{f_Y(y)} = \frac{(x+y)/21}{(6+3y)/21} = \frac{x+y}{6+3y}, \quad x \in \{1, 2, 3\}.$$

$y = 1$ için $f_{X|Y=1}(x) = \frac{x+1}{9}$: değerler $\frac{2}{9}, \frac{3}{9}, \frac{4}{9}$; $y = 2$ için $f_{X|Y=2}(x) = \frac{x+2}{12}$: değerler $\frac{3}{12}, \frac{4}{12}, \frac{5}{12}$. Her ikisinin toplamı 1'dir ve ikisi birbirinden (ve marjinal $\frac{5}{21}, \frac{7}{21}, \frac{9}{21}$ 'den) farklıdır; **Teorem 14.2** bağımlılığı bir kez daha doğrular.

($1 \leq X \leq 3$) olayı X 'in bütün değer kümesini kapsar; dolayısıyla

$$P(1 \leq X \leq 3 | Y = 1) = \sum_{x=1}^3 f_{X|Y=1}(x) = \frac{2+3+4}{9} = 1.$$

Daha bilgilendirici bir olasılık:

$$P(X \leq 2 | Y = 1) = f_{X|Y=1}(1) + f_{X|Y=1}(2) = \frac{2}{9} + \frac{3}{9} = \frac{5}{9}.$$

(d) Satırları satır toplamına bölersek

$$f_{Y|X=x}(y) = \frac{f(x, y)}{f_X(x)} = \frac{x+y}{2x+3}, \quad y \in \{1, 2\}.$$

Somut olarak: $X = 1$ için $\frac{2}{5}, \frac{3}{5}$; $X = 2$ için $\frac{3}{7}, \frac{4}{7}$; $X = 3$ için $\frac{4}{9}, \frac{5}{9}$. Her satırın toplamı 1'dir; X büyüdükçe $Y = 2$ olma koşullu olasılığı azalır.

■

14.5 Sürekli Örnekler

İş akışı hep aynıdır: destek bölgesi çizilir; marjinaler kesitler üzerinde integralle bulunur; koşullu yoğunluk **yalnızca ilgili kesit üzerinde** yazılır; koşullu olasılıklar koşullu yoğunluğun kesit içindeki integralidir. Destek bölgesi bir çarpım kümesi ($D_X \times D_Y$) değilse — örneğin bir üçgen — değişkenler bağımsız olamaz.

Örnek 14.7 (Dikdörtgen Üzerinde Düzgün Ortak Yoğunluk). X ve Y rastgele değişkenlerinin ortak olasılık yoğunluk fonksiyonu

$$f(x, y) = \begin{cases} c, & 0 < x < 2, 2 < y < 4 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun.

- (a) f 'nin bir olasılık yoğunluk fonksiyonu olması için c ne olmalıdır?
- (b) f_X ve f_Y marjinal yoğunluklarını bulunuz. X ve Y bağımsız mıdır?
- (c) $f_{Y|X=x}$ ve $f_{X|Y=y}$ koşullu yoğunluklarını bulunuz.
- (d) $P(2 < Y < 3 \mid X = 1)$, $P(2 < Y < 5 \mid X = 1)$ ve $P(2 < Y < 4 \mid X = 1)$ olasılıklarını hesaplayınız.
- (e) $P(X \leq 1, Y \leq 3)$ ve $P(X + Y \leq 3)$ olasılıklarını hesaplayınız.

Çözüm.

(a) $f \geq 0$ için $c \geq 0$ gerekir. Destek, kenar uzunlukları 2 olan bir karedir; alanı 4'tür:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dy dx = \int_0^2 \int_2^4 c dy dx = 4c = 1 \Rightarrow c = \frac{1}{4}.$$

(b) Sırasıyla $0 < x < 2$ ve $2 < y < 4$ için

$$f_X(x) = \int_2^4 \frac{1}{4} dy = \frac{1}{2}, \quad f_Y(y) = \int_0^2 \frac{1}{4} dx = \frac{1}{2};$$

bu aralıkların dışında marjinaler sıfırdır. X , $(0, 2)$ üzerinde; Y , $(2, 4)$ üzerinde düzgün dağılmıştır. Her (x, y) için $f(x, y) = f_X(x) f_Y(y)$ sağlanır: destek karesinin içinde $\frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2}$, dışında iki taraf da 0'dır. Öyleyse X ve Y **bağımsızdır**.

(c) $0 < x < 2$ için D 'nin x apsisindeki düşey kesiti $(2, 4)$ aralığıdır:

$$f_{Y|X=x}(y) = \frac{f(x, y)}{f_X(x)} = \frac{1/4}{1/2} = \frac{1}{2}, \quad 2 < y < 4;$$

kesit dışında 0. Benzer biçimde $2 < y < 4$ için $f_{X|Y=y}(x) = \frac{1}{2}$, $0 < x < 2$. Koşullu yoğunluklar marjinalere eşittir; **Teorem 14.2** ile uyumludur.

(d) $X = 1$ verilmişken Y 'nin koşullu yoğunluğu $(2, 4)$ üzerinde $\frac{1}{2}$ 'dir:

$$P(2 < Y < 3 \mid X = 1) = \int_2^3 \frac{1}{2} dy = \frac{1}{2}.$$

$(2, 5)$ aralığının $(2, 4)$ dışında kalan kısmında koşullu yoğunluk sıfırdır; integral yalnızca $(2, 4)$ üzerinde alınır:

$$P(2 < Y < 5 \mid X = 1) = \int_2^4 \frac{1}{2} dy = 1, \quad P(2 < Y < 4 \mid X = 1) = \int_2^4 \frac{1}{2} dy = 1.$$

(e) Bağımsızlık sayesinde ilk olasılık çarpıma ayrılır:

$$P(X \leq 1, Y \leq 3) = P(X \leq 1) P(Y \leq 3) = \left(\int_0^1 \frac{1}{2} dx \right) \left(\int_2^3 \frac{1}{2} dy \right) = \frac{1}{4}.$$

İkincisi için bölgeyi belirleyelim: destek karesi içinde $x + y \leq 3$ koşulu $y \leq 3 - x$ demektir; $y > 2$ olduğundan bu ancak $x < 1$ iken mümkündür. Bölge, köşeleri $(0, 2)$, $(1, 2)$, $(0, 3)$ olan dik üçgendir:

$$P(X + Y \leq 3) = \int_0^1 \int_2^{3-x} \frac{1}{4} dy dx = \frac{1}{4} \int_0^1 (1-x) dx = \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8}.$$

■

Örnek 14.8 (Köşegen Altındaki Üçgende Koşullu Yoğunluklar). X ve Y rastgele değişkenlerinin ortak olasılık yoğunluk fonksiyonu

$$f(x, y) = \begin{cases} \frac{1}{x}, & 0 < y < x < 1 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun.

(a) f 'nin bir ortak olasılık yoğunluk fonksiyonu olduğunu gösteriniz.

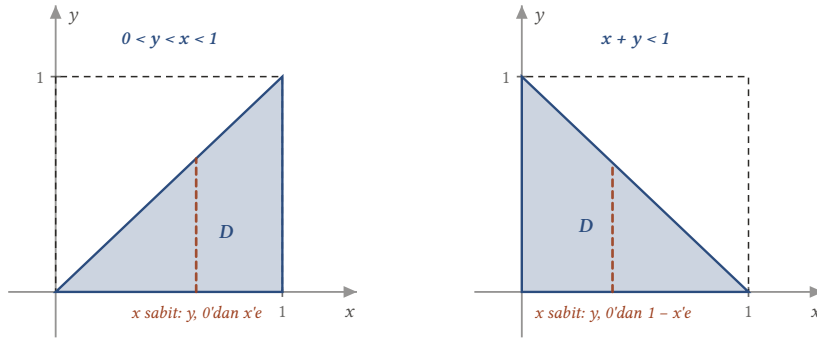
(b) X ve Y 'nin marjinal yoğunluklarını bulunuz.

(c) $(X \leq \frac{1}{2}, Y \leq \frac{1}{4})$ olasılığını hesaplayınız.

(d) X ve Y bağımsız mıdır?

(e) $f_{Y|X=x}$ ve $f_{X|Y=y}$ koşullu yoğunluklarını bulunuz.

Çözüm.



Şekil 14.3: Ortak yoğunluğun sıfırdan farklı olduğu bölge D , marjinal ve koşullu yoğunlukları hesaplamak için integral sınırlarını belirler. Solda $D = \{0 < y < x < 1\}$: sabit bir x için y , 0'dan x 'e gider. Sağda $D = \{x + y < 1\}$: y , 0'dan $1 - x$ 'e gider. Her iki bölge de birim karenin yarısidir.

Destek. $D = \{(x, y) : 0 < y < x < 1\}$, birim karede $y = x$ köşegeninin altında kalan üçgendir. Sabit bir $x \in (0, 1)$ için düşey kesit $0 < y < x$; sabit bir $y \in (0, 1)$ için yatay kesit $y < x < 1$ aralığıdır. Şeklin solundaki panel bu bölgeyi gösterir; sağdaki panel ise aynı yöntemin uygulanacağı başka bir üçgen bölgeye, Örnek 14.10'deki desteğe aittir.

(a) D üzerinde $\frac{1}{x} > 0$ 'dır. İntegral, önce y sonra x üzerinden:

$$\int_0^1 \int_0^x \frac{1}{x} dy dx = \int_0^1 \frac{1}{x} \cdot x dx = \int_0^1 1 dx = 1.$$

Öyleyse f bir ortak olasılık yoğunluk fonksiyonudur (Tanım 13.6).

(b) $0 < x < 1$ için düşey kesit üzerinde integral:

$$f_X(x) = \int_0^x \frac{1}{x} dy = 1, \quad 0 < x < 1.$$

X , $(0, 1)$ üzerinde düzgün dağılmıştır. $0 < y < 1$ için yatay kesit üzerinde integral:

$$f_Y(y) = \int_y^1 \frac{1}{x} dx = \ln 1 - \ln y = -\ln y, \quad 0 < y < 1.$$

$0 < y < 1$ için $\ln y < 0$ olduğundan $f_Y(y) > 0$ 'dır; ayrıca $\int_0^1 (-\ln y) dy = [y - y \ln y]_0^1 = 1$ (burada $y \rightarrow 0^+$ iken $y \ln y \rightarrow 0$ 'dır).

(c) Olayın bölgesi D ile $x \leq \frac{1}{2}$, $y \leq \frac{1}{4}$ yarı düzlemlerinin kesişimidir. Önce x üzerinden integral alalım: $y \in (0, \frac{1}{4})$ sabitken x, y 'den $\frac{1}{2}$ 'ye kadar değişir:

$$\left(PX \leq \frac{1}{2}, Y \leq \frac{1}{4}\right) = \int_0^{1/4} \int_y^{1/2} \frac{1}{x} dx dy = \int_0^{1/4} \left(\ln \frac{1}{2} - \ln y\right) dy = -\frac{\ln 2}{4} + \int_0^{1/4} (-\ln y) dy.$$

Son integral $[y - y \ln y]_0^{1/4} = \frac{1}{4} - \frac{1}{4} \ln \frac{1}{4} = \frac{1}{4} + \frac{\ln 2}{2}$ olur. Böylece

$$\left(PX \leq \frac{1}{2}, Y \leq \frac{1}{4}\right) = -\frac{\ln 2}{4} + \frac{1}{4} + \frac{\ln 2}{2} = \frac{1 + \ln 2}{4} \approx 0,4233.$$

(d) Destek bir üçgendir, çarpım kümesi değildir; dolayısıyla X ve Y **bağımsız olamaz**. Somut olarak $(x, y) = (\frac{1}{4}, \frac{1}{2})$ noktasında $y > x$ olduğundan $f(x, \frac{1}{2}) = 0$, ama $f(\frac{1}{4})f(\frac{1}{2}) = 1 \cdot \ln 2 \neq 0$ 'dır.

(e) $0 < x < 1$ için, düşey kesit $(0, x)$ üzerinde

$$f_{Y|X=x}(y) = \frac{f(x, y)}{f_X(x)} = \frac{1/x}{1} = \frac{1}{x}, \quad 0 < y < x;$$

kesit dışında 0. Yani $X = x$ verilmişken Y , $(0, x)$ üzerinde **düzgün** dağılır. $0 < y < 1$ için, yatay kesit $(y, 1)$ üzerinde

$$f_{X|Y=y}(x) = \frac{f(x, y)}{f_Y(y)} = \frac{1/x}{-\ln y} = -\frac{1}{x \ln y}, \quad y < x < 1;$$

kesit dışında 0. Sağlama: $\int_y^1 \frac{-1}{x \ln y} dx = \frac{-\ln y}{-\ln y} = 1$. Her iki koşullu yoğunluk da koşula bağlıdır; bu, (d)'deki sonucu [Teorem 14.2](#) yoluyla doğrular.

İki aşamalı yorum: önce $(0, 1)$ üzerinde düzgün bir X , sonra $(0, X)$ üzerinde düzgün bir Y seçilir; çarpım kuralı $f(x, y) = f_X(x) f_{Y|X=x}(y) = 1 \cdot \frac{1}{x}$ ortak yoğunluğu verir.

■

Örnek 14.9 (Aynı Üçgende Farklı Bir Yoğunluk). X ve Y rastgele değişkenlerinin ortak olasılık yoğunluk fonksiyonu

$$f(x, y) = \begin{cases} 3x, & 0 < y < x < 1 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun.

(a) Marjinal olasılık yoğunluk fonksiyonlarını bulunuz.

(b) $f_{Y|X=x}$ ve $f_{X|Y=y}$ koşullu yoğunluklarını bulunuz. X ve Y bağımsız mıdır?

Çözüm.

Destek, [Örnek 14.8](#)'deki üçgenin aynısıdır; kesitler de aynıdır. Önce f 'nin yoğunluk olduğunu doğrulayalım: $\int_0^1 \int_0^x 3x dy dx = \int_0^1 3x^2 dx = 1$.

(a) Sırasıyla $0 < x < 1$ ve $0 < y < 1$ için

$$f_X(x) = \int_0^x 3x dy = 3x^2, \quad f_Y(y) = \int_y^1 3x dx = \frac{3}{2}(1 - y^2).$$

Sağlama: $\int_0^1 3x^2 dx = 1$ ve $\int_0^1 \frac{3}{2}(1 - y^2) dy = \frac{3}{2}(1 - \frac{1}{3}) = 1$.

(b) $0 < x < 1$ için düşey kesit $(0, x)$ üzerinde

$$f_{Y|X=x}(y) = \frac{3x}{3x^2} = \frac{1}{x}, \quad 0 < y < x.$$

$0 < y < 1$ için yatay kesit $(y, 1)$ üzerinde

$$f_{X|Y=y}(x) = \frac{3x}{\frac{3}{2}(1-y^2)} = \frac{2x}{1-y^2}, \quad y < x < 1.$$

Sağlama: $\int_y^1 \frac{2x}{1-y^2} dx = \frac{1-y^2}{1-y^2} = 1$. **Örnek 14.8'**deki destek ölçütü burada da geçerlidir: X ve Y **bağımsız değildir**; koşullu yoğunlukların koşula bağlı olması bunu ayrıca gösterir.

$X = x$ verilmişken Y 'nin koşullu dağılımı **Örnek 14.8'**dekiyle aynıdır ($(0, x)$ üzerinde düzgün); ortak yoğunlukların farkı X 'in marjinalinden gelir: $f(x, y) = f_X(x) f_{Y|X=x}(y) = 3x^2 \cdot \frac{1}{x} = 3x$.

■

Örnek 14.10 (Dik Üçgen Bölgede Düzgün Yoğunluk). X ve Y rastgele değişkenlerinin ortak olasılık yoğunluk fonksiyonu

$$f(x, y) = \begin{cases} c, & 0 < x < 1, 0 < y < 1, x + y < 1 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun.

(a) c sabitini ve X ile Y 'nin marjinal yoğunluklarını bulunuz.

(b) $P(X > 2Y)$ olasılığını hesaplayınız.

(c) $f_{Y|X=x}$ ve $f_{X|Y=y}$ koşullu yoğunluklarını bulunuz.

(d) $(\frac{1}{2} \leq Y \leq \frac{7}{8} \mid X = \frac{1}{3})$ olasılığını hesaplayınız.

Çözüm.

Destek. D , köşeleri $(0, 0)$, $(1, 0)$, $(0, 1)$ olan üçgendir. Sabit $x \in (0, 1)$ için düşey kesit $0 < y < 1 - x$; sabit $y \in (0, 1)$ için yatay kesit $0 < x < 1 - y$ aralığıdır. **Örnek 14.8'**deki şeklin sağ paneli tam olarak bu bölgeyi ve düşey kesitini gösterir.

(a) Üçgenin alanı $\frac{1}{2}$ olduğundan yoğunluğun integrali $c \cdot \frac{1}{2} = 1$ olmalıdır; $c = 2$. Doğrudan hesapla:

$$\int_0^1 \int_0^{1-x} c \, dy \, dx = c \int_0^1 (1-x) \, dx = \frac{c}{2} = 1 \Rightarrow c = 2.$$

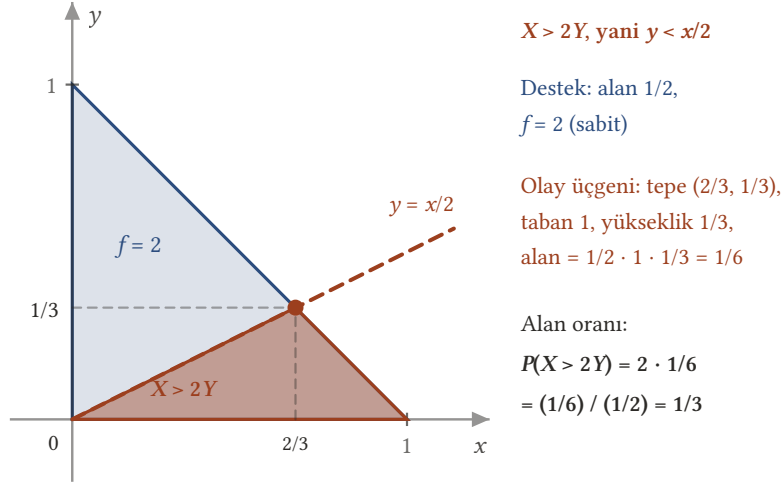
Marjinaller:

$$f_X(x) = \int_0^{1-x} 2 \, dy = 2(1-x), \quad 0 < x < 1; \quad f_Y(y) = \int_0^{1-y} 2 \, dx = 2(1-y), \quad 0 < y < 1.$$

Destek çarpım kümesi olmadığından X ve Y bağımsız değildir; örneğin $f(\frac{3}{4}, \frac{3}{4}) = 0$ ama $f(\frac{3}{4})f(\frac{3}{4}) = \frac{1}{4} \neq 0$.

(b) $X > 2Y$ koşulu $y < \frac{x}{2}$ demektir. $y = \frac{x}{2}$ doğrusu $x + y = 1$ doğrusunu $(\frac{2}{3}, \frac{1}{3})$ noktasında keser. Olayın bölgesi, köşeleri $(0, 0)$, $(1, 0)$, $(\frac{2}{3}, \frac{1}{3})$ olan üçgendir: tabanı 1, yüksekliği $\frac{1}{3}$, alanı $\frac{1}{6}$. Yoğunluk sabit olduğundan

$$P(X > 2Y) = 2 \cdot \frac{1}{6} = \frac{1}{3}.$$



Şekil 14.4: Sabit yoğunlukta olasılık alan oranına döner. Açık taralı üçgen $x + y < 1$ desteğidir; üzerinde $f = 2$. $X > 2Y$ olayı $y < x/2$ demektir: kesikli $y = x/2$ doğrusu desteği $(2/3, 1/3)$ noktasında keser ve olay, köşeleri $(0, 0)$, $(1, 0)$, $(2/3, 1/3)$ olan koyu üçgendir. Alanı $1/6$ olduğundan $P(X > 2Y) = 2 \cdot 1/6 = 1/3$; yani olayın alanının desteğin alanına oranı.

İntegralle doğrulayalım. Düşey kesitin üst sınırı $x < \frac{2}{3}$ iken $\frac{x}{2}$, $x > \frac{2}{3}$ iken $1 - x$ 'tir:

$$P(X > 2Y) = \int_0^{2/3} \int_0^{x/2} 2 \, dy \, dx + \int_{2/3}^1 \int_0^{1-x} 2 \, dy \, dx = \int_0^{2/3} x \, dx + \int_{2/3}^1 2(1-x) \, dx = \frac{2}{9} + \frac{1}{9} = \frac{1}{3}.$$

(c) $0 < x < 1$ için düşey kesit $(0, 1 - x)$ üzerinde

$$f_{Y|X=x}(y) = \frac{2}{2(1-x)} = \frac{1}{1-x}, \quad 0 < y < 1-x;$$

yani $X = x$ verilmişken Y , $(0, 1 - x)$ üzerinde düzgündür. Simetriyle, $0 < y < 1$ için

$$f_{X|Y=y}(x) = \frac{1}{1-y}, \quad 0 < x < 1-y.$$

Koşullu dağılımlar koşula bağlıdır: x büyüdükçe Y 'nin sığabileceği aralık daralır.

(d) $X = \frac{1}{3}$ verilmişken Y 'nin koşullu yoğunluğu $(0, \frac{2}{3})$ aralığında $\frac{1}{1-1/3} = \frac{3}{2}$, dışında 0'dır. İstenen aralığın kesit dışında kalan $(\frac{2}{3}, \frac{7}{8}]$ parçası olasılığa katkı vermez; integral yalnızca $[\frac{1}{2}, \frac{2}{3}]$ üzerinde alınır:

$$\left(\frac{1}{2} \leq Y \leq \frac{7}{8} \mid X = \frac{1}{3} \right) = \int_{1/2}^{2/3} \frac{3}{2} \, dy = \frac{3}{2} \left(\frac{2}{3} - \frac{1}{2} \right) = \frac{1}{4}.$$

Sınırı $\frac{7}{8}$ 'e kadar götürmek $\frac{9}{16}$ gibi yanlış bir sonuç verirdi; $X = \frac{1}{3}$ iken Y , $\frac{2}{3}$ 'ü aşamaz.

■

14.6 Alıştırmalar

Alıştırma 14.1 (Bağımsızlık ve Koşullu Dağılımlar). (a) X ve Y 'nin ortak olasılık fonksiyonu $f(x, y) = \frac{xy}{36}$, $x, y \in \{1, 2, 3\}$ olsun. Marjinal olasılık fonksiyonlarını bulunuz; X ve Y 'nin bağımsız olup olmadığını belirleyiniz; $P(X = 3 \mid Y = 2)$ olasılığını hesaplayınız.

(b) Örnek 14.10'deki ortak yoğunluk için $(\frac{1}{8} \leq Y \leq \frac{1}{2} \mid X = \frac{1}{3})$ ve $(X \leq \frac{1}{2} \mid Y = \frac{1}{4})$ olasılıklarını hesaplayınız.

(c) X ve Y 'nin ortak yoğunluğu $f(x, y) = e^{-x-y}$, $x > 0$, $y > 0$ (diğer durumlarda 0) olsun. X ve Y 'nin bağımsız olduğunu gösteriniz; $f_{X|Y=y}$ koşullu yoğunluğunu yazınız ve $P(X > 1 | Y = 2)$ olasılığını hesaplayınız.

(d) Örnek 14.9'deki ortak yoğunluk için $(P > \frac{1}{2} | X = \frac{3}{4})$ ve $(P > \frac{1}{2} | Y = \frac{1}{4})$ olasılıklarını hesaplayınız.

(e) X ve Y bağımsız kesikli rastgele değişkenler, g ve h reel değerli fonksiyonlar olsun. $U = g(X)$ ve $V = h(Y)$ rastgele değişkenlerinin de bağımsız olduğunu gösteriniz.

Çözüm.

(a) Önce $\sum_{x,y} \frac{xy}{36} = \frac{(1+2+3)(1+2+3)}{36} = 1$, dolayısıyla f bir olasılık fonksiyonudur. Marjinaler:

$$f_X(x) = \sum_{y=1}^3 \frac{xy}{36} = \frac{6x}{36} = \frac{x}{6}, \quad f_Y(y) = \frac{y}{6}, \quad x, y \in \{1, 2, 3\}.$$

Her (x, y) için $f(x, y) = \frac{x}{6} \cdot \frac{y}{6} = f_X(x)f_Y(y)$ sağlanır; destek de $\{1, 2, 3\}^2$ çarpım kümesidir. X ve Y bağımsızdır. Teorem 14.2 gereği $P(X = 3 | Y = 2) = f_X(3) = \frac{1}{2}$; doğrudan: $\frac{f(3,2)}{f_Y(2)} = \frac{6/36}{2/6} = \frac{1}{2}$.

(b) $X = \frac{1}{3}$ verilmişken Y , $(0, \frac{2}{3})$ üzerinde $\frac{3}{2}$ yoğunluğuyla düzgündür. $[\frac{1}{8}, \frac{1}{2}]$ aralığı tamamen kesit içindedir:

$$\left(\frac{1}{8} \leq Y \leq \frac{1}{2} | X = \frac{1}{3}\right) = \frac{3}{2} \left(\frac{1}{2} - \frac{1}{8}\right) = \frac{3}{2} \cdot \frac{3}{8} = \frac{9}{16}.$$

$Y = \frac{1}{4}$ verilmişken X , $(0, \frac{3}{4})$ üzerinde $\frac{4}{1-1/4} = \frac{4}{3}$ yoğunluğuyla düzgündür:

$$\left(PX \leq \frac{1}{2} | Y = \frac{1}{4}\right) = \int_0^{1/2} \frac{4}{3} dx = \frac{2}{3}.$$

(c) $f(x, y) = e^{-x} \cdot e^{-y}$ çarpımıdır ve destek $(0, \infty) \times (0, \infty)$ çarpım kümesidir; çarpanlara ayrılma ölçütü gereği X ve Y bağımsızdır. Gerçekten $f_X(x) = \int_0^\infty e^{-x-y} dy = e^{-x}$ ($x > 0$), $f_Y(y) = e^{-y}$ ($y > 0$) ve $f = f_X f_Y$. Bağımsızlık gereği $f_{X|Y=y}(x) = f_X(x) = e^{-x}$, $x > 0$ (her $y > 0$ için). Böylece

$$P(X > 1 | Y = 2) = \int_1^\infty e^{-x} dx = e^{-1} \approx 0,3679.$$

(d) $X = \frac{3}{4}$ verilmişken Y , $(0, \frac{3}{4})$ üzerinde $\frac{4}{3}$ yoğunluğuyla düzgündür:

$$\left(PY > \frac{1}{2} | X = \frac{3}{4}\right) = \frac{4}{3} \left(\frac{3}{4} - \frac{1}{2}\right) = \frac{1}{3}.$$

$Y = \frac{1}{4}$ verilmişken X 'in koşullu yoğunluğu $\frac{2x}{1-1/16} = \frac{32x}{15}$, $\frac{1}{4} < x < 1$ 'dir:

$$\left(PX > \frac{1}{2} | Y = \frac{1}{4}\right) = \int_{1/2}^1 \frac{32x}{15} dx = \frac{16}{15} \left(1 - \frac{1}{4}\right) = \frac{4}{5}.$$

(e) $(U = u)$ olayı, $g(x) = u$ olan $x \in D_X$ 'ler için $(X = x)$ olaylarının ayrık birleşimidir; $(V = v)$ için de aynısı geçerlidir. Bağımsızlık gereği

$$P(U = u, V = v) = \sum_{\substack{x \in D_X \\ g(x)=u}} \sum_{\substack{y \in D_Y \\ h(y)=v}} f_X(x) f_Y(y) = \left(\sum_{\substack{x \in D_X \\ g(x)=u}} f_X(x) \right) \left(\sum_{\substack{y \in D_Y \\ h(y)=v}} f_Y(y) \right) = P(U = u) P(V = v).$$

Her (u, v) için bu eşitlik sağlandığından U ve V bağımsızdır: bağımsız değişkenlerin fonksiyonları da bağımsızdır.

■

Bağımsızlık ve koşullu dağılım, bir rastgele vektörün bileşenleri arasındaki ilişkiyi tanımlar; bir sonraki adım, bu vektörün fonksiyonlarının (toplam, oran, en büyük değer gibi) dağılımını bulmaktır: **Rastgele Vektörlerin Dönüşümleri**.

15. Rastgele Vektörlerin Dönüşümleri

Bağımsız Rastgele Değişkenler ve Koşullu Dağılımlar bölümünde bir rastgele vektörün ortak dağılımından koşullu dağılımlara nasıl geçildiğini gördük. Bu bölümde soruyu bir adım ileri taşıyoruz: (X_1, X_2) rastgele vektörünün ortak dağılımı biliniyorsa, $X_1 + X_2$, $X_1 X_2$, X_1/X_2 gibi **bileşenlerin fonksiyonlarının** dağılımı nasıl bulunur? Daha genel olarak, $Y_1 = u_1(X_1, X_2)$ ve $Y_2 = u_2(X_1, X_2)$ ile tanımlanan yeni (Y_1, Y_2) vektörünün ortak dağılımı nedir?

Tek bir rastgele değişkenin fonksiyonu için bu soruyu **Rastgele Değişkenin Fonksiyonunun Dağılımı** bölümünde yanıtlamıştık: kesikli durumda olasılıklar toplanır, sürekli durumda ya dağılım fonksiyonu tekniği kullanılır ya da birebir dönüşümler için $f_Y(y) = f_X(h(y)) |h'(y)|$ formülü uygulanır (**Teorem 12.2**). Burada aynı iki yöntemi iki boyuta taşıyacağız. Kesikli durum yine bir toplama işidir. Sürekli durumda ise $|h'(y)|$ çarpanının yerini, ters dönüşümün kısmi türevlerinden oluşan **Jacobian matrisinin determinantının mutlak değeri** alır. Bu, bölümün ana teoremidir ve arkasındaki fikir, çok katlı integralde değişken değiştirme formülünden başka bir şey değildir.

Bölümde ayrıca iki pratik konuya değineceğiz. Birincisi, yalnızca tek bir fonksiyonun, örneğin $Y = X_1 + X_2$ 'nin dağılımı istendiğinde ne yapılacağı: ya **yapay bir ikinci değişken** ekleyip Jacobian yöntemini kullanır ve sonra bu değişkeni integrale atarız ya da doğrudan dağılım fonksiyonu tekniğini uyguluyoruz. İkincisi, **bağımsız** iki değişkenin toplamının yoğunluğunu veren **konvolüsyon** formülüdür. Bütün bu araçlar, bir sonraki kısımda beklenen değeri ve momentleri incelerken ve daha sonra önemli dağılımların toplamalarını hesaplarken sürekli karşımıza çıkacaktır.

15.1 Kesikli Rastgele Vektörlerin Dönüşümleri

Kesikli bir rastgele vektörün fonksiyonu yine kesiklidir ve olasılık fonksiyonu, tek değişkenli durumda olduğu gibi, aynı görüntüyü veren bütün noktaların olasılıklarının toplanmasıyla bulunur. Bunu bir önerme olarak yazalım; ispat, kesikli rastgele vektör tanımı (**Tanım 13.3**) ile olasılık ölçüsünün sayılabilir toplamsallığından ibarettir.

Önerme 15.1 (Kesikli Rastgele Vektörün Dönüşümü). (X_1, X_2) , değer kümesi $D_{(X_1, X_2)}$ ve ortak olasılık fonksiyonu f_{X_1, X_2} olan kesikli bir rastgele vektör olsun. $u_1, u_2 : \mathbb{R}^2 \rightarrow \mathbb{R}$ fonksiyonları için

$$Y_1 = u_1(X_1, X_2), \quad Y_2 = u_2(X_1, X_2)$$

olsun. Bu durumda (Y_1, Y_2) de kesikli bir rastgele vektördür; değer kümesi

$$D_{(Y_1, Y_2)} = \left\{ (u_1(x_1, x_2), u_2(x_1, x_2)) : (x_1, x_2) \in D_{(X_1, X_2)} \right\}$$

ve ortak olasılık fonksiyonu, $(y_1, y_2) \in D_{(Y_1, Y_2)}$ için

$$f_{Y_1, Y_2}(y_1, y_2) = \sum_{\substack{(x_1, x_2) \in D_{(X_1, X_2)} \\ u_1(x_1, x_2) = y_1, u_2(x_1, x_2) = y_2}} f_{X_1, X_2}(x_1, x_2)$$

ile verilir. Özel olarak, tek bir $Y = u(X_1, X_2)$ fonksiyonu için

$$f_Y(y) = \sum_{\substack{(x_1, x_2) \in D_{(X_1, X_2)} \\ u(x_1, x_2) = y}} f_{X_1, X_2}(x_1, x_2), \quad y \in D_Y$$

olur.

İspat.

$D_{(X_1, X_2)}$ sayılabilir bir kümedir; $D_{(Y_1, Y_2)}$ bu kümenin (u_1, u_2) altındaki görüntüsü olduğundan o da sayılabilir. $P((X_1, X_2) \in D_{(X_1, X_2)}) = 1$ olduğundan ve her $(x_1, x_2) \in D_{(X_1, X_2)}$ noktası (u_1, u_2)

altında $D_{(Y_1, Y_2)}$ 'nin bir noktasına gittiğinden $P((Y_1, Y_2) \in D_{(Y_1, Y_2)}) = 1$ olur. Demek ki (Y_1, Y_2) kesiklidir.

Şimdi $(y_1, y_2) \in D_{(Y_1, Y_2)}$ sabit olsun ve

$$A = \{ (x_1, x_2) \in D_{(X_1, X_2)} : u_1(x_1, x_2) = y_1, u_2(x_1, x_2) = y_2 \}$$

kümesini tanımlayalım. $(Y_1 = y_1, Y_2 = y_2)$ olayı ile $((X_1, X_2) \in A)$ olayı, (X_1, X_2) 'nin $D_{(X_1, X_2)}$ dışına düştüğü sıfır olasılıklı N olayı bir yana, aynıdır:

$$(Y_1 = y_1, Y_2 = y_2) = ((X_1, X_2) \in A) \cup N, \quad N \subseteq ((X_1, X_2) \notin D_{(X_1, X_2)}), \quad P(N) = 0.$$

A sayılabilir olduğundan, A 'daki tek noktalı olaylar ayrıktır ve olasılık ölçüsünün sayılabilir toplamsallığıyla

$$f_{Y_1, Y_2}(y_1, y_2) = P((X_1, X_2) \in A) = \sum_{(x_1, x_2) \in A} P(X_1 = x_1, X_2 = x_2) = \sum_{(x_1, x_2) \in A} f_{X_1, X_2}(x_1, x_2)$$

elde edilir. Tek fonksiyonlu durum, u_2 'nin hiç kısıt koymadığı $(Y_2$ 'nin göz ardı edildiği) özel hâldir; aynı akıl yürütme $A = \{(x_1, x_2) \in D_{(X_1, X_2)} : u(x_1, x_2) = y\}$ kümesiyle geçerlidir.

■

Formülün söylediği şudur: $(Y_1, Y_2) = (y_1, y_2)$ olayının olasılığı, bu değeri üreten bütün (x_1, x_2) çiftlerinin olasılıklarının toplamıdır. Uygulamada değer kümesi küçükse en güvenli yol, (X_1, X_2) 'nin bütün noktalarını tek tek dolaşp her birini gittiği (y_1, y_2) hücresine eklemektir.

Örnek 15.1 (İki Kesikli Değişkenin Toplamı). X ve Y rastgele değişkenlerinin ortak olasılık fonksiyonu

$$f(x, y) = \frac{xy}{36}, \quad x \in \{1, 2, 3\}, y \in \{1, 2, 3\}$$

olsun. $Z = X + Y$ rastgele değişkeninin olasılık fonksiyonunu bulunuz.

Çözüm.

Önce f 'nin gerçekten bir olasılık fonksiyonu olduğunu görelim: $\sum_{x,y} \frac{xy}{36} = \frac{(1+2+3)(1+2+3)}{36} = 1$. Ayrıca $f(x, y) = \frac{x}{6} \cdot \frac{y}{6}$ biçiminde çarpanlara ayrıldığından X ile Y bağımsızdır ve marjinal olasılık fonksiyonları $f_X(x) = \frac{x}{6}$, $f_Y(y) = \frac{y}{6}$ 'dir.

Z 'nin değer kümesi $D_Z = \{2, 3, 4, 5, 6\}$ 'dir. **Önerme 15.1**'na göre her z için $x + y = z$ olan çiftlerin olasılıkları toplanır:

$$P(Z = 2) = f(1, 1) = \frac{1}{36},$$

$$P(Z = 3) = f(1, 2) + f(2, 1) = \frac{2}{36} + \frac{2}{36} = \frac{4}{36},$$

$$P(Z = 4) = f(1, 3) + f(2, 2) + f(3, 1) = \frac{3}{36} + \frac{4}{36} + \frac{3}{36} = \frac{10}{36},$$

$$P(Z = 5) = f(2, 3) + f(3, 2) = \frac{6}{36} + \frac{6}{36} = \frac{12}{36},$$

$$P(Z = 6) = f(3, 3) = \frac{9}{36}.$$

z	2	3	4	5	6
$P(Z = z)$	$\frac{1}{36}$	$\frac{4}{36}$	$\frac{10}{36}$	$\frac{12}{36}$	$\frac{9}{36}$

Toplam $\frac{1+4+10+12+9}{36} = 1$ olduğundan sonuç tutarlıdır.

■

İkinci örnekte iki fonksiyonu birlikte alıp (Y_1, Y_2) vektörünün ortak olasılık fonksiyonunu tablo hâlinde çıkaracağız.

Örnek 15.2 (Toplam ve Fark Dönüşümü). (X_1, X_2) rastgele vektörünün ortak olasılık fonksiyonu

$$f(x_1, x_2) = \left(\frac{2}{3}\right)^{x_1+x_2} \left(\frac{1}{3}\right)^{2-x_1-x_2}, \quad x_1 \in \{0, 1\}, x_2 \in \{0, 1\}$$

olsun.

a) $Y_1 = X_1 + X_2$ rastgele değişkeninin olasılık fonksiyonunu bulunuz.

b) $Y_1 = X_1 + X_2$ ve $Y_2 = X_1 - X_2$ ise (Y_1, Y_2) rastgele vektörünün ortak olasılık fonksiyonunu bulunuz.

Çözüm.

Önce f 'nin dört değerini yazalım:

$$f(0, 0) = \frac{1}{9}, \quad f(0, 1) = f(1, 0) = \frac{2}{3} \cdot \frac{1}{3} = \frac{2}{9}, \quad f(1, 1) = \frac{4}{9}.$$

Toplam $\frac{1+2+2+4}{9} = 1$ 'dir. Ayrıca

$$f(x_1, x_2) = \left[\left(\frac{2}{3}\right)^{x_1} \left(\frac{1}{3}\right)^{1-x_1} \right] \left[\left(\frac{2}{3}\right)^{x_2} \left(\frac{1}{3}\right)^{1-x_2} \right]$$

olduğundan X_1 ile X_2 bağımsızdır (Tanım 14.1) ve her biri $\frac{2}{3}$ olasılıkla 1, $\frac{1}{3}$ olasılıkla 0 değerini alır.

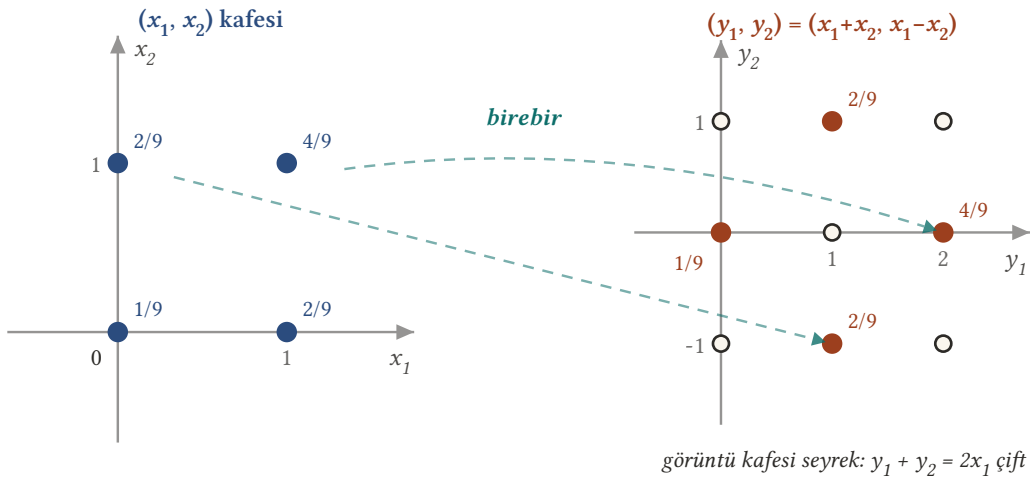
a) $Y_1 = X_1 + X_2 \in \{0, 1, 2\}$. Her değeri üreten çiftleri toplayalım:

$$P(Y_1 = 0) = f(0, 0) = \frac{1}{9}, \quad P(Y_1 = 1) = f(0, 1) + f(1, 0) = \frac{4}{9}, \quad P(Y_1 = 2) = f(1, 1) = \frac{4}{9}.$$

y_1	0	1	2
$P(Y_1 = y_1)$	$\frac{1}{9}$	$\frac{4}{9}$	$\frac{4}{9}$

b) (x_1, x_2) çiftlerinin $(y_1, y_2) = (x_1 + x_2, x_1 - x_2)$ altındaki görüntüleri:

$$(0, 0) \mapsto (0, 0), \quad (0, 1) \mapsto (1, -1), \quad (1, 0) \mapsto (1, 1), \quad (1, 1) \mapsto (2, 0).$$



Şekil 15.1: Toplam ve fark dönüşümü. Solda (x_1, x_2) kafesi $\{0, 1\}^2$ ve her noktaya düşen olasılık; sağda $(y_1, y_2) = (x_1 + x_2, x_1 - x_2)$ altındaki görüntüleri. Dört nokta dört farklı görüntüye gider, dönüşüm birebirdir ve olasılıklar olduğu gibi taşınır. Görüntü kafesinin yalnızca $y_1 + y_2$ çift olan hücreleri dolar; içi boş hücrelerin olasılığı sıfırdır.

Dört nokta dört farklı görüntüye gittiğinden dönüşüm birebirdir ve her hücreye tek bir olasılık düşer; öteki hücreler sıfırdır. Satırlarda y_2 , sütunlarda y_1 olmak üzere ortak olasılık fonksiyonu ve marjinaler şöyledir:

$y_2 \setminus y_1$	0	1	2	$P(Y_2 = y_2)$
-1	0	$\frac{2}{9}$	0	$\frac{2}{9}$
0	$\frac{1}{9}$	0	$\frac{4}{9}$	$\frac{5}{9}$
1	0	$\frac{2}{9}$	0	$\frac{2}{9}$
$P(Y_1 = y_1)$	$\frac{1}{9}$	$\frac{4}{9}$	$\frac{4}{9}$	1

Sütun toplamaları a şıkındaki tabloyu verir; bu, marjinal olasılık fonksiyonu tanımının (Tanım 13.5) beklediği şeydir. Dikkat edilirse X_1 ile X_2 bağımsız olduğu hâlde Y_1 ile Y_2 bağımsız değildir: örneğin $P(Y_1 = 0, Y_2 = -1) = 0$, ama $P(Y_1 = 0)P(Y_2 = -1) = \frac{1}{9} \cdot \frac{2}{9} \neq 0$. Bağımsız değişkenlerin fonksiyonları bağımsız kalmak zorunda değildir.

■

15.2 Sürekli Rastgele Vektörler: Jacobian Yöntemi

Sürekli durumda toplama yerine integral, “aynı görüntüyü veren noktalar” yerine ise dönüşümün tersi ve onun **alanı nasıl ölçeklediği** devreye girer. Tek değişkenli formülü hatırlayalım: $Y = u(X)$ birebir ve $X = h(Y)$ türevlenebilir ise $f_Y(y) = f_X(h(y)) |h'(y)|$ (Teorem 12.2). Buradaki $|h'(y)|$ çarpanı, y ekseninde küçük bir dy uzunluğunun x ekseninde $|h'(y)| dy$ uzunluğuna karşılık gelmesinden doğar. İki boyutta küçük bir $dy_1 dy_2$ dikdörtgeni, ters dönüşüm altında alanı $|J| dy_1 dy_2$ olan küçük bir paralelkenara gider; J , ters dönüşümün kısmi türevlerinden oluşan matrisin determinantıdır. Teorem tam olarak bunu söyler.

Teorem 15.1 (Jacobian Dönüşüm Teoremi). (X_1, X_2) , ortak olasılık yoğunluk fonksiyonu f_{X_1, X_2} olan sürekli bir rastgele vektör ve $D \subseteq \mathbb{R}^2$, dışında f_{X_1, X_2} ’nin sıfır olduğu (yani $P((X_1, X_2) \in D) = 1$ olan) açık bir küme olsun.

$$Y_1 = u_1(X_1, X_2), \quad Y_2 = u_2(X_1, X_2)$$

dönüşümü D üzerinde **birebir** olsun ve D ’yi $E = \{(u_1(x_1, x_2), u_2(x_1, x_2)) : (x_1, x_2) \in D\}$ kümesine götürsün. Ters dönüşüm

$$X_1 = h_1(Y_1, Y_2), \quad X_2 = h_2(Y_1, Y_2)$$

ile gösterilsin ve h_1, h_2 fonksiyonları E üzerinde sürekli birinci mertebe kısmi türevlere sahip olsun. **Jacobian**,

$$J = \frac{\partial(h_1, h_2)}{\partial(y_1, y_2)} = \det \begin{bmatrix} \frac{\partial h_1}{\partial y_1} & \frac{\partial h_1}{\partial y_2} \\ \frac{\partial h_2}{\partial y_1} & \frac{\partial h_2}{\partial y_2} \end{bmatrix}$$

determinantı olsun ve E üzerinde $J \neq 0$ olsun. Bu durumda (Y_1, Y_2) sürekli bir rastgele vektördür ve ortak olasılık yoğunluk fonksiyonu

$$g(y_1, y_2) = \begin{cases} f_{X_1, X_2}(h_1(y_1, y_2), h_2(y_1, y_2)) |J|, & (y_1, y_2) \in E \\ 0, & \text{diğer durumlarda} \end{cases}$$

ile verilir.

İspat.

İspatın özü, çok katlı integralde değişken değiştirme formülüdür. Bu formül şunu söyler: $h = (h_1, h_2) : E \rightarrow D$ birebir, sürekli kısmi türevlere sahip ve Jacobian’ı sıfırdan farklı ise, E ’nin her Borel alt kümesi B ve integrellenebilir her φ için

$$\iint_{h(B)} \varphi(x_1, x_2) dx_1 dx_2 = \iint_B \varphi(h_1(y_1, y_2), h_2(y_1, y_2)) |J(y_1, y_2)| dy_1 dy_2$$

olur. Burada $h(B) = \{h(y_1, y_2) : (y_1, y_2) \in B\} \subseteq D$, B 'nin ters dönüşüm altındaki görüntüsüdür. Formülün geometrik anlamı yukarıda anlatılan alan ölçeklemesidir: $|J|$, h 'nin bir noktanın küçük komşuluğundaki alanı kaç kat büyüttüğünü ölçer.

Şimdi $B \subseteq E$ bir Borel kümesi olsun. Dönüşüm D üzerinde birebir olduğundan, $(Y_1, Y_2) \in B$ olayı ile $(X_1, X_2) \in h(B)$ olayının $(X_1, X_2) \in D$ olayıyla kesişimleri aynıdır; $P((X_1, X_2) \in D) = 1$ olduğundan

$$P((Y_1, Y_2) \in B) = P((X_1, X_2) \in h(B)) = \iint_{h(B)} f_{X_1, X_2}(x_1, x_2) dx_1 dx_2$$

olur (Teorem 13.1 ve Tanım 13.6). Değişken değiştirme formülünü $\varphi = f_{X_1, X_2}$ ile uygularsak

$$P((Y_1, Y_2) \in B) = \iint_B f_{X_1, X_2}(h_1(y_1, y_2), h_2(y_1, y_2)) |J| dy_1 dy_2 = \iint_B g(y_1, y_2) dy_1 dy_2$$

elde ederiz. E dışındaki Borel kümeleri için iki taraf da sıfırdır, çünkü $P((Y_1, Y_2) \in E) = P((X_1, X_2) \in D) = 1$ 'dir. Demek ki (Y_1, Y_2) 'nin her Borel kümesine düşme olasılığı, negatif olmayan g fonksiyonunun o küme üzerindeki integralidir; bu, (Y_1, Y_2) 'nin sürekli olduğu ve yoğunluğunun g olduğu anlamına gelir. $B = E$ alınırsa $\iint_E g(y_1, y_2) dy_1 dy_2 = 1$ olduğu da görülür.

Teoremi uygularken izlenecek adımlar şunlardır:

1. f 'nin sıfırdan farklı olduğu D bölgesini yazın.
2. $y_1 = u_1(x_1, x_2)$, $y_2 = u_2(x_1, x_2)$ denklemlerini x_1, x_2 için çözerek h_1, h_2 'yi bulun; dönüşümün D üzerinde birebir olduğundan emin olun.
3. D 'yi tanımlayan eşitsizliklerde $x_1 = h_1(y_1, y_2)$, $x_2 = h_2(y_1, y_2)$ yazarak görüntü bölgesi E 'yi belirleyin. En çok hata bu adımda yapılır.
4. Jacobian'ı hesaplayın, mutlak değerini alın.
5. $g(y_1, y_2) = f(h_1, h_2) |J|$ yazın; E üzerindeki integralinin 1 olduğunu kontrol edin.

Jacobian'ı ters matristen hesaplamak

Bazen y_1, y_2 'yi x_1, x_2 cinsinden yazmak kolay, ama tersini çözmek zahmetlidir. Ters fonksiyon teoremine göre, h 'nin Jacobian matrisi, $u = (u_1, u_2)$ 'nin Jacobian matrisinin **ters matrisidir**:

$$\begin{bmatrix} \frac{\partial h_1}{\partial y_1} & \frac{\partial h_1}{\partial y_2} \\ \frac{\partial h_2}{\partial y_1} & \frac{\partial h_2}{\partial y_2} \end{bmatrix} = \begin{bmatrix} \frac{\partial u_1}{\partial x_1} & \frac{\partial u_1}{\partial x_2} \\ \frac{\partial u_2}{\partial x_1} & \frac{\partial u_2}{\partial x_2} \end{bmatrix}^{-1} \bigg|_{(x_1, x_2) = (h_1(y_1, y_2), h_2(y_1, y_2))}$$

Bir matrisin tersinin determinanı determinanın tersi olduğundan

$$J = \frac{\partial(h_1, h_2)}{\partial(y_1, y_2)} = \frac{1}{\frac{\partial(u_1, u_2)}{\partial(x_1, x_2)}}$$

olur; sağ taraf $x_1 = h_1(y_1, y_2)$, $x_2 = h_2(y_1, y_2)$ noktasında hesaplanır. Yani Jacobian'ı ileri dönüşümün determinantını alıp tersini çevirerek de bulabiliriz. Örneğin $y_1 = x_1 + x_2$, $y_2 = x_1$ için $\frac{\partial(u_1, u_2)}{\partial(x_1, x_2)} = \det \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = -1$, dolayısıyla $J = -1$ ve $|J| = 1$ 'dir.

İlk örneğimizde dönüşüm doğrusaldır ve görüntü bölgesini çizmek, teoremin “alan ölçekleme” yorumunu doğrudan görmemizi sağlar.

Örnek 15.3 (Birim Kare Üzerinde Toplam ve Birinci Bileşen). (X_1, X_2) rastgele vektörünün ortak olasılık yoğunluk fonksiyonu

$$f_{X_1, X_2}(x_1, x_2) = \begin{cases} 1, & 0 < x_1 < 1, 0 < x_2 < 1 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun. $Y_1 = X_1 + X_2$, $Y_2 = X_1$ dönüşümü ile tanımlanan (Y_1, Y_2) rastgele vektörünün ortak olasılık yoğunluk fonksiyonunu bulunuz. Ardından $Y_1 = X_1 + X_2$ 'nin marjinal yoğunluğunu hesaplayınız.

Çözüm.

Bölge ve ters dönüşüm. $D = \{(x_1, x_2) : 0 < x_1 < 1, 0 < x_2 < 1\}$ birim karedir. $y_1 = x_1 + x_2$, $y_2 = x_1$ denklemlerinden

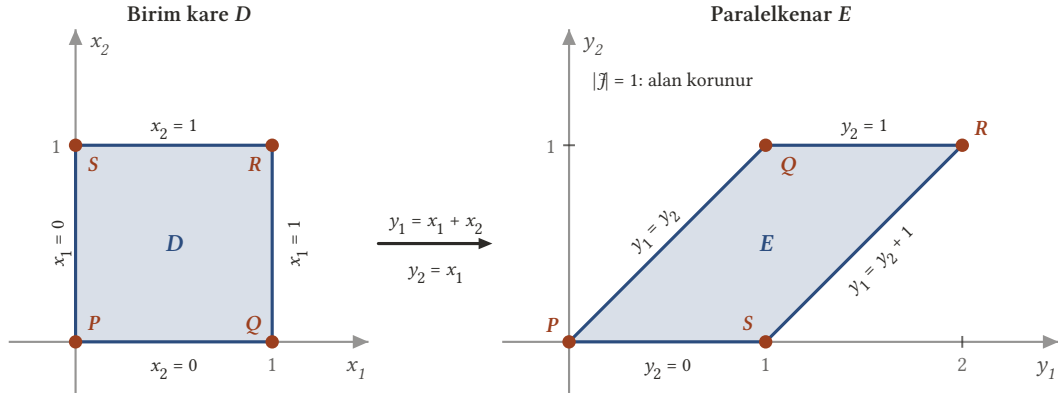
$$x_1 = h_1(y_1, y_2) = y_2, \quad x_2 = h_2(y_1, y_2) = y_1 - y_2$$

bulunur. Her (y_1, y_2) için tek bir (x_1, x_2) elde edildiğinden dönüşüm birebirdir.

Görüntü bölgesi. D 'nin eşitsizliklerinde $x_1 = y_2$, $x_2 = y_1 - y_2$ yazalım:

$$0 < y_2 < 1, \quad 0 < y_1 - y_2 < 1 \Leftrightarrow y_2 < y_1 < y_2 + 1.$$

Demek ki $E = \{(y_1, y_2) : 0 < y_2 < 1, y_2 < y_1 < y_2 + 1\}$ 'dir. Bu, köşeleri $(0, 0)$, $(1, 0)$, $(2, 1)$ ve $(1, 1)$ olan bir paralelkenardır: birim karenin köşeleri sırasıyla $(0, 0) \mapsto (0, 0)$, $(1, 0) \mapsto (1, 1)$, $(0, 1) \mapsto (1, 0)$, $(1, 1) \mapsto (2, 1)$ noktalarına gider.



Şekil 15.2: $Y_1 = X_1 + X_2$, $Y_2 = X_1$ doğrusal dönüşümü birim kare D 'yi paralelkenar E 'ye taşır. Köşeleri $P(0,0) \mapsto (0,0)$, $Q(1,0) \mapsto (1,1)$, $R(1,1) \mapsto (2,1)$, $S(0,1) \mapsto (1,0)$ noktalarına gider; D 'nin $x_1 = 0$ ve $x_1 = 1$ kenarları $y_2 = 0$ ve $y_2 = 1$ doğrularına, $x_2 = 0$ ve $x_2 = 1$ kenarları $y_1 = y_2$ ve $y_1 = y_2 + 1$ doğrularına düşer. $|J| = 1$ olduğundan alan korunur: iki bölgenin de alanı 1'dir ve E üzerinde yoğunluk 1 kalır.

Jacobian.

$$J = \det \begin{bmatrix} \frac{\partial h_1}{\partial y_1} & \frac{\partial h_1}{\partial y_2} \\ \frac{\partial h_2}{\partial y_1} & \frac{\partial h_2}{\partial y_2} \end{bmatrix} = \det \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} = 0 \cdot (-1) - 1 \cdot 1 = -1, \quad |J| = 1.$$

Yoğunluk. Teorem 15.1'a göre

$$g(y_1, y_2) = f_{X_1, X_2}(y_2, y_1 - y_2) |J| = \begin{cases} 1, & 0 < y_2 < 1, y_2 < y_1 < y_2 + 1 \\ 0, & \text{diğer durumlarda.} \end{cases}$$

Paralelkenarın tabanı 1, yüksekliği 1 olduğundan alanı 1'dir ve $\iint_E g(y_1, y_2) dy_1 dy_2 = 1$ sağlanır; $|J| = 1$ olması alanın korunduğunu söyler.

Y_1 'in marjinal yoğunluğu. Marjinal yoğunluk tanımına (Tanım 13.8) göre $g_{Y_1}(y_1) = \int_{-\infty}^{\infty} g(y_1, y_2) dy_2$ 'dir. Sabit bir y_1 için g 'nin 1 olduğu y_2 değerleri, $0 < y_2 < 1$ ve $y_1 - 1 < y_2 < y_1$ koşullarını birlikte sağlayanlardır.

- $0 < y_1 < 1$ ise $y_1 - 1 < 0$ olduğundan alt sınır 0, $y_1 < 1$ olduğundan üst sınır y_1 olur; koşul $0 < y_2 < y_1$ 'e iner: $g_{Y_1}(y_1) = \int_0^{y_1} 1 dy_2 = y_1$.
- $1 \leq y_1 < 2$ ise $y_1 - 1 \geq 0$ olduğundan alt sınır $y_1 - 1$, $y_1 \geq 1$ olduğundan üst sınır 1 olur; koşul $y_1 - 1 < y_2 < 1$ 'e iner: $g_{Y_1}(y_1) = \int_{y_1-1}^1 1 dy_2 = 2 - y_1$.

Böylece

$$g_{Y_1}(y_1) = \begin{cases} y_1, & 0 < y_1 < 1 \\ 2 - y_1, & 1 \leq y_1 < 2 \\ 0, & \text{diğer durumlarda.} \end{cases}$$

Bu, grafiği $(0,0)$, $(1,1)$, $(2,0)$ köşeli bir üçgen olan **üçgen yoğunluktur**; $\int_0^1 y_1 dy_1 + \int_1^2 (2 - y_1) dy_1 = \frac{1}{2} + \frac{1}{2} = 1$.

■

15.3 Tek Bir Dönüşüm: Yapay Değişken ve Dağılım Fonksiyonu Tekniği

Çoğu zaman elimizde tek bir fonksiyon vardır: $Y = u(X_1, X_2)$. Jacobian teoremi iki boyuttan iki boyuta birebir dönüşümler için yazıldığından doğrudan uygulanamaz. Bunun iki yolu vardır; son örnekte aslında birincisini kullandık: $Y_1 = X_1 + X_2$ istiyorduk, $Y_2 = X_1$ 'i biz ekledik, ortak yoğunluğu bulduk ve sonra Y_2 'yi integrale attık.

Önerme 15.2 (Tek Bir Dönüşümün Yoğunluğu). (X_1, X_2) , ortak olasılık yoğunluk fonksiyonu f olan sürekli bir rastgele vektör ve $Y = u(X_1, X_2)$ olsun. Y 'nin yoğunluğu aşağıdaki iki yoldan biriyle bulunur.

(1) Yapay değişken yöntemi. (Y, W) dönüşümü Teorem 15.1'un koşullarını sağlayacak biçimde bir $W = v(X_1, X_2)$ yardımcı değişkeni seçilir (yoğunlukla $W = X_1$ ya da $W = X_2$). Teoreme (Y, W) 'nin ortak yoğunluğu $g(y, w)$ bulunur ve

$$f_Y(y) = \int_{-\infty}^{\infty} g(y, w) dw$$

hesaplanır.

(2) Dağılım fonksiyonu tekniği. Her $y \in \mathbb{R}$ için

$$F_Y(y) = P(u(X_1, X_2) \leq y) = \iint_{\{(x_1, x_2): u(x_1, x_2) \leq y\}} f(x_1, x_2) dx_1 dx_2$$

hesaplanır; F_Y sürekli ve sonlu sayıda nokta dışında türevlenebilir çıkarsa Y sürekli ve türevlenebildiği noktalarda $f_Y(y) = F_Y'(y)$ alınır.

İspat.

(1) Teorem 15.1, (Y, W) 'nin sürekli olduğunu ve yoğunluğunun g olduğunu verir. Sürekli bir rastgele vektörün birinci bileşeninin yoğunluğu, ortak yoğunluğun ikinci değişkene göre integralidir (Tanım 13.8); bu tam olarak yazılan formüldür.

(2) $(Y \leq y)$ olayı, (X_1, X_2) 'nin $A_y = \{(x_1, x_2) : u(x_1, x_2) \leq y\}$ kümesine düşmesi olayıdır; sürekli rastgele vektör tanımına göre bu olasılık f 'nin A_y üzerindeki integralidir (Teorem 13.1). Böylece F_Y elde edilir. F_Y sürekli ve sonlu sayıda nokta dışında türevlenebilir çıktığında türevinin bir yoğunluk olduğu Teorem 11.2'den bilinmektedir; Tanım 12.1'nde tek değişken için yapılan akıl yürütme aynen geçerlidir.

■

Hangi yolun kolay olduğu probleme bağlıdır. Dönüşüm doğrusal ya da çarpım/oran biçimindeyse yapay değişken yöntemi genellikle daha mekaniktir; $\{u(x_1, x_2) \leq y\}$ bölgesi kolay çiziliyorsa dağılım fonksiyonu tekniği daha kısa olabilir. Aşağıda birini birinci, ötekini ikinci yolla çözüyoruz; ikinci örnekte sonucu her iki yolla da doğruluyoruz.

Örnek 15.4 (Üçgen Bölge Üzerinde Toplamın Yoğunluğu). (X_1, X_2) rastgele vektörünün ortak olasılık yoğunluk fonksiyonu

$$f(x_1, x_2) = \begin{cases} 2, & 0 \leq x_1 \leq x_2 \leq 1 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun. $Y = X_1 + X_2$ rastgele değişkeninin olasılık yoğunluk fonksiyonunu bulunuz.

Çözüm.

$D = \{(x_1, x_2) : 0 < x_1 < x_2 < 1\}$ bölgesi, birim karenin köşegeninin üstünde kalan alanı $\frac{1}{2}$ olan üçgendir; $f = 2$ olduğundan toplam olasılık 1'dir. (Sınır doğruları olasılık sıfırlı olduğundan $\leq$ yerine $<$ yazmak sonucu değiştirmez.)

Yapay değişken. $Y = X_1 + X_2$ ve $W = X_2$ alalım. $y = x_1 + x_2$, $w = x_2$ denklemlerinden

$$x_1 = y - w, \quad x_2 = w$$

bulunur; dönüşüm birebirdir. Jacobian

$$J = \det \begin{bmatrix} \frac{\partial x_1}{\partial y} & \frac{\partial x_1}{\partial w} \\ \frac{\partial x_2}{\partial y} & \frac{\partial x_2}{\partial w} \end{bmatrix} = \det \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = 1$$

olup $|J| = 1$ 'dir.

Görüntü bölgesi. $0 < x_1 < x_2 < 1$ eşitsizliklerinde yerine koyarsak

$$0 < y - w < w < 1 \Leftrightarrow \frac{y}{2} < w < y \text{ ve } w < 1.$$

İlk eşitsizlik $(0 < y - w)$ $w < y$, ikincisi $(y - w < w)$ $w > \frac{y}{2}$ verir. Ayrıca $y - w > 0$ ve $w > 0$ olduğundan $y = (y - w) + w > 0$; $\frac{y}{2} < w < 1$ olduğundan da $y < 2$ olmalıdır. Öyleyse

$$E = \left\{ (y, w) : 0 < y < 2, \frac{y}{2} < w < \min(y, 1) \right\}$$

ve E üzerinde $g(y, w) = f(y - w, w) |J| = 2$ 'dir.

Marjinal. $\min(y, 1)$ ifadesi y 'nin 1'den küçük ya da büyük olmasına göre değiştiğinden iki durum vardır.

- $0 < y < 1$ ise w aralığı $(\frac{y}{2}, y)$ 'dir:

$$f_Y(y) = \int_{y/2}^y 2 dw = 2 \left(y - \frac{y}{2} \right) = y.$$

- $1 \leq y < 2$ ise w aralığı $(\frac{y}{2}, 1)$ 'dir:

$$f_Y(y) = \int_{y/2}^1 2 dw = 2 \left(1 - \frac{y}{2} \right) = 2 - y.$$

Sonuç:

$$f_Y(y) = \begin{cases} y, & 0 < y < 1 \\ 2 - y, & 1 \leq y < 2 \\ 0, & \text{diğer durumlarda.} \end{cases}$$

Doğrulama. Yoğunluğun altındaki toplam alan 1'dir:

$$\int_0^1 y dy + \int_1^2 (2 - y) dy = \frac{1}{2} + \left[2y - \frac{y^2}{2} \right]_1^2 = \frac{1}{2} + \left(2 - \frac{3}{2} \right) = 1.$$

Ayrıca dağılım fonksiyonu tekniğiyle bir noktada kontrol edelim: $P(Y \leq 1) = P(X_1 + X_2 \leq 1)$ olasılığı, D üçgeninin $x_1 + x_2 \leq 1$ doğrusunun altında kalan parçası üzerinde f 'nin integralidir. Bu

parça, köşeleri $(0, 0)$, $(0, 1)$, $(\frac{1}{2}, \frac{1}{2})$ olan alanı $\frac{1}{4}$ olan üçgendir; yoğunluk 2 olduğundan $P(Y \leq 1) = \frac{1}{2}$. Bulduğumuz yoğunluktan $\int_0^1 y dy = \frac{1}{2}$, uyuyor.

Bu yoğunluk, **Örnek 15.3**'de bulunan üçgen yoğunlukla aynıdır. Yani birim kare üzerindeki düzgün dağılım ile üçgen üzerindeki $f = 2$ dağılımı, birbirinden çok farklı ortak yoğunluklar olmalarına karşın, $X_1 + X_2$ için aynı dağılımı verir. Toplamın dağılımı ortak dağılımı belirlemez.

■

Örnek 15.5 (Çarpımın Yoğunluğu). (X, Y) rastgele vektörünün ortak olasılık yoğunluk fonksiyonu

$$f(x, y) = \begin{cases} x + y, & 0 < x < 1, 0 < y < 1 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun. $Z = XY$ rastgele değişkeninin olasılık yoğunluk fonksiyonunu bulunuz.

Çözüm.

$\int_0^1 \int_0^1 (x + y) dy dx = \int_0^1 (x + \frac{1}{2}) dx = 1$ olduğundan f geçerli bir yoğunluktur. $0 < X < 1$ ve $0 < Y < 1$ olduğundan $0 < Z < 1$ 'dir.

Dağılım fonksiyonu tekniği. $0 < z < 1$ için $F_Z(z) = P(XY \leq z)$, birim karenin $xy \leq z$ hiperbolünün altında kalan parçası üzerinde f 'nin integralidir. x sabitken $xy \leq z$ koşulu $y \leq \frac{z}{x}$ demektir; $x \leq z$ ise $\frac{z}{x} \geq 1$ olduğundan bütün $y \in (0, 1)$ değerleri uygundur, $x > z$ ise yalnızca $y < \frac{z}{x}$ olanlar. Öyleyse

$$F_Z(z) = \int_0^z \int_0^1 (x + y) dy dx + \int_z^1 \int_0^{z/x} (x + y) dy dx.$$

Birinci integral:

$$\int_0^z \left(x + \frac{1}{2}\right) dx = \frac{z^2}{2} + \frac{z}{2}.$$

İkinci integralde iç integral $\int_0^{z/x} (x + y) dy = z + \frac{z^2}{2x^2}$ olduğundan

$$\int_z^1 \left(z + \frac{z^2}{2x^2}\right) dx = z(1 - z) + \frac{z^2}{2} \left[-\frac{1}{x}\right]_z^1 = z - z^2 + \frac{z^2}{2} \left(\frac{1}{z} - 1\right) = z - z^2 + \frac{z}{2} - \frac{z^2}{2} = \frac{3z}{2} - \frac{3z^2}{2}.$$

Toplayınca

$$F_Z(z) = \frac{z^2}{2} + \frac{z}{2} + \frac{3z}{2} - \frac{3z^2}{2} = 2z - z^2 = z(2 - z), \quad 0 < z < 1.$$

$z \leq 0$ için $F_Z(z) = 0$, $z \geq 1$ için $F_Z(z) = 1$ 'dir ve $F_Z(1) = 1$ ile uyuyor. Türev alarak

$$f_Z(z) = F'_Z(z) = 2 - 2z = 2(1 - z), \quad 0 < z < 1$$

ve diğer durumlarda $f_Z(z) = 0$ bulunur. $\int_0^1 2(1 - z) dz = 1$ olduğu görülür.

Yapay değişkenle doğrulama. $Z = XY$, $W = X$ alalım. $x = w$, $y = \frac{z}{w}$ olur; $w > 0$ olduğundan dönüşüm birebirdir.

$$J = \det \begin{bmatrix} \frac{\partial x}{\partial z} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial z} & \frac{\partial y}{\partial w} \end{bmatrix} = \det \begin{bmatrix} 0 & 1 \\ \frac{1}{w} & -\frac{z}{w^2} \end{bmatrix} = -\frac{1}{w}, \quad |J| = \frac{1}{w}.$$

Görüntü bölgesi: $0 < w < 1$ ve $0 < \frac{z}{w} < 1$, yani $0 < z < w < 1$. Böylece $g(z, w) = (w + \frac{z}{w}) \frac{1}{w}$ ($0 < z < w < 1$) ve

$$f_Z(z) = \int_z^1 \left(1 + \frac{z}{w^2}\right) dw = (1 - z) + z \left(\frac{1}{z} - 1\right) = 1 - z + 1 - z = 2(1 - z), \quad 0 < z < 1.$$

İki yöntem aynı sonucu verdi.

■

💡 Görüntü bölgesini bulmanın güvenli yolu

Yapay değişken yönteminde en sık yapılan hata, f' 'nin sıfırdan farklı olduğu bölgeyi görüntü bölgesine doğru taşıyamamaktır. Güvenli yol şudur: D' 'yi tanımlayan **her** eşitsizliğe $x_1 = h_1(y, w)$, $x_2 = h_2(y, w)$ yazın, elde ettiğiniz eşitsizlikleri w için çözün ve w 'nin alt-üst sınırlarının y 'ye göre nasıl değiştiğini belirleyin. Sınırlar min ya da max içeriyorsa, **Örnek 15.4**'de olduğu gibi, y 'nin aralığını parçalara ayırın. Son olarak bulduğunuz f_Y 'nin integralinin 1 olduğunu kontrol edin; bu kontrol bölge hatalarının çoğunu yakalar.

15.4 Bağımsız Değişkenlerin Toplamı: Konvolüsyon

En sık karşılaşılan dönüşüm, bağımsız iki değişkenin toplamıdır. Bağımsızlık durumunda ortak yoğunluk marjinalerin çarpımı olduğundan (**Tanım 14.1**), yapay değişken yöntemi genel bir formüle indirgenir. Bu formül, iki fonksiyonun **konvolüsyonu** adı verilen işlemidir ve ileride önemli dağılımların toplamlarını hesaplamak için temel araç olacaktır.

Önerme 15.3 (Toplamın Dağılımı: Konvolüsyon Formülü). X ve Y bağımsız rastgele değişkenler ve $Z = X + Y$ olsun.

(a) X ve Y , yoğunlukları f_X ve f_Y olan sürekli rastgele değişkenler ise Z sürekli ve yoğunluğu

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y) dy$$

ile verilir.

(b) X ve Y , olasılık fonksiyonları f_X ve f_Y olan kesikli rastgele değişkenler ise

$$f_Z(z) = \sum_{x \in D_X} f_X(x) f_Y(z-x)$$

olur; burada $y = z - x \notin D_Y$ ise $f_Y(z-x) = 0$ alınır.

İspat.

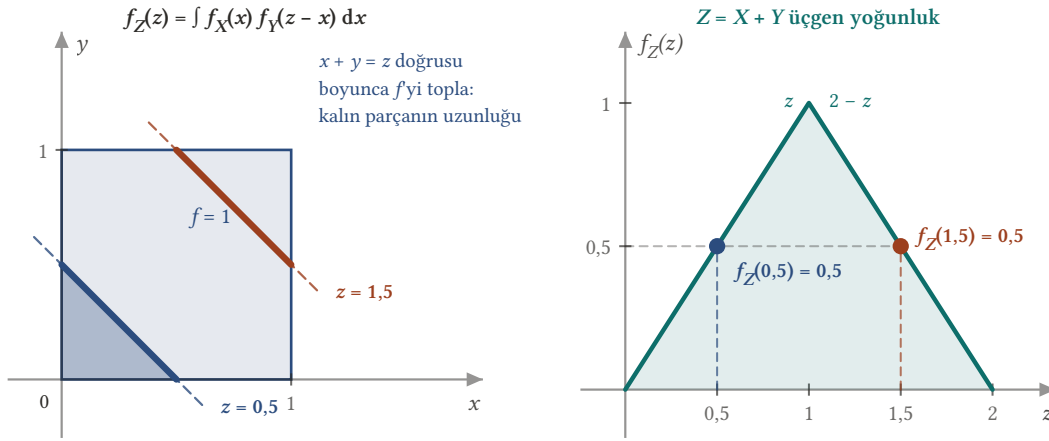
(a) Bağımsızlık gereği (X, Y) 'nin ortak yoğunluğu $f(x, y) = f_X(x) f_Y(y)$ 'dir. $Z = X + Y$ ve $W = X$ yapay değişkeni ile $x = w$, $y = z - w$ olur; dönüşüm bütün $\mathbb{R}^2$ üzerinde birebirdir ve

$$J = \det \begin{bmatrix} \frac{\partial x}{\partial z} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial z} & \frac{\partial y}{\partial w} \end{bmatrix} = \det \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} = -1, \quad |J| = 1.$$

Teorem 15.1'a göre (Z, W) 'nin ortak yoğunluğu $g(z, w) = f_X(w) f_Y(z-w)$ 'dir. **Önerme 15.2**'e göre W 'yi integrale atarsak

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(w) f_Y(z-w) dw$$

elde edilir; bu, ilk formüldür. İkinci biçim, integralde $y = z - w$ ($dw = -dy$, sınırlar yer değiştirir) değişken değişikliğiyle çıkar. Yoğunluklar yalnızca sıfırdan farklı oldukları yerde katkı verdiğinden, integral pratikte $f_X(x) \neq 0$ ve $f_Y(z-x) \neq 0$ koşullarını birlikte sağlayan x değerleri üzerinden alınır.



Şekil 15.3: Konvolüsyon formülünün geometrisi. X ve Y bağımsız, $(0, 1)$ üzerinde düzgün olsun; ortak yoğunluk birim karede $f = 1$ 'dir. $f_Z(z)$, $x + y = z$ doğrusu boyunca f 'nin toplamıdır: yalnızca karenin içindeki kalın parça katkı verir ve bu parçanın x -uzunluğu $z = 0,5$ ile $z = 1,5$ için $0,5$ 'tir. Doğru karede kaydırsa bu uzunluk önce z , sonra $2 - z$ olur; sağdaki üçgen yoğunluk böyle doğar. Sol panelde açık taralı $x + y \leq 0,5$ üçgeni $F_Z(0,5) = 1/8$ olasılığıdır.

(b) Önerme 15.1'ne göre $f_Z(z) = \sum_{x+y=z} f_{X,Y}(x, y)$; toplamda $y = z - x$ olduğundan ve bağımsızlıkla $f_{X,Y}(x, y) = f_X(x)f_Y(y)$ olduğundan

$$f_Z(z) = \sum_{x \in D_X} f_X(x) f_Y(z - x)$$

bulunur. $z - x \notin D_Y$ olan terimler sıfırdır.

Formülü hemen tanıdık bir örnekte sınavalım. Örnek 15.3'de birim kare üzerindeki $f = 1$ ortak yoğunluğu $f_{X_1}(x_1)f_{X_2}(x_2)$ biçiminde çarpanlarına ayrıldığından X_1 ile X_2 bağımsızdır ve her biri $(0, 1)$ üzerinde düzgün dağılır: $f_{X_i}(t) = 1$ ($0 < t < 1$). Konvolüsyon formülünde $f_{X_1}(x)f_{X_2}(y_1 - x) \neq 0$ olması için $0 < x < 1$ ve $0 < y_1 - x < 1$, yani $\max(0, y_1 - 1) < x < \min(1, y_1)$ gerekir. $0 < y_1 < 1$ için integral $\int_0^{y_1} 1 dx = y_1$, $1 \leq y_1 < 2$ için $\int_{y_1-1}^1 1 dx = 2 - y_1$ olur. Bu, orada Jacobian yöntemiyle bulduğumuz üçgen yoğunluğun ta kendisidir.

Kesikli biçim de Örnek 15.1'yi yeniden verir: orada $f_X(x) = \frac{x}{6}$, $f_Y(y) = \frac{y}{6}$ idi ve örneğin $P(Z = 4) = \sum_{x=1}^3 \frac{x}{6} \cdot \frac{4-x}{6} = \frac{3+4+3}{36} = \frac{10}{36}$.

15.5 İki Kapsamlı Örnek

Bu kısımdaki iki örnek, bölümün bütün araçlarını bir arada kullanır. Birincisinde tanım bölgesi dikdörtgen değildir ve yoğunluk simetrik; ikincisinde iki bağımsız değişkenden bir oran ile bir toplam üretilir ve elde edilen değişkenlerin bağımsız olup olmadığı sorulur.

Örnek 15.6 (Parabolik Bölge Üzerinde Çarpımın Yoğunluğu). X ve Y rastgele değişkenlerinin ortak olasılık yoğunluk fonksiyonu

$$f(x, y) = \begin{cases} \frac{21}{4} x^2 y, & x^2 < y < 1 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olarak tanımlansın. $Z = XY$ rastgele değişkeninin olasılık yoğunluk fonksiyonunu bulunuz ve Z 'nin yoğunluğunu kullanarak $\int_{-\infty}^{\infty} z f_Z(z) dz$ değerini hesaplayınız. (Bu integral, bir sonraki kısımda tanımlanacak olan $E(XY)$ beklenen değeridir.)

Çözüm.

Bölge. $x^2 < y < 1$ koşulu $x^2 < 1$, yani $-1 < x < 1$ gerektirir. Öyleyse

$$D = \{(x, y) : -1 < x < 1, x^2 < y < 1\},$$

parabol ile $y = 1$ doğrusu arasında kalan bölgedir. Yoğunluğun geçerliliğini kontrol edelim:

$$\iint_D f(x, y) dx dy = \frac{21}{4} \int_{-1}^1 x^2 \int_{x^2}^1 y dy dx = \frac{21}{4} \int_{-1}^1 x^2 \cdot \frac{1-x^4}{2} dx = \frac{21}{8} \cdot 2 \int_0^1 (x^2 - x^6) dx = \frac{21}{4} \left(\frac{1}{3} - \frac{1}{7} \right) = \frac{21}{4} \cdot \frac{4}{21} = 1.$$

İntegral yalnızca $0 < x < 1$ üzerinden alınsaydı sonuç $\frac{1}{2}$ çıkardı; bölge $x = 0$ doğrusuna göre simetrik iki parçadan oluştuğu için x aralığı $(-1, 1)$ alınmalıdır.

Simetri. $f(-x, y) = f(x, y)$ olduğundan (X, Y) ile $(-X, Y)$ aynı dağılıma sahiptir; dolayısıyla $Z = XY$ ile $-Z = (-X)Y$ aynı dağılıma sahiptir, yani $f_Z(-z) = f_Z(z)$. $|X| < 1$ ve $0 < Y < 1$ olduğundan $-1 < Z < 1$ 'dir. Bu yüzden yoğunluğu yalnızca $0 < z < 1$ için hesaplamak yeterlidir; negatif tarafı simetriyle tamamlarız.

Yapay değişken. $Z = XY$, $W = X$ alalım: $x = w$, $y = \frac{z}{w}$. $\{x = 0\}$ doğrusunun olasılığı sıfır olduğundan onu D 'den atabiliriz; kalan açık küme üzerinde dönüşüm birebirdir. **Örnek 15.5'**de olduğu gibi $|J| = \frac{1}{|w|}$ 'dir. Görüntü bölgesi için D 'nin eşitsizliklerinde yerine koyalım:

$$-1 < w < 1, \quad w^2 < \frac{z}{w} < 1.$$

$z > 0$ olsun. $\frac{z}{w} > w^2 > 0$ olması için $w > 0$ gerekir; o zaman $w^2 < \frac{z}{w}$ eşitsizliği $w^3 < z$, yani $w < z^{1/3}$; $\frac{z}{w} < 1$ eşitsizliği $w > z$ verir. $0 < z < 1$ için $z < z^{1/3}$ olduğundan aralık boş değildir:

$$E_+ = \{(z, w) : 0 < z < 1, z < w < z^{1/3}\}.$$

Bu bölgede

$$g(z, w) = \frac{21}{4} w^2 \cdot \frac{z}{w} \cdot \frac{1}{w} = \frac{21}{4} z.$$

Yoğunluk. $0 < z < 1$ için

$$f_Z(z) = \int_z^{z^{1/3}} \frac{21}{4} z dw = \frac{21}{4} z (z^{1/3} - z) = \frac{21}{4} (z^{4/3} - z^2).$$

Simetriyle $-1 < z < 0$ için $f_Z(z) = f_Z(-z) = \frac{21}{4} ((-z)^{4/3} - z^2)$. Mutlak değerle tek formüle:

$$f_Z(z) = \begin{cases} \frac{21}{4} (|z|^{4/3} - z^2), & 0 < |z| < 1 \\ 0, & \text{diğer durumlarda.} \end{cases}$$

Kontrol: $\int_0^1 \frac{21}{4} (z^{4/3} - z^2) dz = \frac{21}{4} \left(\frac{3}{7} - \frac{1}{3} \right) = \frac{21}{4} \cdot \frac{2}{21} = \frac{1}{2}$; simetrik parçayla birlikte toplam 1'dir. Ayrıca $0 < z < 1$ için $z^{4/3} > z^2$ olduğundan $f_Z \geq 0$ 'dır.

$\int z f_Z(z) dz$ hesabı. f_Z sınırlı ve desteği $(-1, 1)$ olduğundan $\int_{-1}^1 |z| f_Z(z) dz < \infty$ 'dur; yani integral mutlak yakınsaktır. $z f_Z(z)$ tek fonksiyondur (f_Z çift olduğundan); simetrik $(-1, 1)$ aralığında integrali sıfırdır:

$$\int_{-1}^1 z f_Z(z) dz = 0.$$

Aynı sonucu doğrudan ortak yoğunluktan da görebiliriz: $\iint_D xy f(x, y) dx dy$ integralinde integrand x 'e göre tek, bölge $x \mapsto -x$ altında simetrik olduğundan sonuç 0'dır. Yani $E(XY) = 0$. Bu, X ile Y 'nin bağımsız olduğu anlamına gelmez; tanım bölgesi dikdörtgen olmadığından bağımsız değildirler. Bir sonraki kısımda beklenen değeri tanımladıktan sonra bu ayrıma yeniden döneceğiz.

Karşılaştırma için, Z 'nin mutlak değerinin ortalaması sıfır değildir:

$$\int_{-1}^1 |z| f_Z(z) dz = 2 \cdot \frac{21}{4} \int_0^1 (z^{7/3} - z^3) dz = \frac{21}{2} \left(\frac{3}{10} - \frac{1}{4} \right) = \frac{21}{40}.$$

■

Örnek 15.7 (Oran ve Toplam Dönüşümü). Y_1 ve Y_2 bağımsız rastgele değişkenler ve her birinin olasılık yoğunluk fonksiyonu

$$f(y) = \begin{cases} \frac{1}{y^2}, & y > 1 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun. $U_1 = \frac{Y_2}{Y_1 + Y_2}$ ve $U_2 = Y_1 + Y_2$ rastgele değişkenleri tanımlansın. (Bu örnekte dönüşümün girdisi (Y_1, Y_2) , çıktısı (U_1, U_2) vektörüdür.)

- (Y_1, Y_2) rastgele vektörünün ortak olasılık yoğunluk fonksiyonunu bulunuz.
- (U_1, U_2) rastgele vektörünün ortak olasılık yoğunluk fonksiyonunu bulunuz.
- U_1 rastgele değişkeninin olasılık yoğunluk fonksiyonunu bulunuz.
- U_1 ve U_2 rastgele değişkenlerinin bağımsız olup olmadığını inceleyiniz.

Çözüm.

Önce f 'nin bir yoğunluk olduğunu görelim: $\int_1^\infty y^{-2} dy = 1$.

a) Bağımsızlık gereği (Tanım 14.1)

$$f_{Y_1, Y_2}(y_1, y_2) = f(y_1)f(y_2) = \begin{cases} \frac{1}{y_1^2 y_2^2}, & y_1 > 1, y_2 > 1 \\ 0, & \text{diğer durumlarda.} \end{cases}$$

$D = \{(y_1, y_2) : y_1 > 1, y_2 > 1\}$ 'dir.

b) $u_1 = \frac{y_2}{y_1 + y_2}$, $u_2 = y_1 + y_2$ denklemlerini çözelim: $y_2 = u_1 u_2$ ve $y_1 = u_2 - y_2 = u_2(1 - u_1)$. Yani

$$h_1(u_1, u_2) = u_2(1 - u_1), \quad h_2(u_1, u_2) = u_1 u_2.$$

Her (u_1, u_2) için tek bir (y_1, y_2) çıktığından dönüşüm birebirdir. Jacobian:

$$J = \det \begin{bmatrix} \frac{\partial h_1}{\partial u_1} & \frac{\partial h_1}{\partial u_2} \\ \frac{\partial h_2}{\partial u_1} & \frac{\partial h_2}{\partial u_2} \end{bmatrix} = \det \begin{bmatrix} -u_2 & 1 - u_1 \\ u_2 & u_1 \end{bmatrix} = -u_1 u_2 - u_2(1 - u_1) = -u_2, \quad |J| = u_2.$$

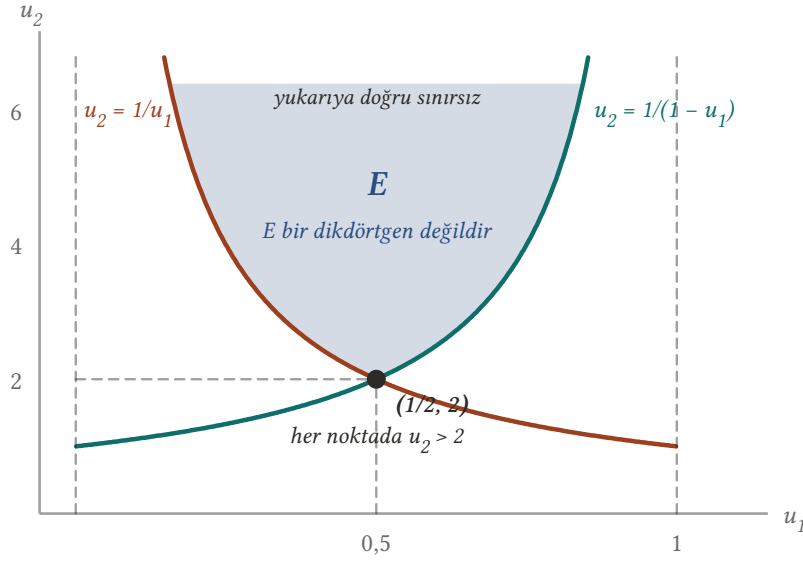
Görüntü bölgesi: $y_1 > 1$ ve $y_2 > 1$ koşulları

$$u_2(1 - u_1) > 1, \quad u_1 u_2 > 1$$

olur. $y_1, y_2 > 0$ olduğundan $u_2 > 0$ ve $0 < u_1 < 1$ 'dir; koşullar $u_2 > \frac{1}{1-u_1}$ ve $u_2 > \frac{1}{u_1}$ biçiminde yazılır. Öyleyse

$$E = \left\{ (u_1, u_2) : 0 < u_1 < 1, u_2 > \max\left(\frac{1}{u_1}, \frac{1}{1-u_1}\right) \right\}.$$

$u_1 < \frac{1}{2}$ iken $\frac{1}{u_1} > \frac{1}{1-u_1}$, $u_1 > \frac{1}{2}$ iken tersi geçerlidir; $u_1 = \frac{1}{2}$ 'de ikisi eşittir: iki sınır orada $u_2 = 2$ değerinde birleşir ve E 'nin her noktasında $u_2 > 2$ 'dir (ki bu $Y_1 + Y_2 > 2$ ile uyudur).



Şekil 15.4: $Y_1, Y_2 > 1$ bölgesinin $U_1 = Y_2/(Y_1 + Y_2)$, $U_2 = Y_1 + Y_2$ dönüşümü altındaki görüntüsü. İki sınır eğrisi $(1/2, 2)$ noktasında kesişir; E , ikisinin de üstünde kalan bölgedir ve her noktasında $u_2 > 2$ 'dir. E iki aralığın çarpımı olmadığından $-u_1$ 'in alabileceği değerler u_2 'ye bağlıdır $-U_1$ ile U_2 bağımsız olamaz.

Teorem 15.1'a göre E üzerinde

$$g(u_1, u_2) = f_{Y_1, Y_2}(u_2(1-u_1), u_1 u_2) u_2 = \frac{1}{u_2^2(1-u_1)^2 \cdot u_1^2 u_2^2} \cdot u_2 = \frac{1}{u_1^2(1-u_1)^2 u_2^3},$$

E dışında $g = 0$ 'dır.

c) U_1 'in yoğunluğu, g 'nin u_2 'ye göre integralidir; alt sınır u_1 'in $\frac{1}{2}$ 'nin hangi tarafında olduğuna bağlıdır. Her iki durumda da $\int_a^\infty u_2^{-3} du_2 = \frac{1}{2a^2}$ kullanılır.

- $0 < u_1 < \frac{1}{2}$: alt sınır $a = \frac{1}{u_1}$,

$$f_{U_1}(u_1) = \frac{1}{u_1^2(1-u_1)^2} \cdot \frac{u_1^2}{2} = \frac{1}{2(1-u_1)^2}.$$

- $\frac{1}{2} \leq u_1 < 1$: alt sınır $a = \frac{1}{1-u_1}$,

$$f_{U_1}(u_1) = \frac{1}{u_1^2(1-u_1)^2} \cdot \frac{(1-u_1)^2}{2} = \frac{1}{2u_1^2}.$$

Sonuç:

$$f_{U_1}(u_1) = \begin{cases} \frac{1}{2(1-u_1)^2}, & 0 < u_1 < \frac{1}{2} \\ \frac{1}{2u_1^2}, & \frac{1}{2} \leq u_1 < 1 \\ 0, & \text{diğer durumlarda.} \end{cases}$$

Kontrol: $\int_0^{1/2} \frac{du_1}{2(1-u_1)^2} = \frac{1}{2} \left[\frac{1}{1-u_1} \right]_0^{1/2} = \frac{1}{2}(2-1) = \frac{1}{2}$ ve $\int_{1/2}^1 \frac{du_1}{2u_1^2} = \frac{1}{2} \left[-\frac{1}{u_1} \right]_{1/2}^1 = \frac{1}{2}(-1+2) = \frac{1}{2}$; toplam 1. Yoğunluk $u_1 = \frac{1}{2}$ etrafında simetriktir ($f_{U_1}(1-u_1) = f_{U_1}(u_1)$) ve orada en büyük değeri olan 2'yi alır. Simetri beklenen bir şeydir: Y_1 ile Y_2 aynı dağılımlı olduğundan $U_1 = \frac{Y_2}{Y_1+Y_2}$ ile $1-U_1 = \frac{Y_1}{Y_1+Y_2}$ aynı dağılıma sahiptir. Aynı nedenle, oranın payında Y_2 yerine Y_1 alınsaydı bütün sonuçlar değişmezdi.

d) Bağımsız değildirler. U_1 ile U_2 bağımsız olsaydı her (u_1, u_2) için $g(u_1, u_2) = f_{U_1}(u_1)f_{U_2}(u_2)$ olması gerekirdi (Tanım 14.1); özellikle g 'nin sıfırdan farklı olduğu bölge iki aralığın kartezyen çarpımı olurdu. Oysa E öyle değildir: u_2 'nin alabileceği değerler u_1 'e bağlıdır. Somut bir çift alalım: $(u_1, u_2) = (\frac{1}{10}, 3)$. Burada $\frac{1}{u_1} = 10 > 3$ olduğundan $(u_1, u_2) \notin E$ ve $g(\frac{1}{10}, 3) = 0$ 'dır. Öte yandan $f_{U_1}(\frac{1}{10}) = \frac{1}{2(\frac{9}{10})^2} = \frac{50}{81} > 0$ ve

$$f_{U_2}(3) = \int_{-\infty}^{\infty} g(u_1, 3) du_1 = \frac{1}{27} \int_{1/3}^{2/3} \frac{du_1}{u_1^2(1-u_1)^2} > 0$$

(integrand pozitif, aralık boş değil; $u_2 = 3$ için E 'nin kesiti $\frac{1}{3} < u_1 < \frac{2}{3}$ 'tür). Demek ki $g(\frac{1}{10}, 3) = 0 \neq f_{U_1}(\frac{1}{10})f_{U_2}(3)$; U_1 ile U_2 bağımsız değildir. Sezgisel olarak da açıktır: toplam U_2 küçükse (2 'ye yakınsa) iki değişken de 1 'e yakın olmak zorundadır, dolayısıyla oran U_1 zorunlu olarak $\frac{1}{2}$ 'ye yakındır; toplam büyüdükçe oran daha serbestleşir.

⚠ Bağımsızlık dönüşümle korunmayabilir

Örnek 15.2 ve Örnek 15.7'de girdi değişkenleri bağımsızdı, ama toplam-fark ve oran-toplam çiftleri bağımsız çıkmadı. Genel kural şudur: X_1 ile X_2 bağımsızsa $u(X_1)$ ile $v(X_2)$ de bağımsızdır; ancak **her ikisi de** hem X_1 'e hem X_2 'ye bağlı olan $u_1(X_1, X_2)$ ile $u_2(X_1, X_2)$ genellikle bağımsız değildir. Bağımsızlığı korumak için özel bir yapı gerekir; bunun tanınmış bir örneğini alıştırmalarda, birim disk üzerinde düzgün dağılımın kutupsal koordinatlarında göreceğiz.

15.6 Alıştırmalar

Alıştırma 15.1 (Rastgele Vektörlerin Dönüşümleri Üzerine Alıştırmalar). a) Örnek 15.1'deki $f(x, y) = \frac{xy}{36}$ ($x, y \in \{1, 2, 3\}$) ortak olasılık fonksiyonu için $W = XY$ rastgele değişkeninin olasılık fonksiyonunu bulunuz.

b) X ve Y bağımsız ve aynı dağılımlı olsun; $P(X = k) = \frac{1}{2^k}$, $k \in \{1, 2, 3, \dots\}$. Konvolüsyon formülüyle $Z = X + Y$ 'nin olasılık fonksiyonunu bulunuz ve olasılıkların toplamının 1 olduğunu gösteriniz.

c) (X_1, X_2) birim kare üzerinde $f = 1$ yoğunluğuna sahip olsun. $Y = \frac{X_1}{X_2}$ oranının yoğunluğunu bulunuz.

d) X ve Y bağımsız olsun ve her birinin yoğunluğu $f(t) = e^{-t}$ ($t > 0$) olsun. $Z = X + Y$ 'nin yoğunluğunu bulunuz.

e) (X, Y) birim kare üzerinde $f = 1$ yoğunluğuna sahip olsun. $U = X + Y$, $V = X - Y$ dönüşümü için (U, V) 'nin ortak yoğunluğunu ve V 'nin marjinal yoğunluğunu bulunuz.

f) (X, Y) , birim disk üzerinde düzgün dağılsın: $f(x, y) = \frac{1}{\pi}(x^2 + y^2 < 1)$. $R = \sqrt{X^2 + Y^2}$ ve Θ , (X, Y) noktasının kutupsal açısı olsun ($X = R \cos \Theta$, $Y = R \sin \Theta$, $0 < \Theta < 2\pi$). (R, Θ) 'nin ortak yoğunluğunu bulunuz; R ile Θ 'nın bağımsız olduğunu gösteriniz ve R 'nin yoğunluğunu yazınız.

Çözüm.

a) $W = XY \in \{1, 2, 3, 4, 6, 9\}$. Her değer için çarpımı o değeri veren çiftlerin olasılıkları toplanır:

$$P(W = 1) = f(1, 1) = \frac{1}{36}, \quad P(W = 2) = f(1, 2) + f(2, 1) = \frac{4}{36},$$

$$P(W = 3) = f(1, 3) + f(3, 1) = \frac{6}{36}, \quad P(W = 4) = f(2, 2) = \frac{4}{36},$$

$$P(W = 6) = f(2, 3) + f(3, 2) = \frac{12}{36}, \quad P(W = 9) = f(3, 3) = \frac{9}{36}.$$

w	1	2	3	4	6	9
$P(W = w)$	$\frac{1}{36}$	$\frac{4}{36}$	$\frac{6}{36}$	$\frac{4}{36}$	$\frac{12}{36}$	$\frac{9}{36}$

Toplam $\frac{1+4+6+4+12+9}{36} = 1$.

b) $Z \in \{2, 3, 4, \dots\}$. Önerme 15.3 (b) ile, $z - x \geq 1$ olması için $x \leq z - 1$ gerekir:

$$P(Z = z) = \sum_{x=1}^{z-1} \frac{1}{2^x} \cdot \frac{1}{2^{z-x}} = \sum_{x=1}^{z-1} \frac{1}{2^z} = \frac{z-1}{2^z}, \quad z \geq 2.$$

Toplamı doğrulayalım:

$$\sum_{z=2}^{\infty} \frac{z-1}{2^z} = \sum_{m=1}^{\infty} \frac{m}{2^{m+1}} = \frac{1}{2} \sum_{m=1}^{\infty} \frac{m}{2^m} = \frac{1}{2} \cdot 2 = 1.$$

Burada $|q| < 1$ için $\sum_{m \geq 1} mq^m = \frac{q}{(1-q)^2}$ özdeşliği $q = \frac{1}{2}$ ile kullanıldı.

c) $Y = \frac{X_1}{X_2}$, $W = X_2$ alalım: $x_1 = yw$, $x_2 = w$;

$$J = \det \begin{bmatrix} w & y \\ 0 & 1 \end{bmatrix} = w, \quad |J| = w(w > 0).$$

Bölge: $0 < w < 1$ ve $0 < yw < 1$, yani $y > 0$ ve $w < \min(1, \frac{1}{y})$. $g(y, w) = 1 \cdot w = w$ olduğundan

$$f_Y(y) = \int_0^{\min(1, 1/y)} w dw = \frac{1}{2} \min\left(1, \frac{1}{y}\right)^2 = \begin{cases} \frac{1}{2}, & 0 < y \leq 1 \\ \frac{1}{2y^2}, & y > 1 \\ 0, & \text{diğer durumlarda.} \end{cases}$$

Kontrol: $\frac{1}{2} + \int_1^{\infty} \frac{dy}{2y^2} = \frac{1}{2} + \frac{1}{2} = 1$. Oranın 1'den küçük ya da büyük olma olasılıkları eşittir; bu, X_1 ile X_2 'nin aynı dağılımlı olmasından beklenen simetridir.

d) Konvolüsyon formülünde $f_X(x)f_Y(z-x) \neq 0$ için $x > 0$ ve $z-x > 0$, yani $0 < x < z$ gerekir; dolayısıyla $z > 0$ için

$$f_Z(z) = \int_0^z e^{-x} e^{-(z-x)} dx = \int_0^z e^{-z} dx = ze^{-z},$$

$z \leq 0$ için $f_Z(z) = 0$. Kontrol: $\int_0^{\infty} ze^{-z} dz = [-ze^{-z} - e^{-z}]_0^{\infty} = 1$.

e) $u = x + y$, $v = x - y$ denklemlerinden $x = \frac{u+v}{2}$, $y = \frac{u-v}{2}$;

$$J = \det \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} = -\frac{1}{4} - \frac{1}{4} = -\frac{1}{2}, \quad |J| = \frac{1}{2}.$$

Bölge: $0 < \frac{u+v}{2} < 1$ ve $0 < \frac{u-v}{2} < 1$, yani $-u < v < 2-u$ ve $u-2 < v < u$; birleştirince $0 < u < 2$ ve $|v| < \min(u, 2-u)$. Bu, köşeleri $(0, 0)$, $(1, 1)$, $(2, 0)$, $(1, -1)$ olan bir karedir (alanı 2). Ortak yoğunluk

$$g(u, v) = \begin{cases} \frac{1}{2}, & 0 < u < 2, |v| < \min(u, 2-u) \\ 0, & \text{diğer durumlarda.} \end{cases}$$

V 'nin marjinali: sabit $v \in (-1, 1)$ için u , $|v| < u < 2 - |v|$ aralığında dolaşır;

$$f_V(v) = \int_{|v|}^{2-|v|} \frac{1}{2} du = \frac{1}{2} (2 - 2|v|) = 1 - |v|, \quad -1 < v < 1.$$

Bu, $(-1, 1)$ üzerinde ortalanmış üçgen yoğunluktur; $\int_{-1}^1 (1 - |v|) dv = 1$. (U 'nun marjinali de aynı hesapla $1 - |u - 1|$, yani Örnek 15.3'deki üçgen yoğunluktur.)

f) Ters dönüşüm $x = r \cos \theta$, $y = r \sin \theta$;

$$J = \det \begin{bmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{bmatrix} = r \cos^2 \theta + r \sin^2 \theta = r, \quad |J| = r.$$

$(0, 1) \times (0, 2\pi)$ dikdörtgeni, $\theta = 0$ ışını dışında (olasılığı sıfır) birim diske birebir gider. Ortak yoğunluk

$$g(r, \theta) = \frac{1}{\pi} \cdot r = \frac{r}{\pi}, \quad 0 < r < 1, 0 < \theta < 2\pi.$$

Marjinaller: $f_R(r) = \int_0^{2\pi} \frac{r}{\pi} d\theta = 2r$ ($0 < r < 1$) ve $f_\Theta(\theta) = \int_0^1 \frac{r}{\pi} dr = \frac{1}{2\pi}$ ($0 < \theta < 2\pi$); bu aralıkların dışında marjinaller sıfırdır. Her (r, θ) için $g(r, \theta) = f_R(r) f_\Theta(\theta)$ sağlanır: dikdörtgenin içinde $\frac{r}{\pi} = (2r) \cdot \frac{1}{2\pi}$, dışında iki taraf da 0'dır. Öyleyse R ile Θ bağımsızdır ([Tanım 14.1](#)). Yani açı $(0, 2\pi)$ üzerinde düzgün, yarıçap ise $2r$ yoğunluğuna sahiptir; $\int_0^1 2r dr = 1$. Yarıçapın düzgün olmaması doğaldır: diskin dış halkalarında daha fazla alan vardır.

■

Rastgele vektörlerin dönüşümleriyle rastgele vektörler kısmını tamamlamış olduk. Bundan sonra dağılımları tek tek betimlemek yerine onları birkaç sayıyla özetlemeyi öğreneceğiz: [Beklenen Değer, Momentler ve Varyans](#).

16. Beklenen Değer, Momentler ve Varyans

Rastgele Vektörlerin Dönüşümleri bölümüyle birlikte bir rastgele değişkenin ya da rastgele vektörün dağılımını betimlemenin bütün araçlarını edindik: dağılım fonksiyonu, olasılık fonksiyonu, yoğunluk fonksiyonu ve bunların bir dönüşüm altında nasıl değiştiği. Ancak bir dağılımın tamamı çoğu zaman gereğinden fazla bilgidir. Bir zarın “ortalama” kaç geldiğini, bir makine parçasının ömrünün ortalama ne kadar sürdüğünü, ölçümlerin bu ortalamadan ne kadar saptığını sormak isteriz. Bu bölümde bir dağılımı birkaç sayıya sıkıştıran ilk ve en önemli araçları tanıtıyoruz: **beklenen değer**, **momentler** ve **varyans**. Fikir basittir. Bir zar N kez atılsın ve $1, 2, \dots, 6$ değerlerinin sırasıyla $n_1, \dots, n_6$ kez geldiğini varsayalım. Gelen sayıların aritmetik ortalaması

$$\frac{1 \cdot n_1 + 2 \cdot n_2 + \dots + 6 \cdot n_6}{N} = 1 \cdot \frac{n_1}{N} + 2 \cdot \frac{n_2}{N} + \dots + 6 \cdot \frac{n_6}{N}$$

biçiminde yazılabilir; yani her değer, gözlenme sıklığıyla ağırlıklandırılarak toplanmıştır. Atış sayısı büyüdükçe sıklıkların olasılıklara yaklaşmasını bekleriz. Öyleyse “uzun vadedeki ortalama” için doğal aday, her değer kendi olasılığıyla ağırlıklandırılmış toplamı $\sum_{x \in D_X} x f_X(x)$ olur. Beklenen değer tam olarak budur; sürekli durumda toplamın yerini integral alır.

Bu bölümde şunları öğreneceğiz: bir rastgele değişkenin fonksiyonunun beklenen değerinin tanımı ve bu tanımın anlamlı olması için gereken mutlak yakınsaklık koşulu; beklenen değeri olmayan rastgele değişkenlerin gerçekten var olduğu; momentler, varyans ve standart sapma; beklenen değer en çok kullanılan özelliği olan lineerlik; doğrusal bir dönüşümün beklenen değeri ve varyansı nasıl değiştirdiği; varyansı hesaplamının pratik formülü; ve varyansın ne ölçtüğünü kesin bir eşitsizliğe dönüştüren Markov ve Chebyshev eşitsizlikleri.

16.1 Bir Rastgele Değişkenin Fonksiyonunun Beklenen Değeri

Beklenen değeri yalnız X için değil, X 'in bir fonksiyonu $g(X)$ için tanımlamak daha kullanışlıdır; çünkü X^2 , $|X - c|$, e^{tX} gibi nesnelerin beklenen değerleri ilerleyen bölümlerde sürekli karşımıza çıkacak. $g(X)$ 'in bir rastgele değişken olabilmesi için her Borel kümesinin g altındaki ters görüntüsünün yine bir Borel kümesi olması gerekir; bu koşulu sağlayan fonksiyonlara **Borel ölçülebilir** fonksiyon denir. Sürekli fonksiyonlar, parçalı sürekli fonksiyonlar ve bunların bileşkeleri, toplamaları, çarpımları hep Borel ölçülebilirdir; bu kitapta karşılaşacağımız her g bu sınıftadır.

Tanım 16.1 (Beklenen Değer). X bir rastgele değişken ve $g : \mathbb{R} \rightarrow \mathbb{R}$, her $B \in \mathcal{B}(\mathbb{R})$ için $\{x : g(x) \in B\} \in \mathcal{B}(\mathbb{R})$ özelliğini taşıyan bir fonksiyon olsun.

(i) X kesikli bir rastgele değişken ve

$$\sum_{x \in D_X} |g(x)| f_X(x) < \infty$$

ise

$$E[g(X)] = \sum_{x \in D_X} g(x) f_X(x),$$

(ii) X sürekli bir rastgele değişken ve

$$\int_{-\infty}^{\infty} |g(x)| f_X(x) dx < \infty$$

ise

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$$

sayısına $g(X)$ rastgele değişkeninin **beklenen değeri** denir. Yakınsaklık koşulu sağlanmıyorsa $g(X)$ 'in beklenen değeri **yoktur** denir.

Tanım üzerine birkaç açıklama gerekir.

- $g(x) = x$ seçimi $E(X)$ 'i verir: kesikli durumda $E(X) = \sum_{x \in D_X} x f_X(x)$, sürekli durumda $E(X) = \int_{-\infty}^{\infty} x f_X(x) dx$. Girişteki “olasılıkla ağırlıklandırılmış ortalama” tam olarak budur.
- Kesikli durumda D_X sonlu ise koşul kendiliğinden sağlanır; toplam sonlu terimlidir. D_X sayılabilir sonsuz ise $\sum_{x \in D_X} g(x) f_X(x)$ bir seridir ve D_X 'in elemanlarını numaralandırmanın ayrıcalıklı bir yolu yoktur. Serinin değeri sıralamaya bağlı olmasın diye **mutlak yakınsaklık** istenir; mutlak yakınsak bir serinin terimleri hangi sırayla toplanırsa toplansın aynı sayı çıkar.
- Sürekli durumda integral yalnız D_X üzerinden, yani $\int_{D_X} g(x) f_X(x) dx$ biçiminde de yazılabilir; D_X dışında f_X sıfır olduğundan ikisi aynıdır. Mutlak yakınsaklık koşulu burada da integralin “ $\infty - \infty$ ” belirsizliğine düşmemesini sağlar.
- Yakınsaklık koşulu sağlandığında, üçgen eşitsizliği gereği $|E[g(X)]| \leq E[|g(X)|]$ olur; yani beklenen değer sonlu bir gerçel sayıdır.

i Tanım neden X 'in dağılımıyla yazılıyor?

$Y = g(X)$ kendisi bir rastgele değişkendir (Tanım 8.1). Öyleyse beklenen değerini, tanımdaki $g(y) = y$ seçimiyle, Y 'nin kendi olasılık fonksiyonu f_Y üzerinden de yazabilirdik: $E(Y) = \sum_{y \in D_Y} y f_Y(y)$. İki hesabın aynı sayıyı verdiğini kesikli durumda gösterelim. Tanım 10.2 gereği

$$f_Y(y) = P(Y = y) = P(g(X) = y) = \sum_{x \in D_X: g(x)=y} f_X(x)$$

olduğundan

$$\sum_{y \in D_Y} y f_Y(y) = \sum_{y \in D_Y} \sum_{x \in D_X: g(x)=y} y f_X(x) = \sum_{y \in D_Y} \sum_{x \in D_X: g(x)=y} g(x) f_X(x) = \sum_{x \in D_X} g(x) f_X(x)$$

bulunur. Terimler mutlak toplanabilir olduğundan, yani $\sum_{x \in D_X} |g(x)| f_X(x) < \infty$ olduğundan, iki katlı toplam istenen sırada düzenlenebilir; son adımda D_X , g 'nin aldığı değerlere göre ayrık parçalara bölünmüştür. Demek ki $E[g(X)]$ 'i hesaplamak için $g(X)$ 'in dağılımını bulmaya gerek yoktur; X 'in dağılımı yeter. Tanımın asıl gücü budur: X^2 'nin ya da e^{tX} 'in dağılımını hiç bilmeden beklenen değerleri yazılabilir. Sürekli durumda da aynı sonuç geçerlidir; ispatı dönüşüm formülünün genel biçimini gerektirdiğinden burada verilmeyecektir.

i Varlık anlaşıması

Bundan sonra, aksi söylenmedikçe, göz önüne alınan $g(X)$ rastgele değişkenlerinin beklenen değerlerinin **var olduğu varsayılacaktır**. Bir teoremden “ $E(X)$ ve $E(X^2)$ var olsun” gibi bir hipotez yazılmamışsa, ifadede geçen bütün beklenen değerlerin var olduğu anlaşılmalıdır.

Beklenen değerin var olup olmadığına her seferinde tanımdan bakmak zahmetlidir. Aşağıdaki gözlemler, çoğu durumda varlığı bir bakışta görmeyi sağlar.

Önerme 16.1 (Beklenen Değerin Varlığı İçin Yeter Koşullar). X bir rastgele değişken, g ve h Borel ölçülebilir fonksiyonlar olsun.

(a) g sınırlı ise, yani her x için $|g(x)| \leq M$ olacak biçimde bir M varsa, $E[g(X)]$ vardır ve $|E[g(X)]| \leq M$ 'dir.

(b) Her $x \in D_X$ için $|g(x)| \leq |h(x)|$ ve $E[h(X)]$ varsa $E[g(X)]$ de vardır.

(c) Bir $k \in \mathbb{N}$ için $E(X^k)$ varsa, her $j \in \{1, 2, \dots, k\}$ için $E(X^j)$ vardır.

İspat.

İspatı kesikli durum için yazıyoruz; sürekli durumda toplamaların yerine integraller gelir, hiçbir adım değişmez.

- (a) $\sum_{x \in D_X} |g(x)| f_X(x) \leq M \sum_{x \in D_X} f_X(x) = M < \infty$. Böylece $E[g(X)]$ vardır ve $|E[g(X)]| \leq \sum_{x \in D_X} |g(x)| f_X(x) \leq M$ olur.
- (b) $\sum_{x \in D_X} |g(x)| f_X(x) \leq \sum_{x \in D_X} |h(x)| f_X(x) < \infty$.
- (c) Her gerçel x ve her $j \leq k$ için $|x|^j \leq 1 + |x|^k$ eşitsizliği geçerlidir: $|x| \leq 1$ ise $|x|^j \leq 1$, $|x| > 1$ ise $|x|^j \leq |x|^k$ olur. Öyleyse

$$\sum_{x \in D_X} |x|^j f_X(x) \leq \sum_{x \in D_X} f_X(x) + \sum_{x \in D_X} |x|^k f_X(x) = 1 + \sum_{x \in D_X} |x|^k f_X(x) < \infty$$

olur; yani $E(X^j)$ vardır.

■

(c) maddesinin karşıt tersi de düşünmeye değer: bir moment yoksa, ondan daha yüksek mertebeden hiçbir moment yoktur. Aşağıdaki örnek en alt basamağın bile boş kalabileceğini gösterir.

Örnek 16.1 (Beklenen Değeri Olmayan Bir Rastgele Değişken). X sürekli rastgele değişkeninin yoğunluk fonksiyonu

$$f_X(x) = \begin{cases} \frac{1}{x^2}, & x > 1 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun. f_X 'in gerçekten bir yoğunluk fonksiyonu olduğunu, ama $E(X)$ 'in var olmadığını gösteriniz.

Çözüm.

$f_X \geq 0$ açıktır ve

$$\int_{-\infty}^{\infty} f_X(x) dx = \int_1^{\infty} \frac{dx}{x^2} = \lim_{b \rightarrow \infty} \left[-\frac{1}{x} \right]_1^b = \lim_{b \rightarrow \infty} \left(1 - \frac{1}{b} \right) = 1$$

olduğundan f_X bir yoğunluk fonksiyonudur (Tanım 11.2). Beklenen değer için tanımdaki koşula bakalım; $x > 1$ için $|x| = x$ olduğundan

$$\int_{-\infty}^{\infty} |x| f_X(x) dx = \int_1^{\infty} x \cdot \frac{1}{x^2} dx = \int_1^{\infty} \frac{dx}{x} = \lim_{b \rightarrow \infty} \ln b = \infty.$$

Koşul sağlanmaz; $E(X)$ yoktur. **Önerme 16.1**'in (c) maddesi gereği X 'in hiçbir $k \geq 1$ mertebeli momenti yoktur.

Bunun anlamı şudur: X 'in dağılımı kusursuz biçimde tanımlıdır, her $a > 1$ için $P(X > a) = \int_a^{\infty} x^{-2} dx = \frac{1}{a}$ hesaplanabilir; ama bu kuyruk o kadar yavaş söner ki “ortalama değer” sonlu bir sayı olarak var olmaz. X 'in bağımsız gözlemlerinin aritmetik ortalaması, gözlem sayısı arttıkça bir sayıya yerleşmez; sınırsız büyür.

■

⚠ Simetri tek başına beklenen değeri sıfır yapmaz

$f(x) = \frac{1}{\pi(1+x^2)}$, $x \in \mathbb{R}$, bir yoğunluk fonksiyonudur (Cauchy yoğunluğu): $\int_{-\infty}^{\infty} f(x) dx = \frac{1}{\pi} [\arctan x]_{-\infty}^{\infty} = \frac{1}{\pi} \left(\frac{\pi}{2} + \frac{\pi}{2} \right) = 1$. Grafiği 0'a göre simetriktir ve $xf(x)$ tek fonksiyondur; bu yüzden $E(X) = 0$ demek çekicidir. Ama

$$\int_{-\infty}^{\infty} |x| f(x) dx = \frac{2}{\pi} \int_0^{\infty} \frac{x}{1+x^2} dx = \frac{1}{\pi} \lim_{b \rightarrow \infty} \ln(1+b^2) = \infty$$

olduğundan $E(X)$ **yoktur**. Simetriden $E(X) = 0$ sonucunu çıkarmak ancak mutlak yakınsaklık koşulu sağlandığında geçerlidir: o zaman integral iki yarımın toplamı olarak parçalanabilir ve yarım-lar birbirini götürür. Koşul sağlanmıyorsa integral “ $\infty - \infty$ ” biçimindedir ve hiçbir değer taşımaz.

16.2 Momentler, Varyans ve Standart Sapma

g fonksiyonunun belirli seçimleri o kadar sık kullanılır ki özel adlar almıştır.

Tanım 16.2 (Momentler). X bir rastgele değişken, $c \in \mathbb{R}$ ve $k \in \mathbb{N}$ olsun.

(a) $E[(X - c)^k]$ değerine X 'in c 'ye göre k . **mertebeden momenti**,

(b) $E[X^k]$ değerine X 'in k . **mertebeden momenti** (sıfıra göre k . moment),

(c) $\mu_X = \mu = E(X)$ değerine X 'in **beklenen değeri** (birinci moment),

(d) $E[X(X-1)(X-2)\cdots(X-k+1)]$ değerine X 'in k . **çarpımsal momenti**

denir. Beklenen değere göre momentlere, yani $E[(X - \mu)^k]$ değerlerine, **merkezi momentler** de denir.

Merkezi momentlerin ilki her zaman sıfırdır: birazdan ispatlayacağımız lineerlik gereği $E(X - \mu) = E(X) - \mu = 0$. Bilgi taşıyan ilk merkezi moment ikincisidir ve kendi adı vardır.

Tanım 16.3 (Varyans). X , beklenen değeri $\mu = E(X)$ olan bir rastgele değişken olsun. X 'in beklenen değerine göre ikinci momentine, yani

$$\sigma_X^2 = \sigma^2 = \text{Var}(X) = E[(X - E(X))^2]$$

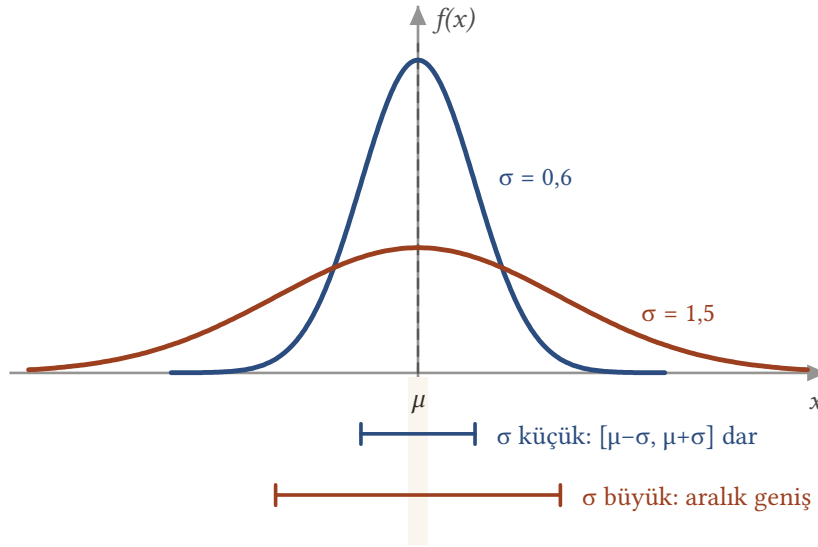
değerine X 'in **varyansı** denir.

Tanım 16.4 (Standart Sapma). X 'in varyansının negatif olmayan kareköküne X 'in **standart sapması** denir ve

$$\sigma_X = \sqrt{\text{Var}(X)}$$

ile gösterilir.

Bu iki büyüklüğün yorumu şöyledir. Kesikli bir dağılımı, sayı doğrusu üzerinde $x \in D_X$ noktalarına yerleştirilmiş $f_X(x)$ ağırlıklarında kütleler gibi düşünelim; toplam kütle 1'dir. Bu kütle sisteminin ağırlık merkezi $\frac{\sum_{x \in D_X} x f_X(x)}{\sum_{x \in D_X} f_X(x)} = E(X)$ noktasıdır. Yani **beklenen değer, dağılımın “merkezi”, dağılımın dengede durduğu noktadır**. Varyans ise kütlelerin bu merkezden uzaklıklarının karelerinin ağırlıklı ortalamasıdır, mekanikteki eylemsizlik momentinin karşılığıdır: kütleler merkeze yığılmışsa küçük, merkezden uzağa saçılmışsa büyüktür. **Varyans, dağılımın merkez etrafındaki “yayılımının” bir ölçüsüdür**. Varyansın birimi X 'in biriminin karesidir; X santimetreyle ölçülüyorsa $\text{Var}(X)$ santimetrekaredir. Standart sapma karekök alarak X ile aynı birime döner; bu yüzden yayılımı somut olarak ifade etmek istendiğinde σ_X kullanılır.



Şekil 16.1: Varyans, dağılımın ortalama etrafındaki yayılımını ölçer. İki yoğunluğun beklenen değeri aynı μ noktasıdır; dar eğrinin standart sapması $\sigma = 0,6$, geniş eğrininki $\sigma = 1,5$ 'tir. Eksenin altındaki parantezler $[\mu - \sigma, \mu + \sigma]$ aralıklarını gösterir: kütle merkeze yığılmışsa varyans küçük, merkezden uzağa saçılmışsa büyüktür.

Bir dağılımın momentleri hakkında ilk soru, hangilerinin var olduğudur. **Önerme 16.1'in (c)** maddesi bunun düzenli bir yapı taşıdığını söyler: $E(X^k)$ varsa daha düşük mertebeli bütün momentler de vardır. Özellikle $E(X^2)$ varsa $E(X)$ de vardır ve birazdan göreceğimiz gibi varyans da var olur.

16.3 Beklenen Değerin Lineerliği

Beklenen değerin hesaplarda en çok kullanılan özelliği, toplamın beklenen değerinin beklenen değerlerin toplamı olmasıdır. Bu özellik olmadan $E[(X - \mu)^2]$ gibi ifadeleri açmak bile mümkün olmazdı. Aşağıdaki teoremden $g_1, \dots, g_n$ fonksiyonlarının toplamı da Borel ölçülebilirdir; bu, ölçülebilir fonksiyonların sonlu toplamlarının ölçülebilir olduğu genel gerçeğinin bir sonucudur ve burada ispatsız kullanılacaktır.

Teorem 16.1 (Beklenen Değerin Lineerliği). X bir rastgele değişken ve $g_i : \mathbb{R} \rightarrow \mathbb{R}, i = 1, 2, \dots, n$, her $B \in \mathcal{B}(\mathbb{R})$ için $\{x : g_i(x) \in B\} \in \mathcal{B}(\mathbb{R})$ özelliğini taşıyan fonksiyonlar olsun. Her i için $E[g_i(X)]$ beklenen değeri varsa $\sum_{i=1}^n g_i(X)$ rastgele değişkeninin beklenen değeri vardır ve

$$E\left[\sum_{i=1}^n g_i(X)\right] = \sum_{i=1}^n E[g_i(X)]$$

olur.

İspat.

$g = g_1 + g_2 + \dots + g_n$ yazalım.

X kesikli ise. Önce varlık: her x için üçgen eşitsizliği $|g(x)| \leq \sum_{i=1}^n |g_i(x)|$ verir. Buradan

$$\sum_{x \in D_X} |g(x)| f_X(x) \leq \sum_{x \in D_X} \sum_{i=1}^n |g_i(x)| f_X(x) = \sum_{i=1}^n \sum_{x \in D_X} |g_i(x)| f_X(x) < \infty$$

olur; sonlu sayıda yakınsak serinin toplamı yakınsak olduğundan sonlu toplam ile seri yer değiştirebilir. Demek ki $E[g(X)]$ vardır. Şimdi beklenen değerin kendisini hesaplayalım:

$$E[g(X)] = \sum_{x \in D_X} \left(\sum_{i=1}^n g_i(x) \right) f_X(x) = \sum_{i=1}^n \sum_{x \in D_X} g_i(x) f_X(x) = \sum_{i=1}^n E[g_i(X)].$$

Yer değiştirme yine geçerlidir, çünkü içindeki n serinin her biri mutlak yakınsaktır.
 X sürekli ise. Aynı adımlar integralle yazılır:

$$\int_{-\infty}^{\infty} |g(x)| f_X(x) dx \leq \sum_{i=1}^n \int_{-\infty}^{\infty} |g_i(x)| f_X(x) dx < \infty$$

ve

$$E[g(X)] = \int_{-\infty}^{\infty} \sum_{i=1}^n g_i(x) f_X(x) dx = \sum_{i=1}^n \int_{-\infty}^{\infty} g_i(x) f_X(x) dx = \sum_{i=1}^n E[g_i(X)].$$

Burada da sonlu sayıda yakınsak integralin toplamı terim terim alınabilir.

■

Lineerliğin öteki yarısı, sabitlerin dışarı alınabilmesidir.

Önerme 16.2 (Sabitin Beklenen Değeri ve Sabitle Çarpım). X bir rastgele değişken, $c \in \mathbb{R}$ ve g Borel ölçülebilir bir fonksiyon olsun.

(a) $E(c) = c$; yani sabit c değerini alan rastgele değişkenin beklenen değeri c 'dir.

(b) $E[cg(X)]$ varsa $E[cg(X)]$ de vardır ve $E[cg(X)] = cE[g(X)]$ olur.

İspat.

(a) $g(x) = c$ sabit fonksiyonu Borel ölçülebilirdir: $\{x : c \in B\}$ kümesi $c \in B$ ise $\mathbb{R}$, değilse $\emptyset$ 'dir ve ikisi de Borel kümesidir. Kesikli durumda $\sum_{x \in D_X} |c| f_X(x) = |c| < \infty$ olduğundan beklenen değer vardır ve

$$E(c) = \sum_{x \in D_X} c f_X(x) = c \sum_{x \in D_X} f_X(x) = c \cdot 1 = c.$$

Sürekli durumda $E(c) = \int_{-\infty}^{\infty} c f_X(x) dx = c \int_{-\infty}^{\infty} f_X(x) dx = c$.

(b) $\sum_{x \in D_X} |cg(x)| f_X(x) = |c| \sum_{x \in D_X} |g(x)| f_X(x) < \infty$ olduğundan $E[cg(X)]$ vardır ve

$$E[cg(X)] = \sum_{x \in D_X} cg(x) f_X(x) = c \sum_{x \in D_X} g(x) f_X(x) = cE[g(X)].$$

Sürekli durum, toplam yerine integrale aynıdır.

■

Teorem 16.1 ile **Önerme 16.2** birleşince, $c_1, \dots, c_n$ sabitleri için

$$E\left[\sum_{i=1}^n c_i g_i(X)\right] = \sum_{i=1}^n c_i E[g_i(X)]$$

elde edilir. Bu yüzden **beklenen değer, ya da E operatörü, lineer bir dönüşümdür.** Bundan sonraki hesapların çoğu bu tek cümleye dayanır.

Lineerliğin en basit ve en sık kullanılan hâli, X 'in yerine $aX + b$ konulmasıdır. Bir sıcaklık ölçümünü Celsius'tan Fahrenheit'a çevirmek, bir fiyata sabit bir vergi eklemek, bir ölçümü başka birime dönüştürmek hep bu türdendir. Beklenen değer ve varyans bu dönüşüm altında öngörülebilir biçimde değişir.

Teorem 16.2 (Doğrusal Dönüşümün Beklenen Değeri ve Varyansı). X bir rastgele değişken ve $a, b \in \mathbb{R}$ olsun.

(a) $E(X)$ varsa $E(aX + b)$ de vardır ve $E(aX + b) = aE(X) + b$ olur.

(b) $\text{Var}(X)$ varsa $\text{Var}(aX + b)$ de vardır ve $\text{Var}(aX + b) = a^2 \text{Var}(X)$ olur.

İspat.

(a) Önce varlık. $|ax + b| \leq |a||x| + |b|$ olduğundan kesikli durumda

$$\sum_{x \in D_X} |ax + b| f_X(x) \leq |a| \sum_{x \in D_X} |x| f_X(x) + |b| \sum_{x \in D_X} f_X(x) = |a| \sum_{x \in D_X} |x| f_X(x) + |b| < \infty$$

olur; sağdaki seri $E(X)$ 'in varlık koşulundan yakınsaktır. Sürekli durumda aynı sınır integralle yazılır. Demek ki $E(aX + b)$ vardır. Değeri doğrudan hesaplayalım. X kesikli ise

$$E(aX + b) = \sum_{x \in D_X} (ax + b) f_X(x) = a \underbrace{\sum_{x \in D_X} x f_X(x)}_{=E(X)} + b \underbrace{\sum_{x \in D_X} f_X(x)}_{=1} = a E(X) + b.$$

X sürekli ise

$$E(aX + b) = \int_{-\infty}^{\infty} (ax + b) f_X(x) dx = a \int_{-\infty}^{\infty} x f_X(x) dx + b \int_{-\infty}^{\infty} f_X(x) dx = a E(X) + b.$$

(b) $\mu = E(X)$ yazalım. **(a)** gereği $E(aX + b) = a\mu + b$ olduğundan, varyansın tanımına göre

$$\text{Var} (aX + b) = E \left[(aX + b - E(aX + b))^2 \right] = E \left[(aX + b - a\mu - b)^2 \right] = E \left[a^2 (X - \mu)^2 \right].$$

$E[(X - \mu)^2] = \text{Var} (X)$ var olduğundan **Önerme 16.2**'in **(b)** maddesi sabit a^2 'yi dışarı alır:

$$\text{Var} (aX + b) = a^2 E \left[(X - \mu)^2 \right] = a^2 \text{Var} (X).$$

■

Sonucun iki yanı ayrı ayrı yorumlanmalıdır. b sabitini eklemek dağılımı bütünüyle b kadar kaydırır: merkez b kadar kayar, ama yayılım değişmez; $\text{Var} (X + b) = \text{Var} (X)$. a ile çarpmak ise dağılımı $|a|$ kat gerer ya da sıkıştırır; uzaklıklar $|a|$ katına, uzaklıkların kareleri a^2 katına çıkar. Standart sapma cinsinden yazılırsa $\sigma_{aX+b} = |a| \sigma_X$ olur. Bu gözlemin en yararlı uygulaması **standartlaştırmadır**: $\sigma_X > 0$ ise

$$Z = \frac{X - \mu_X}{\sigma_X} = \frac{1}{\sigma_X} X - \frac{\mu_X}{\sigma_X}$$

rastgele değişkeni için $E(Z) = \frac{1}{\sigma_X} \mu_X - \frac{\mu_X}{\sigma_X} = 0$ ve $\text{Var} (Z) = \frac{1}{\sigma_X^2} \text{Var} (X) = 1$ olur. Yani her rastgele değişken, doğrusal bir dönüşümle beklenen değeri 0 ve varyansı 1 olan bir değişkene çevrilebilir.

16.4 Varyansın Hesaplanması

Varyansı tanımından, yani $\sum_{x \in D_X} (x - \mu)^2 f_X(x)$ toplamından hesaplamak, önce μ 'yü bulmayı, sonra her terimde bir fark ve bir kare almayı gerektirir. Aşağıdaki formül bu işi $E(X)$ ve $E(X^2)$ gibi iki bağımsız hesaba indirir; uygulamada varyans hemen her zaman bu yolla bulunur.

Önerme 16.3 (Varyans Hesap Formülü). X , ikinci momenti $E(X^2)$ var olan bir rastgele değişken ve $\mu = E(X)$ olsun.

(a) $\text{Var} (X)$ vardır ve

$$\text{Var} (X) = E (X^2) - (E(X))^2$$

olur.

(b) Her $c \in \mathbb{R}$ için

$$E [(X - c)^2] = \text{Var} (X) + (\mu - c)^2$$

olur. Dolayısıyla c 'ye göre ikinci moment en küçük değerini yalnız $c = \mu$ 'de alır ve bu en küçük değer $\text{Var} (X)$ 'tir.

İspat.

(a) $E(X^2)$ var olduğundan **Önerme 16.1**'in **(c)** maddesi gereği $E(X)$ de vardır; μ sonlu bir sayıdır. $(X - \mu)^2 = X^2 - 2\mu X + \mu^2$ açılımındaki üç terimin de beklenen değeri vardır, öyleyse **Teorem 16.1** ve **Önerme 16.2** gereği $E[(X - \mu)^2]$ vardır ve

$$\text{Var} (X) = E [X^2 - 2\mu X + \mu^2] = E (X^2) - 2\mu E(X) + \mu^2 = E (X^2) - 2\mu^2 + \mu^2 = E (X^2) - \mu^2.$$

Kesikli durumda bu hesap açıkça

$$\sum_{x \in D_X} (x - \mu)^2 f_X(x) = \sum_{x \in D_X} x^2 f_X(x) - 2\mu \sum_{x \in D_X} x f_X(x) + \mu^2 \sum_{x \in D_X} f_X(x) = E(X^2) - 2\mu \cdot \mu + \mu^2 = E(X^2) - \mu^2$$

biçimindedir; sürekli durumda toplamaların yerini integraller alır.

(b) $X - c = (X - \mu) + (\mu - c)$ yazıp kare alalım:

$$(X - c)^2 = (X - \mu)^2 + 2(\mu - c)(X - \mu) + (\mu - c)^2.$$

Beklenen değer alındığında lineerlik gereği

$$E[(X - c)^2] = E[(X - \mu)^2] + 2(\mu - c)E(X - \mu) + (\mu - c)^2$$

olur. $E(X - \mu) = E(X) - \mu = 0$ olduğundan ortadaki terim düşer ve $E[(X - c)^2] = \text{Var}(X) + (\mu - c)^2$ kalır. $(\mu - c)^2 \geq 0$ ve yalnız $c = \mu$ 'de sıfır olduğundan en küçük değer $c = \mu$ 'de alınır.

■

(b) maddesi beklenen değer “merkez” yorumunu kesinleştirir: X 'i tek bir c sayısı ile temsil etmek isteyip hatayı $(X - c)^2$ 'nin beklenen değeriyle ölçersek, en iyi seçim $c = \mu$ 'dür ve o zaman kalan hata tam olarak varyanstır.

Bir başka sonuç da varlıkla ilgilidir: $E(X^2)$ varsa varyans vardır. Terside doğrudur; μ ve $\text{Var}(X)$ varsa $X^2 = (X - \mu)^2 + 2\mu X - \mu^2$ eşitliği $E(X^2)$ 'nin varlığını verir. Yani **varyansın var olması ile ikinci momentin var olması aynı şeydir**.

Varyans, karelerin ağırlıklı ortalaması olduğundan negatif olamaz. Bu basit gözlemin hesap formülüyle birleşmesi, sık kullanılan bir eşitsizlik verir.

Önerme 16.4 (Varyans Negatif Olamaz). X , ikinci momenti var olan bir rastgele değişken olsun.

(a) $\text{Var}(X) \geq 0$.

(b) $E(X^2) \geq (E(X))^2$.

(c) X kesikli ise, $\text{Var}(X) = 0$ olması için gerek ve yeter koşul $P(X = E(X)) = 1$ olmasıdır.

İspat.

(a) $\mu = E(X)$ olsun. Kesikli durumda $\text{Var}(X) = \sum_{x \in D_X} (x - \mu)^2 f_X(x)$ toplamının her terimi, negatif olmayan $(x - \mu)^2$ ile negatif olmayan $f_X(x)$ 'in çarpımıdır; öyleyse toplam ≥ 0 'dır. Sürekli durumda $\int_{-\infty}^{\infty} (x - \mu)^2 f_X(x) dx$ integralinin integrandı her yerde ≥ 0 olduğundan integral ≥ 0 'dır.

(b) **Önerme 16.3** gereği $0 \leq \text{Var}(X) = E(X^2) - (E(X))^2$; taraflar düzenlenince $E(X^2) \geq (E(X))^2$ çıkar.

(c) X kesikli olsun ve D_X , olasılığı pozitif olan değerlerden oluşsun. $\text{Var}(X) = 0$ ise, negatif olmayan terimlerden oluşan $\sum_{x \in D_X} (x - \mu)^2 f_X(x)$ toplamı sıfırdır; bu ancak her terim sıfırda olur. $f_X(x) > 0$ olduğundan her $x \in D_X$ için $(x - \mu)^2 = 0$, yani $x = \mu$ olmalıdır. Demek ki $D_X = \{\mu\}$ ve $P(X = \mu) = f_X(\mu) = 1$. Tersine $P(X = \mu) = 1$ ise $D_X = \{\mu\}$ ve $\text{Var}(X) = (\mu - \mu)^2 \cdot 1 = 0$ olur.

■

(c) maddesi varyansın yayılım ölçüsü olduğunu en uç durumda doğrular: varyansın sıfır olması, rastgele değişkenin aslında rastgele olmaması, tek bir değere çakılmış olması demektir. Sürekli bir rastgele değişkenin varyansı hiçbir zaman sıfır olamaz; bunu Chebyshev eşitsizliğinden sonra göstereceğiz.

16.5 Örnekler

Şimdi tanımları ve formülleri somut dağılımlar üzerinde işletelim. Her örnekte aynı üç adım tekrarlanır: $E(X)$, $E(X^2)$ ve $\text{Var}(X) = E(X^2) - (E(X))^2$.

Örnek 16.2 (Kesikli Düzgün Dağılımın Momentleri). X kesikli rastgele değişkeninin olasılık fonksiyonu

$$f_X(x) = \frac{1}{6}, \quad x \in \{-2, -1, 0, 1, 2, 3\}$$

olsun. $E(X)$, $E(X^2)$, $\text{Var}(X)$ ve σ_X değerlerini bulunuz.

Çözüm.

D_X sonlu olduğundan bütün momentler vardır. Beklenen değer:

$$E(X) = \sum_{x=-2}^3 x f_X(x) = (-2) \cdot \frac{1}{6} + (-1) \cdot \frac{1}{6} + 0 \cdot \frac{1}{6} + 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} = \frac{-2 - 1 + 0 + 1 + 2 + 3}{6} = \frac{3}{6} = \frac{1}{2}.$$

İkinci moment:

$$E(X^2) = \sum_{x=-2}^3 x^2 f_X(x) = \frac{(-2)^2 + (-1)^2 + 0^2 + 1^2 + 2^2 + 3^2}{6} = \frac{4 + 1 + 0 + 1 + 4 + 9}{6} = \frac{19}{6}.$$

Varyans, **Önerme 16.3** ile:

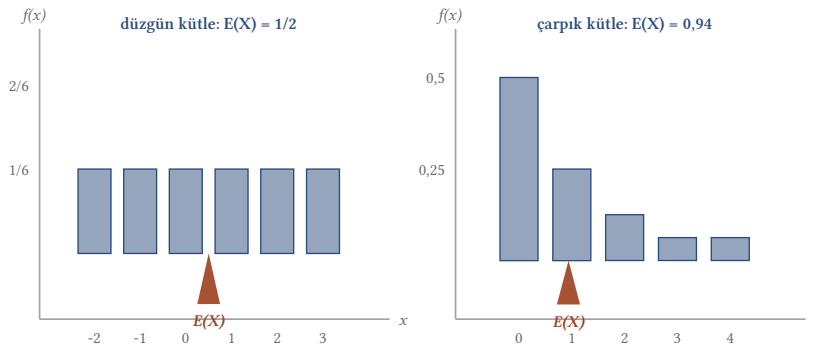
$$\text{Var}(X) = E(X^2) - (E(X))^2 = \frac{19}{6} - \left(\frac{1}{2}\right)^2 = \frac{38}{12} - \frac{3}{12} = \frac{35}{12}.$$

Aynı sonucu tanımdan da doğrulayalım. $\mu = \frac{1}{2}$ için $x - \mu$ farkları $-\frac{5}{2}, -\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}, \frac{5}{2}$; kareleri $\frac{25}{4}, \frac{9}{4}, \frac{1}{4}, \frac{1}{4}, \frac{9}{4}, \frac{25}{4}$ tür ve

$$\text{Var}(X) = \frac{1}{6} \cdot \frac{25 + 9 + 1 + 1 + 9 + 25}{4} = \frac{1}{6} \cdot \frac{70}{4} = \frac{35}{12}.$$

Standart sapma $\sigma_X = \sqrt{\frac{35}{12}} \approx 1,71$ 'dir.

Beklenen değer $\frac{1}{2}$ çıkması şaşırtıcı değildir: $\frac{1}{2}$, $D_X = \{-2, \dots, 3\}$ kümesinin tam ortasıdır; altı eşit kütle -2 ile 3 arasına eşit aralıklarla dizilmiştir ve sistem bu iki ucun ortasında dengede durur. $\frac{1}{2}$ 'nin X 'in alabileceği bir değer olmadığına dikkat ediniz; beklenen değer bir "olası sonuç" değil, dağılımın ağırlık merkezidir.



Şekil 16.2: Beklenen değer, olasılık kütlelerinin denge noktasıdır: eksen bir çubuk, olasılıklar üzerine konmuş ağırlıklar gibi düşünülürse çubuk tam $E(X)$ noktasında dengede durur. Solda simetrik kütle ortada dengelenir; sağda kütle sola yığıldığından denge noktası sola kayar ve alınan değerlerden hiçbirine eşit olmak zorunda değildir.

■

Örnek 16.3 (Sürekli Düzgün Dağılımın Momentleri). X sürekli rastgele değişkeninin yoğunluk fonksiyonu

$$f_X(x) = \begin{cases} \frac{1}{2}, & 0 < x < 2 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun. $E(X)$, $E(X^2)$ ve $\text{Var}(X)$ değerlerini bulunuz.

Çözüm.

f_X sınırlı bir aralık dışında sıfır olduğundan bütün momentler vardır. Beklenen değer:

$$E(X) = \int_0^2 x \cdot \frac{1}{2} dx = \frac{1}{2} \left[\frac{x^2}{2} \right]_0^2 = \frac{1}{2} \cdot 2 = 1.$$

İkinci moment:

$$E(X^2) = \int_0^2 x^2 \cdot \frac{1}{2} dx = \frac{1}{2} \left[\frac{x^3}{3} \right]_0^2 = \frac{1}{2} \cdot \frac{8}{3} = \frac{4}{3}.$$

Varyans:

$$\text{Var}(X) = E(X^2) - (E(X))^2 = \frac{4}{3} - 1 = \frac{1}{3}.$$

Tanımından doğrulama: $\mu = 1$ için

$$\text{Var}(X) = \int_0^2 (x-1)^2 \cdot \frac{1}{2} dx = \frac{1}{2} \left[\frac{(x-1)^3}{3} \right]_0^2 = \frac{1}{2} \left(\frac{1}{3} - \left(-\frac{1}{3} \right) \right) = \frac{1}{3}.$$

Standart sapma $\sigma_X = \frac{1}{\sqrt{3}} \approx 0,58$ 'dir. Beklenen değer $(0, 2)$ aralığının orta noktası olan 1 çıkması, yoğunluğun simetrisinin doğal bir sonucudur.

Teorem 16.2'i bu örnekte deneyelim: $Y = 3X - 1$ için hiçbir yeni integral hesaplamadan $E(Y) = 3 \cdot 1 - 1 = 2$ ve $\text{Var}(Y) = 9 \cdot \frac{1}{3} = 3$ bulunur.

■

Örnek 16.4 (Kuyruğu Kuvvet Gibi Sönen Bir Yoğunluk). X sürekli rastgele değişkeninin yoğunluk fonksiyonu

$$f_X(x) = \begin{cases} 3x^{-4}, & x > 1 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun. $E(X)$, $E(X^2)$ ve $\text{Var}(X)$ değerlerini bulunuz. X 'in hangi mertebeden momentleri vardır?

Çözüm.

Önce f_X 'in bir yoğunluk olduğunu görelim: $f_X \geq 0$ ve

$$\int_1^\infty 3x^{-4} dx = \lim_{b \rightarrow \infty} \left[-x^{-3} \right]_1^b = \lim_{b \rightarrow \infty} \left(1 - \frac{1}{b^3} \right) = 1.$$

Beklenen değer. İntegrand $x > 1$ üzerinde pozitif olduğundan mutlak yakınsaklık ile yakınsaklık aynı şeydir:

$$E(X) = \int_1^\infty x \cdot 3x^{-4} dx = \int_1^\infty 3x^{-3} dx = \lim_{b \rightarrow \infty} \left[-\frac{3}{2}x^{-2} \right]_1^b = \frac{3}{2}.$$

İkinci moment:

$$E(X^2) = \int_1^\infty x^2 \cdot 3x^{-4} dx = \int_1^\infty 3x^{-2} dx = \lim_{b \rightarrow \infty} \left[-3x^{-1} \right]_1^b = 3.$$

Varyans:

$$\text{Var}(X) = E(X^2) - (E(X))^2 = 3 - \left(\frac{3}{2} \right)^2 = 3 - \frac{9}{4} = \frac{3}{4}.$$

Standart sapma $\sigma_X = \frac{\sqrt{3}}{2} \approx 0,87$ 'dir.
Üçüncü momente bakalım:

$$\int_1^\infty x^3 \cdot 3x^{-4} dx = \int_1^\infty \frac{3}{x} dx = \lim_{b \rightarrow \infty} 3 \ln b = \infty.$$

Öyleyse $E(X^3)$ yoktur; **Önerme 16.1'in (c)** maddesi gereği $k \geq 3$ için hiçbir $E(X^k)$ yoktur. Genel olarak x^{-4} gibi kuvvet biçiminde sönen bir kuyruk, ancak sönmeye hızının izin verdiği kadar momente sahiptir: $x^k \cdot x^{-4}$ integrallenebilir olsun diye $k - 4 < -1$, yani $k < 3$ gerekir. **Örnek 16.1**'taki x^{-2} kuyruğu için aynı ölçüt $k < 1$ verir; bu yüzden orada birinci moment bile yoktu.

■

Örnek 16.5 (Üçgen Yoğunluğun Momentleri). X sürekli rastgele değişkeninin yoğunluk fonksiyonu

$$f_X(x) = \begin{cases} \frac{1}{4} (2 - |x|), & -2 \leq x \leq 2 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun. $E(X)$ ve $\text{Var}(X)$ değerlerini bulunuz.

Çözüm.

Grafiği $(-2, 0)$, $(0, \frac{1}{2})$ ve $(2, 0)$ köşeli bir üçgen olan bu fonksiyon negatif değildir ve altında kalan alan üçgenin alanıdır: $\frac{1}{2} \cdot 4 \cdot \frac{1}{2} = 1$. İntegrale de doğrulayalım; f_X çift fonksiyon olduğundan

$$\int_{-2}^2 f_X(x) dx = 2 \int_0^2 \frac{1}{4}(2-x) dx = \frac{1}{2} \left[2x - \frac{x^2}{2} \right]_0^2 = \frac{1}{2}(4-2) = 1.$$

Beklenen değer. $xf_X(x)$ tek fonksiyondur ve $[-2, 2]$ sınırlı aralığında integrallenebilir; öyleyse mutlak yakınsaklık koşulu sağlanır ve simetri kullanılabilir: $E(X) = 0$. Açık hesapla da görelim:

$$E(X) = \int_{-2}^0 x \cdot \frac{2+x}{4} dx + \int_0^2 x \cdot \frac{2-x}{4} dx = \frac{1}{4} \left[x^2 + \frac{x^3}{3} \right]_{-2}^0 + \frac{1}{4} \left[x^2 - \frac{x^3}{3} \right]_0^2 = \frac{1}{4} \left(-\frac{4}{3} \right) + \frac{1}{4} \cdot \frac{4}{3} = 0.$$

İkinci moment. $x^2 f_X(x)$ çift fonksiyondur:

$$E(X^2) = 2 \int_0^2 x^2 \cdot \frac{1}{4}(2-x) dx = \frac{1}{2} \int_0^2 (2x^2 - x^3) dx = \frac{1}{2} \left[\frac{2x^3}{3} - \frac{x^4}{4} \right]_0^2 = \frac{1}{2} \left(\frac{16}{3} - 4 \right) = \frac{1}{2} \cdot \frac{4}{3} = \frac{2}{3}.$$

Varyans. $E(X) = 0$ olduğundan

$$\text{Var}(X) = E(X^2) - 0^2 = \frac{2}{3}, \quad \sigma_X = \sqrt{\frac{2}{3}} \approx 0,82.$$

Karşılaştırma için $[-2, 2]$ üzerindeki düzgün yoğunluğu, $f(x) = \frac{1}{4}$, düşünelim: beklenen değeri yine 0, ikinci momenti $\int_{-2}^2 \frac{x^2}{4} dx = \frac{4}{3}$ 'tür. Üçgen yoğunluk kütlelerini merkeze yığdığından varyansı düzgün yoğunluğunun yarısıdır. Varyans, aynı aralık üzerindeki iki dağılımı yayılımlarına göre ayırt etmektedir.

■

16.6 Markov ve Chebyshev Eşitsizlikleri

Varyansın yayılımı ölçtüğünü söyledik; ama “küçük varyans, merkeze yakın değerler” cümlesi şimdilik bir slogandan ibarettir. Bunu kesin bir ifadeye çevirmek istiyoruz: varyansı bilinen bir rastgele değişkenin, beklenen değerinden en az ε uzağa düşme olasılığı en çok ne olabilir? Dağılımın kendisi hakkında hiçbir şey bilmeden yanıt verilebilmesi şaşırtıcıdır. Önce daha temel bir eşitsizlik gerekir.

Teorem 16.3 (Markov Eşitsizliği). X , değer kümesi $D_X \subseteq [0, \infty)$ olan ve beklenen değeri var olan bir rastgele değişken olsun. Her $a > 0$ için

$$P(X \geq a) \leq \frac{E(X)}{a}$$

olur.

İspat.

X kesikli ise. D_X 'teki her x negatif olmadığından $xf_X(x) \geq 0$ 'dır; toplamdan $x < a$ olan terimleri atmak toplamı küçültür ya da değiştirmez:

$$E(X) = \sum_{x \in D_X} x f_X(x) \geq \sum_{x \in D_X, x \geq a} x f_X(x) \geq \sum_{x \in D_X, x \geq a} a f_X(x) = a \sum_{x \in D_X, x \geq a} f_X(x) = a P(X \geq a).$$

İkinci eşitsizlikte, kalan terimlerde $x \geq a$ olduğundan x 'in yerine daha küçük olan a yazılmıştır. Her iki taraf $a > 0$ 'a bölünürse istenen çıkar.

X sürekli ise. $x < 0$ için $f_X(x) = 0$ olduğundan integrand $[0, \infty)$ dışında sıfırdır ve aynı adımlar integralle yazılır:

$$E(X) = \int_0^\infty x f_X(x) dx \geq \int_a^\infty x f_X(x) dx \geq \int_a^\infty a f_X(x) dx = a P(X \geq a).$$

■

Markov eşitsizliği yalnız beklenen değeri kullanır ve $X \geq 0$ koşuluna dayanır. Örneğin ortalama geliri $E(X)$ olan bir toplulukta, gelirin dağılımı ne olursa olsun, geliri ortalamanın 10 katından fazla olanların oranı $\frac{1}{10}$ 'u geçemez. Şimdi bunu $(X - \mu)^2$ rastgele değişkenine uygulayalım.

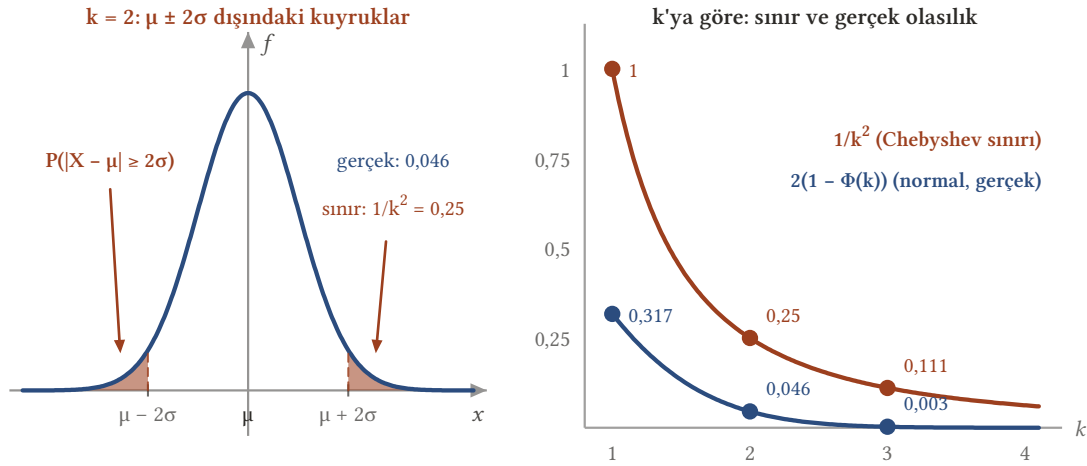
Teorem 16.4 (Chebyshev Eşitsizliği). X , varyansı var olan bir rastgele değişken, $\mu = E(X)$ ve $\sigma^2 = \text{Var}(X)$ olsun. Her $\varepsilon > 0$ için

$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}, \quad \text{eşdeğer olarak} \quad P(|X - \mu| < \varepsilon) \geq 1 - \frac{\sigma^2}{\varepsilon^2}$$

olur. Özel olarak $\sigma > 0$ ve $k > 0$ için $\varepsilon = k\sigma$ alınırsa

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

bulunur.



Şekil 16.3: Chebyshev eşitsizliği, beklenen değerden en az $k\sigma$ uzağa düşme olasılığını dağılımdan bağımsız olarak $1/k^2$ ile sınırlar. Solda bir yoğunluğun $\mu \pm 2\sigma$ dışındaki iki kuyruğu taralıdır: sınır $1/4$ 'tür, normal dağılım için gerçek olasılık 0,046'dır. Sağda k büyüdükçe iki eğri de sıfıra iner; ama sınır (turuncu) gerçek olasılığın (mavi) hep üstünde kalır ve aradaki açık, sınırın genelliğinin bedelidir.

İspat.

$g(x) = (x - \mu)^2$ Borel ölçülebilir bir fonksiyondur (süreklî); $Y = g(X) = (X - \mu)^2$ rastgele değişkeninin değerleri negatif değildir ve $E(Y) = \text{Var}(X) = \sigma^2$ vardır. $|X - \mu| \geq \varepsilon$ olayı ile $Y \geq \varepsilon^2$ olayı aynıdır, çünkü her iki taraf da negatif olmayan sayılardır ve kare alma $[0, \infty)$ üzerinde artandır. Y 'ye $a = \varepsilon^2$ ile Teorem 16.3 uygulanırsa

$$P(|X - \mu| \geq \varepsilon) = P(Y \geq \varepsilon^2) \leq \frac{E(Y)}{\varepsilon^2} = \frac{\sigma^2}{\varepsilon^2}$$

çıkar. İkinci biçim tümleyen olayın olasılığıdır; üçüncüsü $\varepsilon = k\sigma$ yazılarak elde edilir.

■

Chebyshev eşitsizliği varyansın anlamını kesinleştirir: X 'in beklenen değerinden 2σ 'dan fazla sapma olasılığı en çok $\frac{1}{4}$, 3σ 'dan fazla sapma olasılığı en çok $\frac{1}{9}$ 'dur. Standart sapma, dağılımın “doğal uzunluk birimi”dir; merkezden birkaç σ uzağa gidildiğinde olasılık kütesinin büyük bölümü geride kalmak zorundadır. Eşitsizliğin gücü genelliğindedir: X 'in dağılımı hakkında μ ve σ^2 dışında hiçbir bilgi gerektirmez. Bedeli de aynı genelliştir; somut bir dağılım için verdiği sınır çoğu zaman gerçek olasılıktan çok kabadır. Aşağıdaki örnek her iki yönü de gösterir.

Örnek 16.6 (Chebyshev Eşitsizliğinin Uygulanması). (a) Bir dolum makinesinin doldurduğu paketlerin ağırlığı X (gram) rastgele değişkeninin dağılımı bilinmiyor; yalnız $E(X) = 500$ ve $\sigma_X = 10$ olduğu biliniyor. Bir paketin ağırlığının 480 ile 520 gram arasında olma olasılığı için bir alt sınır, 500'den en az 30 gram sapma olasılığı için bir üst sınır bulunuz.

(b) Örnek 16.3'teki X için Chebyshev eşitsizliğinin $\varepsilon = 1$ 'de verdiği sınırı gerçek olasılıkla karşılaştırınız.

(c) $k \geq 1$ bir tam sayı olmak üzere, X kesikli rastgele değişkeninin olasılık fonksiyonu

$$f_X(-1) = f_X(1) = \frac{1}{2k^2}, \quad f_X(0) = 1 - \frac{1}{k^2}$$

olsun. Chebyshev eşitsizliğinin bu değişken için $\varepsilon = 1$ 'de eşitlikle sağlandığını gösteriniz.

Çözüm.

(a) $\mu = 500$, $\sigma^2 = 100$. $480 < X < 520$ olayı $|X - 500| < 20$ olayıdır; $\varepsilon = 20$ ile

$$P(480 < X < 520) = P(|X - 500| < 20) \geq 1 - \frac{100}{400} = \frac{3}{4}.$$

$\varepsilon = 30$ ile

$$P(|X - 500| \geq 30) \leq \frac{100}{900} = \frac{1}{9}.$$

Demek ki makine hakkında yalnız ortalama ve standart sapma bilinse bile, paketlerin en az dörtte üçünün 480–520 gram aralığında olduğu ve en çok dokuzda birinin 500 gramdan en az 30 gram saptığı söylenebilir.

(b) X , $(0, 2)$ üzerinde düzgün dağılmıştı: $\mu = 1$, $\sigma^2 = \frac{1}{3}$. Chebyshev eşitsizliği $\varepsilon = 1$ için

$$P(|X - 1| \geq 1) \leq \frac{1/3}{1} = \frac{1}{3}$$

der. Oysa X değerlerini yalnız $(0, 2)$ aralığında aldığından $|X - 1| \geq 1$ olayı, yani $X \leq 0$ ya da $X \geq 2$ olayı, sıfır olasılıklıdır: gerçek olasılık 0'dır. Sınır doğru ama kabadır.

(c) Önce f_X 'in bir olasılık fonksiyonu olduğunu görelim: değerler negatif değildir ve toplamaları $\frac{1}{2k^2} + \frac{1}{2k^2} + 1 - \frac{1}{k^2} = 1$ 'dir. Beklenen değer, simetri gereği ya da doğrudan,

$$\mu = (-1) \cdot \frac{1}{2k^2} + 0 \cdot \left(1 - \frac{1}{k^2}\right) + 1 \cdot \frac{1}{2k^2} = 0;$$

varyans

$$\sigma^2 = E(X^2) - \mu^2 = E(X^2) = (-1)^2 \cdot \frac{1}{2k^2} + 0^2 \cdot \left(1 - \frac{1}{k^2}\right) + 1^2 \cdot \frac{1}{2k^2} = \frac{1}{k^2}.$$

$\varepsilon = 1$ için Chebyshev sınırı $\frac{\sigma^2}{\varepsilon^2} = \frac{1}{k^2}$ 'dir. Gerçek olasılık ise

$$P(|X - \mu| \geq 1) = P(X = -1) + P(X = 1) = \frac{1}{2k^2} + \frac{1}{2k^2} = \frac{1}{k^2}.$$

İki taraf eşittir. Demek ki Chebyshev eşitsizliğindeki $\frac{\sigma^2}{\varepsilon^2}$ sınırı genel olarak iyileştirilemez: kütlelerini merkeze ve iki uca dağıtan bu dağılım sınırı tam olarak doldurur.

■

Söz verdiğimiz gibi, sürekli bir rastgele değişkenin varyansının sıfır olamayacağını Chebyshev eşitsizliğiyle gösterebiliriz. X sürekli ve $\text{Var}(X) = 0$ olsa, her $\varepsilon > 0$ için $P(|X - \mu| \geq \varepsilon) \leq 0$, yani $P(|X - \mu| < \varepsilon) = 1$ olurdu. $A_n = (|X - \mu| < \frac{1}{n})$ olayları azalan bir dizi oluşturur ve kesişimleri $(X = \mu)$ olayıdır; Teorem 3.5 gereği

$$P(X = \mu) = \lim_{n \rightarrow \infty} P(A_n) = 1$$

olurdu. Oysa sürekli bir rastgele değişkende her nokta sıfır olasılıklıdır: $P(X = \mu) = 0$. Bu çelişki gösterir ki sürekli bir rastgele değişkenin varyansı her zaman pozitifdir.

16.7 Alıştırmalar

Alıştırma 16.1 (Beklenen Değer, Momentler ve Varyans). (a) X kesikli rastgele değişkeninin olasılık fonksiyonu $f_X(x) = \frac{x}{10}$, $x \in \{1, 2, 3, 4\}$ olsun. $E(X)$, $E(X^2)$, $\text{Var}(X)$ ve σ_X değerlerini bulunuz.

(b) X sürekli rastgele değişkeninin yoğunluk fonksiyonu $f_X(x) = 2x$ ($0 < x < 1$), diğer durumlarda 0 olsun. $E(X)$, $\text{Var}(X)$ ve $E(X^2)$ değerlerini bulunuz.

(c) X sürekli rastgele değişkeninin yoğunluk fonksiyonu $f_X(x) = e^{-x}$ ($x > 0$), diğer durumlarda 0 olsun. Her $k \in \mathbb{N}$ için $E(X^k) = k!$ olduğunu gösteriniz; buradan $\text{Var}(X)$ 'i bulunuz.

(d) $E(X) = 2$ ve $\text{Var}(X) = 5$ olsun. $Y = 3X - 2$ için $E(Y)$, $\text{Var}(Y)$ ve $E(Y^2)$ değerlerini bulunuz.

(e) X kesikli rastgele değişkeninin olasılık fonksiyonu $f_X(x) = \frac{1}{x(x+1)}$, $x \in \{1, 2, 3, \dots\}$ olsun. f_X 'in bir olasılık fonksiyonu olduğunu, ama $E(X)$ 'in var olmadığını gösteriniz.

(f) İkinci momenti var olan her X için $\text{Var}(X) = E[X(X-1)] + E(X) - (E(X))^2$ olduğunu gösteriniz. Örnek 16.2'teki X için ikinci çarpımsal moment $E[X(X-1)]$ 'i hesaplayıp formülü doğrulayınız.

(g) Bir ölçüm aletinin verdiği değer X için $E(X) = 50$ ve $\sigma_X = 5$ biliniyor. $P(35 < X < 65)$ için bir alt sınır bulunuz. $P(|X - 50| < k\sigma_X) \geq 0,99$ olmasını garanti eden en küçük k tam sayısını ve buna karşılık gelen aralığı belirleyiniz.

Çözüm.

(a) $\sum_{x=1}^4 \frac{x}{10} = \frac{10}{10} = 1$, gerçekten bir olasılık fonksiyonudur.

$$E(X) = \sum_{x=1}^4 x \cdot \frac{x}{10} = \frac{1+4+9+16}{10} = \frac{30}{10} = 3, \quad E(X^2) = \sum_{x=1}^4 x^2 \cdot \frac{x}{10} = \frac{1+8+27+64}{10} = \frac{100}{10} = 10.$$

$\text{Var}(X) = 10 - 3^2 = 1$ ve $\sigma_X = 1$.

(b) $\int_0^1 2x \, dx = 1$, bir yoğunluktur.

$$E(X) = \int_0^1 x \cdot 2x \, dx = \left[\frac{2x^3}{3} \right]_0^1 = \frac{2}{3}, \quad E(X^2) = \int_0^1 x^2 \cdot 2x \, dx = \left[\frac{x^4}{2} \right]_0^1 = \frac{1}{2}.$$

$\text{Var}(X) = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{9-8}{18} = \frac{1}{18}$. $g(x) = \frac{1}{x}$ için

$$\left(E\left(\frac{1}{X}\right)\right) = \int_0^1 \frac{1}{x} \cdot 2x \, dx = \int_0^1 2 \, dx = 2.$$

g fonksiyonu $(0,1)$ üzerinde sınırsız olduğu hâlde $g(x)f_X(x) = 2$ sınırlı kaldığından beklenen değer vardır. $\left(E\left(\frac{1}{X}\right)\right) = 2 \neq \frac{3}{2} = \frac{1}{E(X)}$ olduğuna dikkat ediniz; beklenen değer doğrusal olmayan fonksiyonların içine girmez.

(c) $I_k = \int_0^\infty x^k e^{-x} \, dx$ yazalım; integrand pozitif olduğundan mutlak yakınsaklık ile yakınsaklık aynıdır. $I_0 = \int_0^\infty e^{-x} \, dx = 1$ 'dir; bu aynı zamanda f_X 'in bir yoğunluk olduğunu gösterir. $k \geq 1$ için $u = x^k$, $dv = e^{-x} \, dx$ ile kısmi integrasyon:

$$I_k = \left[-x^k e^{-x} \right]_0^\infty + k \int_0^\infty x^{k-1} e^{-x} \, dx = 0 + k I_{k-1},$$

çünkü $x \rightarrow \infty$ iken $x^k e^{-x} \rightarrow 0$ 'dır. Tümevarımla $I_k = k \cdot (k-1) \cdots 1 \cdot I_0 = k!$ bulunur; yani her merbeden moment vardır ve $E(X^k) = k!$ 'dir. Özel olarak $E(X) = 1$, $E(X^2) = 2$ ve

$$\text{Var}(X) = 2 - 1^2 = 1.$$

(d) Teorem 16.2 ile $E(Y) = 3 \cdot 2 - 2 = 4$ ve $\text{Var}(Y) = 3^2 \cdot 5 = 45$. Önerme 16.3'nü Y için ters yönde kullanırsak

$$E(Y^2) = \text{Var}(Y) + (E(Y))^2 = 45 + 16 = 61.$$

Aynı sonuca $E(X^2) = 5 + 4 = 9$ ve lineerlik yoluyla da ulaşılır: $E(Y^2) = E(9X^2 - 12X + 4) = 9 \cdot 9 - 12 \cdot 2 + 4 = 61$.

(e) Değerler pozitifdir. $\frac{1}{x(x+1)} = \frac{1}{x} - \frac{1}{x+1}$ olduğundan kısmi toplamlar iç içe sadeleşir:

$$\sum_{x=1}^n \frac{1}{x(x+1)} = \sum_{x=1}^n \left(\frac{1}{x} - \frac{1}{x+1} \right) = 1 - \frac{1}{n+1} \xrightarrow{n \rightarrow \infty} 1.$$

Demek ki f_X bir olasılık fonksiyonudur. Beklenen değer için mutlak yakınsaklık koşuluna bakalım; terimler pozitif olduğundan koşul serinin yakınsaması demektir:

$$\sum_{x=1}^\infty x \cdot \frac{1}{x(x+1)} = \sum_{x=1}^\infty \frac{1}{x+1} = \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots = \infty,$$

çünkü bu, harmonik serinin ilk terimi atılmış hâlidir ve harmonik seri iraksaktır. $E(X)$ yoktur. Bu, Örnek 16.1'un kesikli karşılığıdır: $f_X(x)$, x^{-2} hızıyla sönmektedir.

(f) $X(X-1) = X^2 - X$ olduğundan lineerlik gereği $E[X(X-1)] = E(X^2) - E(X)$, yani $E(X^2) = E[X(X-1)] + E(X)$ 'tir. Bunu Önerme 16.3'nde yerine koyarsak

$$\text{Var}(X) = E(X^2) - (E(X))^2 = E[X(X-1)] + E(X) - (E(X))^2.$$

Örnek 16.2'te $E(X) = \frac{1}{2}$ ve $E(X^2) = \frac{19}{6}$ bulunmuştu; öyleyse $E[X(X-1)] = \frac{19}{6} - \frac{1}{2} = \frac{16}{6} = \frac{8}{3}$. Doğrudan da hesaplanabilir: $x(x-1)$ değerleri $x = -2, \dots, 3$ için 6, 2, 0, 0, 2, 6'dır ve ortalamaları $\frac{16}{6} = \frac{8}{3}$ 'tür. Formül:

$$\frac{8}{3} + \frac{1}{2} - \frac{1}{4} = \frac{32 + 6 - 3}{12} = \frac{35}{12} = \text{Var}(X).$$

Çarpımsal momentler, kesikli dağılımlarda $E(X^2)$ 'den daha kolay hesaplanır; bunun nedenini [sonraki bölümde](#) çarpımsal moment üreten fonksiyonla göreceğiz.

(g) $35 < X < 65$ olayı $|X - 50| < 15 = 3\sigma_X$ olayıdır. [Teorem 16.4](#) ile

$$P(35 < X < 65) \geq 1 - \frac{1}{3^2} = \frac{8}{9}.$$

$P(|X - 50| < k\sigma_X) \geq 1 - \frac{1}{k^2}$ olduğundan, $1 - \frac{1}{k^2} \geq 0,99$, yani $k^2 \geq 100$ olması yeter; en küçük tam sayı $k = 10$ 'dur. Karşılık gelen aralık $|X - 50| < 50$, yani $(0, 100)$ 'dür: dağılım ne olursa olsun, ölçümlerin en az yüzde 99'u 0 ile 100 arasındadır.

■

Beklenen değer ve varyans, bir dağılımın iki sayıya sığan özetidir; ama dağılımın bütün momentlerini tek bir fonksiyonda toplamak ve bu fonksiyondan dağılımı geri okumak da mümkündür: [Karakteristik ve Moment Üreten Fonksiyonlar](#).

17. Karakteristik ve Moment Üreten Fonksiyonlar

Beklenen Değer, Momentler ve Varyans bölümünde bir rastgele değişkenin momentlerini $E(X^k)$ tek tek, her seferinde ayrı bir toplam ya da integral hesaplayarak bulduk. Bu bölümde bütün momentleri **tek bir fonksiyonun içine sıkıştırarak** üç araç tanıtaacağız: karakteristik fonksiyon $\varphi_X(t)$, moment üreten fonksiyon $M_X(t)$ ve çarpımsal moment üreten fonksiyon $N_X(t)$. Üçü de aynı fikre dayanır: X yerine e^{itX} , e^{tX} ya da t^X gibi bir üstel ifadenin beklenen değeri alınır ve t değişkeninin bir fonksiyonu elde edilir. Bu fonksiyonun $t = 0$ (ya da $t = 1$) noktasındaki türevleri momentleri verir.

Bu araçların değeri yalnızca hesap kolaylığı değildir. Karakteristik fonksiyon, bir rastgele değişkenin dağılımını **tek biçimde belirler**: iki rastgele değişkenin karakteristik fonksiyonları aynıysa dağılımları da aynıdır. Bu yüzden karakteristik fonksiyon, dağılımın bir tür “parmak izi”dir. Bir dağılımı bulmak için dağılım fonksiyonunu ya da yoğunluğu doğrudan hesaplamak yerine karakteristik fonksiyonunu hesaplayıp tanımak çoğu zaman çok daha kolaydır. Sonraki bölümlerde bağımsız rastgele değişkenlerin toplamını incelerken bu yaklaşımın gücünü göreceğiz.

Bölümde önce karakteristik fonksiyonu tanımlayacak, her rastgele değişken için var olduğunu göstereceğiz; temel özelliklerini ve momentlerle ilişkisini ispatlayacak, teklik teoremini ifade edeceğiz. Ardından moment üreten fonksiyonu, karakteristik fonksiyonla bağıntı ve moment üretme mekanizmasını ele alacağız. Son olarak çarpımsal moment üreten fonksiyonu ve çarpımsal momentleri inceleyeceğiz. Düzgün, iki noktalı, Laplace, üstel ve Poisson dağılımlarında bütün hesapları sonuna kadar yapacağız.

17.1 Karakteristik Fonksiyon

Karakteristik fonksiyon, karmaşık değerli bir rastgele değişkenin beklenen değeridir. Bu yüzden önce bir küçük hazırlık gerekir. U ve V gerçel rastgele değişkenler olmak üzere $Z = U + iV$ biçimindeki bir nesneye **karmaşık değerli rastgele değişken** diyelim. $E(U)$ ve $E(V)$ varsa Z 'nin beklenen değeri

$$E(Z) = E(U) + i E(V)$$

ile tanımlanır. Euler formülü $e^{i\theta} = \cos \theta + i \sin \theta$ gereği, her $t \in \mathbb{R}$ için

$$e^{itX} = \cos(tX) + i \sin(tX)$$

karmaşık değerli bir rastgele değişkendir ve gerçel ile sanal kısımları $\cos(tX)$ ve $\sin(tX)$ 'dir.

Tanım 17.1 (Karakteristik Fonksiyon). X bir rastgele değişken olsun. Her $t \in \mathbb{R}$ için

$$\varphi_X(t) = (E^{itX}) = E(\cos(tX)) + i E(\sin(tX))$$

ile tanımlanan $\varphi_X : \mathbb{R} \rightarrow \mathbb{C}$ fonksiyonuna X rastgele değişkeninin **karakteristik fonksiyonu** denir.

Beklenen değer tanımları (**Tanım 16.1**) $g(x) = \cos(tx)$ ve $g(x) = \sin(tx)$ fonksiyonlarına uygulanırsa karakteristik fonksiyonun hesap formülleri ortaya çıkar. X kesikli ve olasılık fonksiyonu f_X ise

$$\varphi_X(t) = \sum_{x \in D_X} e^{itx} f_X(x);$$

X sürekli ve yoğunluk fonksiyonu f_X ise

$$\varphi_X(t) = \int_{-\infty}^{\infty} e^{itx} f_X(x) dx.$$

Her rastgele değişkenin karakteristik fonksiyonu vardır. Bunun nedeni $|e^{itx}| = 1$, dolayısıyla $|\cos(tx)| \leq 1$ ve $|\sin(tx)| \leq 1$ olmasıdır: kesikli durumda

$$\sum_{x \in D_X} |\cos(tx)| f_X(x) \leq \sum_{x \in D_X} f_X(x) = 1 < \infty$$

olduğundan $E(\cos(tx))$ vardır; sin için ve sürekli durumda integral için aynı hesap geçerlidir. Bu, karakteristik fonksiyonu daha sonra tanımlayacağımız moment üreten fonksiyondan ayıran en önemli özelliktir: moment üreten fonksiyon bazı dağılımlar için yoktur, karakteristik fonksiyon ise her zaman vardır.

Karmaşık değerli beklenen değerle ilgili üç basit özelliğe sık sık gerek duyacağız.

İ Karmaşık değerli beklenen değer üç özelliği

$Z = U + iV$ karmaşık değerli bir rastgele değişken ve $E(Z)$ var olsun.

(i) **Doğrusallık.** $c = \alpha + i\beta \in \mathbb{C}$ sabit olmak üzere $E(cZ) = c E(Z)$ ve $E(Z_1 + Z_2) = E(Z_1) + E(Z_2)$ olur. Gerçekten $cZ = (\alpha U - \beta V) + i(\alpha V + \beta U)$ 'dir. Gerçek beklenen değer doğrusallığı (Teorem 16.1 ve Önerme 16.2) gerçek ve sanal kısımlara ayrı ayrı uygulanınca sonuç şudur:

$$E(cZ) = \alpha E(U) - \beta E(V) + i(\alpha E(V) + \beta E(U)) = (\alpha + i\beta)(E(U) + iE(V)).$$

(ii) **Eşlenik.** $\overline{E(Z)} = E(\overline{Z})$; çünkü $\overline{Z} = U - iV$ ve $\overline{E(Z)} = E(U) - iE(V)$.

(iii) **Modül eşitsizliği.** $|E(Z)| \leq E|Z|$. Bunu görmek için $E(Z) = re^{i\theta}$ ($r \geq 0$) yazalım. (i) gereği $r = e^{-i\theta} E(Z) = E(e^{-i\theta} Z)$ olur. Sol taraf gerçek olduğundan sağ tarafın sanal kısmı sıfırdır; öyleyse $r = E(\operatorname{Re}(e^{-i\theta} Z))$. Her karmaşık sayı için $\operatorname{Re} w \leq |w|$ olduğundan ve $|e^{-i\theta} Z| = |Z|$ olduğundan $r \leq E|Z|$ elde edilir.

17.1.1 Temel Özellikler

Bir fonksiyonun bir karakteristik fonksiyon olup olamayacağını hızla anlamak, karakteristik fonksiyonu bilinen bir değişkenden $aX + b$ gibi türetilmiş değişkenlere geçmek ve ileride teklik teoremini kullanmak için aşağıdaki özellikler gerekir.

Teorem 17.1 (Karakteristik Fonksiyonun Temel Özellikleri). X bir rastgele değişken ve φ_X karakteristik fonksiyonu olsun.

(a) $\varphi_X(0) = 1$.

(b) Her $t \in \mathbb{R}$ için $|\varphi_X(t)| \leq 1$.

(c) $\varphi_X, \mathbb{R}$ üzerinde düzgün süreklidir.

(d) $a, b \in \mathbb{R}$ sabit olmak üzere, her $t \in \mathbb{R}$ için

$$\varphi_{aX+b}(t) = e^{itb} \varphi_X(at).$$

(e) Her $t \in \mathbb{R}$ için $\varphi_X(-t) = \overline{\varphi_X(t)}$. Özel olarak X ile $-X$ aynı dağılıma sahipse φ_X gerçel değerlidir.

İspat.

(a) $e^{i \cdot 0 \cdot X} = 1$ olduğundan $\varphi_X(0) = E(1) = 1$.

(b) Modül eşitsizliği ile $|\varphi_X(t)| = |E(e^{itX})| \leq E|e^{itX}| = E(1) = 1$.

(c) $t \in \mathbb{R}$ ve $h \in \mathbb{R}$ olsun. $e^{i(t+h)X} - e^{itX} = e^{itX}(e^{ihX} - 1)$ olduğundan, doğrusallık ve modül eşitsizliği ile

$$|\varphi_X(t+h) - \varphi_X(t)| = |E(e^{itX}(e^{ihX} - 1))| \leq E(|e^{itX}| |e^{ihX} - 1|) = E|e^{ihX} - 1|$$

bulunur. Sağ taraf t 'ye bağlı değildir; ona $\delta(h)$ diyelim. $h \rightarrow 0$ iken her ω için $e^{ihX(\omega)} \rightarrow 1$, yani $|e^{ihX(\omega)} - 1| \rightarrow 0$ olur; üstelik $|e^{ihX} - 1| \leq |e^{ihX}| + 1 = 2$ olduğundan bu fonksiyonlar h 'den bağımsız bir sabitle sınırlıdır. Sınırlı yakınsaklık teoremi (sınırlı ve noktasal yakınsayan bir dizinin beklenen değerinin limite yakınsaması) gereği $\delta(h) \rightarrow 0$ olur. Öyleyse verilen $\varepsilon > 0$ için $|h| < h_0$ olduğunda $\delta(h) < \varepsilon$ olacak biçimde bir $h_0 > 0$ vardır ve bu h_0 her t için işe yarar: $|h| < h_0$ olan her $t, t+h$ çifti için $|\varphi_X(t+h) - \varphi_X(t)| < \varepsilon$. Bu, düzgün sürekliliğin tanımıdır.

(d) $e^{it(aX+b)} = e^{itb} e^{i(at)X}$ yazılır; e^{itb} bir sabit olduğundan doğrusallık ile

$$\varphi_{aX+b}(t) = E(e^{itb} e^{i(at)X}) = e^{itb} E(e^{i(at)X}) = e^{itb} \varphi_X(at).$$

(e) $\overline{e^{itX}} = \cos(tX) - i \sin(tX) = e^{-itX}$ olduğundan, eşlenik özelliği ile $\overline{\varphi_X(t)} = E(e^{-itX}) = \varphi_X(-t)$. Şimdi X ile $-X$ aynı dağılıma sahip olsun. Karakteristik fonksiyon yalnızca dağılıma (olasılık ya da yoğunluk fonksiyonuna) bağlı olduğundan $\varphi_{-X} = \varphi_X$ olur. Öte yandan (d) şıkkı $a = -1$, $b = 0$ ile $\varphi_{-X}(t) = \varphi_X(-t) = \overline{\varphi_X(t)}$ verir. Böylece $\varphi_X(t) = \overline{\varphi_X(t)}$, yani $\varphi_X(t)$ gerçeldir. ■

Özelliklerden (a) ve (b)'nin pratik bir kullanımı vardır: verilen bir fonksiyon $t = 0$ 'da 1 değerini almıyorsa ya da modülü bir yerde 1'i aşıyorsa, o fonksiyon hiçbir rastgele değişkenin karakteristik fonksiyonu olamaz. Örneğin $\varphi(t) = 2 \cos t$ ya da $\varphi(t) = e^t$ karakteristik fonksiyon değildir. Bu koşulların yeterli olmadığını da not edelim: (a) ve (b)'yi sağlayan ama karakteristik fonksiyon olmayan fonksiyonlar vardır.

17.1.2 Karakteristik Fonksiyondan Momentler

Karakteristik fonksiyonun asıl işlevlerinden biri momentleri türev olarak üretmesidir. Biçimsel olarak e^{itX} 'in t 'ye göre k . türevi $(iX)^k e^{itX}$ 'tir; $t = 0$ 'da bu $i^k X^k$ olur. Türev ile beklenen değerin yerinin değiştirilebildiğini görmek ispatın tek işidir.

Teorem 17.2 (Karakteristik Fonksiyondan Momentler). X rastgele değişkeninin n . mertebeden momentleri var, yani $E(|X|^n) < \infty$ olsun. O zaman $\varphi_X, \mathbb{R}$ üzerinde n kez türevlenebilir ve $k = 1, 2, \dots, n$ için

$$\varphi_X^{(k)}(t) = i^k (\mathbb{E} X^k e^{itX}), \quad \text{özel olarak} \quad \varphi_X^{(k)}(0) = i^k (\mathbb{E} X^k)$$

olur. Başka bir deyişle $E(X^k) = i^{-k} \varphi_X^{(k)}(0)$ 'dır.

İspat.

Önce $k \leq n$ için $E(|X|^k) < \infty$ olduğunu görelim: her gerçel x için $|x|^k \leq 1 + |x|^n$ olduğundan ($|x| \leq 1$ ise $|x|^k \leq 1$, $|x| > 1$ ise $|x|^k \leq |x|^n$) $E(|X|^k) \leq 1 + E(|X|^n) < \infty$.

Şimdi $k = 1$ durumunu ele alalım. $h \neq 0$ için fark oranı

$$\frac{\varphi_X(t+h) - \varphi_X(t)}{h} = \left(\mathbb{E} e^{itX} \frac{e^{ihX} - 1}{h} \right)$$

biçimindedir. $h \rightarrow 0$ iken her ω için $\frac{e^{ihX(\omega)} - 1}{h} \rightarrow iX(\omega)$ olur (bu, $u \mapsto e^{iuX(\omega)}$ fonksiyonunun $u = 0$ 'daki türevidir). Ayrıca her gerçel u için $|e^{iu} - 1| = 2 \left| \sin \frac{u}{2} \right| \leq |u|$ olduğundan

$$\left| e^{itX} \frac{e^{ihX} - 1}{h} \right| \leq \frac{|hX|}{|h|} = |X|$$

olur; sağ taraf h 'den bağımsızdır ve $E|X| < \infty$ 'dur. Baskın yakınsaklık teoremi gereği limit ile beklenen değerin yeri değiştirilebilir ve

$$\varphi_X'(t) = (\mathbb{E} X e^{itX}) = i(\mathbb{E} X^2 e^{itX})$$

bulunur. Aynı akıl yürütme $X e^{itX}$ için tekrarlanır: bu kez fark oranı $|X|^2$ ile sınırlıdır ve $E(|X|^2) < \infty$ 'dur; böylece $\varphi_X''(t) = i^2 \mathbb{E}(X^2 e^{itX})$ olur. k adımda $\varphi_X^{(k)}(t) = i^k \mathbb{E}(X^k e^{itX})$ elde edilir. $t = 0$ konursa $e^{i \cdot 0 \cdot X} = 1$ olduğundan $\varphi_X^{(k)}(0) = i^k \mathbb{E}(X^k)$ çıkar. ■

Teoremin kullanışlı bir sonucu, φ_X 'in $t = 0$ çevresindeki Taylor açılımıdır: $E(|X|^n) < \infty$ ise

$$\varphi_X(t) = \sum_{k=0}^n \frac{(it)^k}{k!} (\mathbb{E} X^k) + o(t^n) = 1 + it E(X) - \frac{t^2}{2} (\mathbb{E} X^2) - \frac{it^3}{6} (\mathbb{E} X^3) + \dots$$

olur. Bu yüzden φ_X 'in kapalı bir formülü biliniyorsa, momentleri türev almak yerine formülün kuvvet serisi açılımındaki katsayılardan okumak çoğu zaman daha pratiktir; aşağıdaki örneklerde bunu yapacağız.

17.1.3 Teklik ve Fourier Dönüşümü

Karakteristik fonksiyonun kuramsal önemi şu sorunun yanıtından gelir: φ_X bilindiğinde X 'in dağılımı bilinmiş olur mu? Yanıt evettir. Bunun arka planında karakteristik fonksiyonun bir Fourier dönüşümü olması yatar.

i Karakteristik fonksiyon bir Fourier dönüşümüdür

X sürekli ve yoğunluğu f_X ise $\varphi_X(t) = \int_{-\infty}^{\infty} e^{itx} f_X(x) dx$, yani φ_X tam olarak f_X 'in (işaret ve sabit uzlaşımı bir yana) **Fourier dönüşümü**dür. Fourier dönüşümü tersinirdir: $\int_{-\infty}^{\infty} |\varphi_X(t)| dt < \infty$ ise X sürekli bir rastgele değişkendir ve yoğunluğu **ters Fourier dönüşümü** ile geri kazanılır:

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi_X(t) dt.$$

Değer kümesi tam sayılardan oluşan kesikli bir X için de benzer bir formül vardır: $f_X(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-itx} \varphi_X(t) dt$. Her rastgele değişkenin karakteristik fonksiyonu var olduğundan ve bu dönüşümler için hazır tablolar bulunduğundan, uygulamada dağılımlar çoğu zaman karakteristik fonksiyonlarından tanınır. Aşağıdaki teklik teoremi bu ters çevirme formüllerinin (ve genel dağılım fonksiyonları için geçerli daha genel bir ters çevirme formülünün) bir sonucudur.

Teorem 17.3 (Karakteristik Fonksiyon için Teklik Teoremi). X ve Y rastgele değişkenlerinin karakteristik fonksiyonları her $t \in \mathbb{R}$ için eşitse, yani $\varphi_X(t) = \varphi_Y(t)$ ise, X ile Y aynı dağılıma sahiptir: her $x \in \mathbb{R}$ için $F_X(x) = F_Y(x)$.

Teoremin tersi açıktır: karakteristik fonksiyon yalnızca dağılıma bağlı olduğundan aynı dağılıma sahip iki değişkenin karakteristik fonksiyonları aynıdır. Teoremin kendisinin ispatı ters çevirme formülüne dayanır ve bu kitabın sınırları dışındadır; biz onu bir araç olarak kullanacağız. Kullanım biçimi şudur: bir Y rastgele değişkeninin dağılımı aranıyorsa φ_Y hesaplanır; sonuç, dağılımı bilinen bir X 'in karakteristik fonksiyonuyla çakışıyorsa Y 'nin dağılımı X 'inkiyle aynıdır.

17.1.4 Karakteristik Fonksiyon Örnekleri

Aşağıdaki hesaplarda, gerçel değişkenli karmaşık değerli $x \mapsto e^{itx}$ fonksiyonunun türevini ve integralini kullanacağız. Gerçel ve sanal kısımlar ayrı ayrı türevlenir:

$$\frac{d}{dx} e^{itx} = \frac{d}{dx} (\cos(tx) + i \sin(tx)) = -t \sin(tx) + it \cos(tx) = it (\cos(tx) + i \sin(tx)) = it e^{itx}.$$

Dolayısıyla $t \neq 0$ için $\int_a^b e^{itx} dx = \left[\frac{e^{itx}}{it} \right]_a^b = \frac{e^{itb} - e^{ita}}{it}$ olur. Ayrıca $e^{iu} - e^{-iu} = 2i \sin u$ ve $e^{iu} + e^{-iu} = 2 \cos u$ özdeşliklerini kullanacağız.

Örnek 17.1 (Düzgün Dağılımın Karakteristik Fonksiyonu). $a < b$ olmak üzere X rastgele değişkeninin yoğunluk fonksiyonu

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun ((a, b) aralığında düzgün dağılım). φ_X 'i bulunuz; özel olarak $a = -1$, $b = 1$ durumunu yazınız ve buradan $E(X)$ ile $\text{Var}(X)$ 'i okuyunuz.

Çözüm.

$t \neq 0$ için

$$\varphi_X(t) = \int_a^b e^{itx} \frac{1}{b-a} dx = \frac{1}{b-a} \left[\frac{e^{itx}}{it} \right]_a^b = \frac{e^{ibt} - e^{iat}}{it(b-a)}.$$

$t = 0$ için formül $\frac{0}{0}$ biçimindedir; ama tanımdan doğrudan $\varphi_X(0) = 1$ 'dir ve Teorem 17.1 (c) gereği φ_X sürekli olduğundan formülün $t \rightarrow 0$ limiti de 1'dir. Özetle

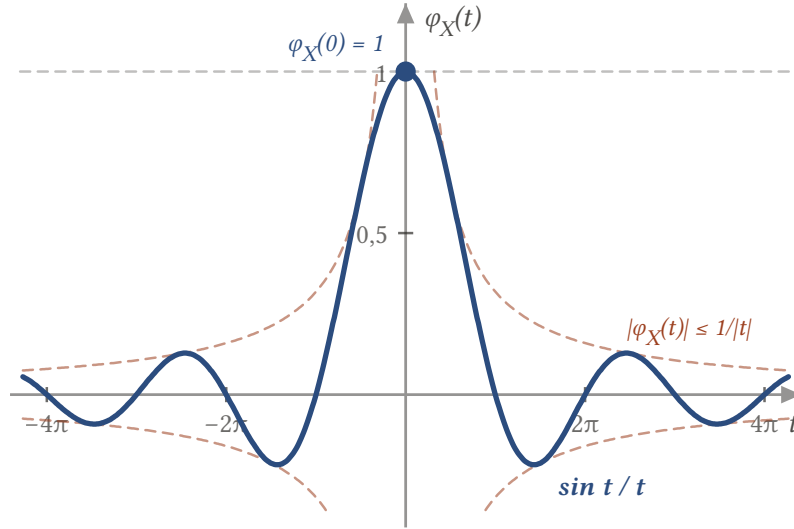
$$\varphi_X(t) = \begin{cases} \frac{e^{ibt} - e^{iat}}{it(b-a)}, & t \neq 0 \\ 1, & t = 0. \end{cases}$$

Aynı sonuca Teorem 17.1 (d) ile de ulaşılabilir: $U, (0, 1)$ aralığında düzgün dağılımlı ise $\varphi_U(t) = \frac{e^{it} - 1}{it}$ ve $(b-a)U + a$ rastgele değişkeni X ile aynı dağılıma sahip olduğundan $\varphi_X(t) = e^{iat} \varphi_U((b-a)t) = e^{iat} \frac{e^{i(b-a)t} - 1}{i(b-a)t} = \frac{e^{ibt} - e^{iat}}{it(b-a)}$.

Özel durum $a = -1, b = 1$. $t \neq 0$ için

$$\varphi_X(t) = \frac{e^{it} - e^{-it}}{2it} = \frac{2i \sin t}{2it} = \frac{\sin t}{t}, \quad \varphi_X(0) = 1.$$

X ile $-X$ aynı dağılıma sahip olduğundan sonucun gerçel çıkması Teorem 17.1 (e) ile uyumludur. Fonksiyon $t = 0$ 'da 1 değerini alır, $|\sin t/t| \leq 1$ 'dir ve $t \rightarrow \pm\infty$ iken sifıra sönerek salınır.



Şekil 17.1: $(-1, 1)$ üzerinde düzgün dağılımın karakteristik fonksiyonu $\varphi_X(t) = \sin t / t$. Sıfırda 1'e eşittir, her yerde mutlak değerce 1'i aşmaz ve sönümlü salınımla sifıra gider; 2π 'nin katlarında sıfırlanır. Kesikli zarflar $\pm 1/t$ sönümün hızını gösterir.

Momentler. $\sin t = t - \frac{t^3}{6} + \frac{t^5}{120} - \dots$ olduğundan

$$\varphi_X(t) = \frac{\sin t}{t} = 1 - \frac{t^2}{6} + \frac{t^4}{120} - \dots$$

Bunu $\varphi_X(t) = 1 + it E(X) - \frac{t^2}{2} E(X^2) - \frac{it^3}{6} E(X^3) + \frac{t^4}{24} E(X^4) - \dots$ açılımıyla karşılaştıralım: t 'nin katsayısı sıfır olduğundan $E(X) = 0$; t^2 'nin katsayısından $-\frac{E(X^2)}{2} = -\frac{1}{6}$, yani $E(X^2) = \frac{1}{3}$ bulunur. Böylece Önerme 16.3 ile $\text{Var}(X) = \frac{1}{3} - 0 = \frac{1}{3}$. Doğrudan hesap da aynı sonucu verir: $\int_{-1}^1 \frac{x^2}{2} dx = \frac{1}{3}$. Aynı yolla t^4 katsayısından $\frac{E(X^4)}{24} = \frac{1}{120}$, yani $E(X^4) = \frac{1}{5}$ okunur.

Örnek 17.2 (İki Noktalı Dağılımın Karakteristik Fonksiyonu). X rastgele değişkeninin olasılık fonksiyonu $P(X = 1) = P(X = -1) = \frac{1}{2}$ ile verilsin. φ_X 'i bulunuz; buradan X 'in momentlerini okuyunuz ve $Y = \frac{X+1}{2}$ değişkeninin karakteristik fonksiyonunu yazınız.

Çözüm.

Kesikli formül ile

$$\varphi_X(t) = \sum_{x \in \{-1, 1\}} e^{itx} f_X(x) = \frac{1}{2}e^{it} + \frac{1}{2}e^{-it} = \frac{e^{it} + e^{-it}}{2} = \cos t.$$

Dağılım 0 çevresinde bakışlı olduğundan sonuç yine gerçeldir.

Momentler. $\cos t = 1 - \frac{t^2}{2} + \frac{t^4}{24} - \dots$ açılımını $1 + it E(X) - \frac{t^2}{2} E(X^2) - \frac{it^3}{6} E(X^3) + \frac{t^4}{24} E(X^4) - \dots$ ile karşılaştırınca $E(X) = 0$, $E(X^2) = 1$, $E(X^3) = 0$, $E(X^4) = 1$ bulunur. Bu, $X^2 = 1$ olmasıyla tutarlıdır: X 'in bütün çift momentleri 1, tek momentleri 0'dır. Türevle de doğrulanabilir: $\varphi_X''(0) = -\cos 0 = -1 = -E(X^2)$.

$Y = \frac{X+1}{2}$ için. Y , 0 ve 1 değerlerini $\frac{1}{2}$ olasılıkla alır. **Teorem 17.1** (d) ile $a = \frac{1}{2}$, $b = \frac{1}{2}$ alınarak

$$\varphi_Y(t) = e^{it/2} \varphi_{\frac{X-1}{2}}\left(\frac{t}{2}\right) = e^{it/2} \cos \frac{t}{2} = e^{it/2} \cdot \frac{e^{it/2} + e^{-it/2}}{2} = \frac{e^{it} + 1}{2}.$$

Doğrudan hesap da aynı sonucu verir: $\varphi_Y(t) = \frac{1}{2}e^{it \cdot 0} + \frac{1}{2}e^{it} = \frac{1+e^{it}}{2}$.

■

Örnek 17.3 (Laplace Dağılımının Karakteristik Fonksiyonu). X rastgele değişkeninin yoğunluk fonksiyonu

$$f_X(x) = \frac{1}{2}e^{-|x|}, \quad x \in \mathbb{R}$$

olsun (Laplace dağılımı). φ_X 'i bulunuz ve buradan $E(X)$, $\text{Var}(X)$ ve $E(X^4)$ 'ü okuyunuz.

Çözüm.

Önce f_X 'in bir yoğunluk olduğunu görelim: $f_X \geq 0$ ve bakışım nedeniyle $\int_{-\infty}^{\infty} \frac{1}{2}e^{-|x|} dx = 2 \int_0^{\infty} \frac{1}{2}e^{-x} dx = 1$.

$|x|$ 'in işaretine göre integrali ikiye ayıralım:

$$\varphi_X(t) = \frac{1}{2} \int_{-\infty}^0 e^{itx} e^x dx + \frac{1}{2} \int_0^{\infty} e^{itx} e^{-x} dx = \frac{1}{2} \int_{-\infty}^0 e^{(1+it)x} dx + \frac{1}{2} \int_0^{\infty} e^{-(1-it)x} dx.$$

İkinci integrali hesaplayalım. $\frac{d}{dx} e^{-(1-it)x} = -(1-it)e^{-(1-it)x}$ olduğundan $R > 0$ için

$$\int_0^R e^{-(1-it)x} dx = \left[\frac{e^{-(1-it)x}}{-(1-it)} \right]_0^R = \frac{1 - e^{-(1-it)R}}{1-it}.$$

$|e^{-(1-it)R}| = e^{-R}$ olduğundan $R \rightarrow \infty$ iken integral $\frac{1}{1-it}$ 'ye yakınsar. Aynı biçimde

$$\int_{-R}^0 e^{(1+it)x} dx = \left[\frac{e^{(1+it)x}}{1+it} \right]_{-R}^0 = \frac{1 - e^{-(1+it)R}}{1+it} \rightarrow \frac{1}{1+it}.$$

Böylece

$$\varphi_X(t) = \frac{1}{2} \left(\frac{1}{1+it} + \frac{1}{1-it} \right) = \frac{1}{2} \cdot \frac{(1-it) + (1+it)}{(1+it)(1-it)} = \frac{1}{2} \cdot \frac{2}{1+t^2} = \frac{1}{1+t^2}, \quad t \in \mathbb{R}.$$

Sonuç gerçeldir; bu yine bakışımın sonucudur. Aslında bakışımdan yararlanarak sanal kısmı hiç hesaplamadan da ilerlenebilirdi: $\sin(tx)f_X(x)$ tek fonksiyon olduğundan $E(\sin(tX)) = 0$ ve $\varphi_X(t) = E(\cos(tX)) = \int_0^{\infty} e^{-x} \cos(tx) dx = \frac{1}{1+t^2}$.

Momentler. $|t| < 1$ için geometrik seri ile

$$\varphi_X(t) = \frac{1}{1+t^2} = 1 - t^2 + t^4 - t^6 + \dots$$

Bunu $1 + it E(X) - \frac{t^2}{2} E(X^2) - \frac{it^3}{6} E(X^3) + \frac{t^4}{24} E(X^4) - \dots$ ile karşılaştıralım: $E(X) = 0$; $-\frac{E(X^2)}{2} = -1$, yani $E(X^2) = 2$; $E(X^3) = 0$; $\frac{E(X^4)}{24} = 1$, yani $E(X^4) = 24$. Dolayısıyla $\text{Var}(X) = E(X^2) - (E(X))^2 = 2$. Doğrudan hesapla karşılaştırma: $E(X^4) = 2 \int_0^\infty \frac{x^4}{2} e^{-x} dx = 4! = 24$.

■

Laplace örneği ters Fourier dönüşümü notunu somutlaştırır: $\varphi_X(t) = \frac{1}{1+t^2}$ mutlak integrallenebilir olduğundan $f_X(x) = \frac{1}{2\pi} \int_{-\infty}^\infty \frac{e^{-itx}}{1+t^2} dt = \frac{1}{2} e^{-|x|}$ olmak zorundadır; yani bu integral, karakteristik fonksiyon kuramıyla hesaplanmış olur.

17.2 Moment Üreten Fonksiyon

Karakteristik fonksiyondaki i sayısı hesapları karmaşık sayılara taşır. e^{itX} yerine gerçel e^{tX} alınırsa karmaşık sayılardan kurtulunur; bedeli, e^{tX} sınırlı olmadığından beklenen değerin her zaman var olmasıdır.

Tanım 17.2 (Moment Üreten Fonksiyon). X bir rastgele değişken olsun. Bir $h > 0$ sayısı için her $t \in (-h, h)$ 'de (E^{tX}) beklenen değeri (sonlu olarak) varsa

$$M_X(t) = (E^{tX}), \quad -h < t < h$$

fonksiyonuna X rastgele değişkeninin **moment üreten fonksiyonu** (moment çıkaran fonksiyonu) denir. X kesikli ise $M_X(t) = \sum_{x \in D_X} e^{tx} f_X(x)$, sürekli ise $M_X(t) = \int_{-\infty}^\infty e^{tx} f_X(x) dx$ olur.

Tanımdaki $(-h, h)$ koşulu önemlidir: moment üreten fonksiyonun var sayılması için $E(e^{tX})$ 'in $t = 0$ 'ın **bir komşuluğundaki bütün** t 'ler için sonlu olması istenir. $t = 0$ 'da $M_X(0) = E(1) = 1$ her zaman sonludur; sorun, $t > 0$ için e^{tx} 'in $x \rightarrow \infty$ iken, $t < 0$ için ise $x \rightarrow -\infty$ iken üstel hızla büyümesidir. Beklenen değerin sonlu kalması için dağılımın her iki kuyruğunun üstel hızla sönmesi gerekir. Aşağıda bunun sağlanmadığı bir örnek göreceğiz.

Moment üreten fonksiyon var olduğunda karakteristik fonksiyonla aradaki bağ, beklendiği gibi, t yerine it koymaktır. Bunu tam olarak ifade edelim; çünkü bu bağ, moment üreten fonksiyon için teklik teoreminin dayanağıdır.

Önerme 17.1 (Moment Üreten Fonksiyon ile Karakteristik Fonksiyon Arasındaki İlişki). X 'in moment üreten fonksiyonu $(-h, h)$ üzerinde var olsun. O zaman $|\text{Re } z| < h$ koşulunu sağlayan her karmaşık $z = s + iu$ için (E^{zX}) vardır; böylece $M_X, \{z \in \mathbb{C} : |\text{Re } z| < h\}$ şeridinde $M_X(z) = (E^{zX})$ ile genişletilir ve her $t \in \mathbb{R}$ için

$$\varphi_X(t) = M_X(it)$$

olur.

İspat.

$z = s + iu$ olsun. $e^{zX} = e^{sX} e^{iuX} = e^{sX} \cos(uX) + i e^{sX} \sin(uX)$ yazılır. $|e^{sX} \cos(uX)| \leq e^{sX}$ ve $|e^{sX} \sin(uX)| \leq e^{sX}$ olduğundan, $|s| < h$ için $E(e^{sX}) = M_X(s) < \infty$ olması gerçel ve sanal kısımların beklenen değerlerinin var olmasını sağlar. Öyleyse $E(e^{zX})$ tanımlıdır. $z = it$ alınırsa $s = 0, u = t$ olur ve $E(e^{itX}) = \varphi_X(t)$ elde edilir.

■

Uygulamada bu, M_X 'in kapalı formülünde t yerine it yazmak demektir. Örneğin aşağıda üstel dağılım için $M_X(t) = \frac{1}{1-t}$ bulacağız; buradan $\varphi_X(t) = \frac{1}{1-it}$ okunur ve bu, Laplace örneğindeki $\int_0^\infty e^{-(1-it)x} dx = \frac{1}{1-it}$ hesabıyla doğrulanır.

17.2.1 Moment Üreten Fonksiyondan Momentler

Adının söylediği gibi, moment üreten fonksiyon momentleri üretir. Mekanizma karakteristik fonksiyondakiyle aynıdır, ama bu kez i çarpanı yoktur ve sonuç daha temizdir.

Teorem 17.4 (Moment Üreten Fonksiyondan Momentler). X 'in moment üreten fonksiyonu $(-h, h)$ üzerinde var olsun. O zaman X 'in her mertebeden momenti vardır, M_X bu aralıkta sonsuz kez türevlenebilir ve

$$M_X^{(n)}(0) = (E(X^n)), \quad n = 1, 2, \dots$$

olur. Ayrıca $|t| < h$ için

$$M_X(t) = \sum_{n=0}^{\infty} \frac{E(X^n)}{n!} t^n$$

kuvvet serisi açılımı geçerlidir.

İspat.

Adım 1: Bütün momentler vardır. $0 < s < h$ sabit olsun. Her gerçel x ve her n için üstel serinin bir terimi seriden büyük olamaz: $\frac{(s|x|)^n}{n!} \leq e^{s|x|}$. Ayrıca $e^{s|x|} \leq e^{sx} + e^{-sx}$ olduğundan

$$|x|^n \leq \frac{n!}{s^n} (e^{sx} + e^{-sx})$$

bulunur. Beklenen değer alınırsa $E(|X|^n) \leq \frac{n!}{s^n} (M_X(s) + M_X(-s)) < \infty$.

Adım 2: Türev. $\frac{d^n}{dt^n} e^{tx} = x^n e^{tx}$ 'dir. $|t| < h$ verilsin ve $|t| \leq s' < s < h$ olacak biçimde s' ile s seçelim. $|x^n e^{tx}| \leq |x|^n e^{s'|x|}$ ve Adım 1'deki eşitsizlik $s - s'$ ile uygulanırsa $|x|^n \leq \frac{n!}{(s-s')^n} e^{(s-s')|x|}$, dolayısıyla

$$|x^n e^{tx}| \leq \frac{n!}{(s-s')^n} e^{s|x|} \leq \frac{n!}{(s-s')^n} (e^{sx} + e^{-sx})$$

olur. Sağ taraftan bağımsızdır ve beklenen değeri sonludur. Bu, $|t| \leq s'$ aralığında türev ile beklenen değerin yerinin (baskın yakınsaklık teoremi gereği) değiştirilebilmesi için yeterlidir:

$$M_X^{(n)}(t) = (E(X^n e^{tX})), \quad |t| < h.$$

$t = 0$ konursa $M_X^{(n)}(0) = E(X^n)$ elde edilir.

Adım 3: Seri açılımı. $e^{tx} = \sum_{n=0}^{\infty} \frac{(tx)^n}{n!}$ serisini terim terim beklenen değere taşıyabilmek için mutlak yakınsaklığı kontrol edelim: $|t| < h$ için

$$\sum_{n=0}^{\infty} \frac{|t|^n E(|X|^n)}{n!} = \left(E \sum_{n=0}^{\infty} \frac{|tX|^n}{n!} \right) = (E|e^{tX}|) \leq M_X(|t|) + M_X(-|t|) < \infty.$$

(Negatif olmayan terimli serilerde toplam ile beklenen değerini her zaman değiştirilebilir.) Mutlak yakınsaklık, toplam ile beklenen değerinin yerinin genel durumda da değiştirilebilmesini sağlar; böylece $M_X(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} E(X^n)$ olur. Bu seri M_X 'in 0 çevresindeki Taylor serisidir ve katsayı karşılaştırması Adım 2'nin sonucunu bir kez daha verir.

■

Serideki t^n 'nin katsayısı $\frac{E(X^n)}{n!}$ olduğundan, M_X 'in kapalı formülü seriye açıldığında momentler doğrudan okunur. Karakteristik fonksiyondan farkı, katsayılarda i^n çarpanının bulunmamasıdır.

17.2.2 Teklik

Moment üreten fonksiyon var olduğunda, karakteristik fonksiyon gibi, dağılımı tek biçimde belirler.

Teorem 17.5 (Moment Üreten Fonksiyon için Teklik Teoremi). X ve Y rastgele değişkenlerinin moment üreten fonksiyonları var ve bir $h > 0$ için her $t \in (-h, h)$ 'de $M_X(t) = M_Y(t)$ ise, X ile Y aynı dağılıma sahiptir: her x için $F_X(x) = F_Y(x)$.

Teoremin arkasındaki fikir Önerme 17.1 ile Teorem 17.3'in birleştirilmesidir: M_X ve M_Y , $|\operatorname{Re} z| < h$ şeridinde karmaşık değişkenin analitik fonksiyonlarıdır; gerçel eksenin bir parçasında çakışan iki analitik

fonksiyon bütün şeritte çakışır, dolayısıyla sanal ekseninde de çakışır ve $\varphi_X(t) = M_X(it) = M_Y(it) = \varphi_Y(t)$ olur. Analitiklik adımının ayrıntısına girmiyoruz; teoremi olduğu gibi kullanacağız. Teoremin bir uyarıyla birlikte okunması gerekir: teklik yalnızca **moment üreten fonksiyonu var olan** rastgele değişkenler için söz konusudur. Moment üreten fonksiyonu olmayan dağılımlar vardır ve bunlar için tek güvenilir araç karakteristik fonksiyondur.

⚠ Her rastgele değişkenin moment üreten fonksiyonu yoktur

X rastgele değişkeninin yoğunluğu $f_X(x) = \frac{1}{\pi(1+x^2)}$, $x \in \mathbb{R}$ olsun (Cauchy dağılımı). $t > 0$ için $x \geq 0$ iken $e^{tx} \geq \frac{(tx)^3}{6}$ olduğundan

$$(E^{tX}) \geq \int_0^\infty \frac{e^{tx}}{\pi(1+x^2)} dx \geq \frac{t^3}{6\pi} \int_0^\infty \frac{x^3}{1+x^2} dx = \infty;$$

çünkü $x \geq 1$ için $\frac{x^3}{1+x^2} \geq \frac{x}{2}$ 'dir. $t < 0$ için aynı şey sol kuyrukta olur. Yani $E(e^{tX})$ yalnızca $t = 0$ 'da sonludur ve X 'in moment üreten fonksiyonu **yoktur**; zaten $E|X| = \infty$ olduğundan beklenen değeri bile yoktur. Buna karşın karakteristik fonksiyonu, her rastgele değişkende olduğu gibi vardır ve $\varphi_X(t) = e^{-|t|}$ 'dir. (Bu, Laplace örneğinin ters Fourier dönüşümüyle okunmuş biçimdir: $\int_{-\infty}^\infty \frac{e^{-itx}}{1+t^2} dt = \pi e^{-|x|}$ eşitliğinde x ile t 'nin rolleri değiştirilirse $\int_{-\infty}^\infty \frac{e^{-itx}}{1+x^2} dx = \pi e^{-|t|}$ olur; $\frac{1}{1+x^2}$ çift bir fonksiyon olduğundan $x \mapsto -x$ değişken değiştirmesiyle üsteldeki işaret çevrilebilir ve $\varphi_X(t) = \frac{1}{\pi} \int_{-\infty}^\infty \frac{e^{itx}}{1+x^2} dx = e^{-|t|}$ elde edilir.) Moment üreten fonksiyon için teklik teoremi Cauchy dağılımını bu yüzden kapsamaz: Cauchy dağılımının moment üreten fonksiyonu tanımlı değildir; dağılımını belirleyen tek araç karakteristik fonksiyondur.

17.2.3 Moment Üreten Fonksiyon Örnekleri

Örnek 17.4 (Üstel Dağılımın Moment Üreten Fonksiyonu). X rastgele değişkeninin yoğunluk fonksiyonu

$$f_X(x) = \begin{cases} e^{-x}, & x > 0 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun (üstel dağılım). M_X 'in var olduğunu gösteriniz, bulunuz ve buradan bütün momentleri, özel olarak $E(X)$ ile $\text{Var}(X)$ 'i hesaplayınız.

Çözüm.

$t \in \mathbb{R}$ için

$$(E^{tX}) = \int_0^\infty e^{tx} e^{-x} dx = \int_0^\infty e^{-(1-t)x} dx.$$

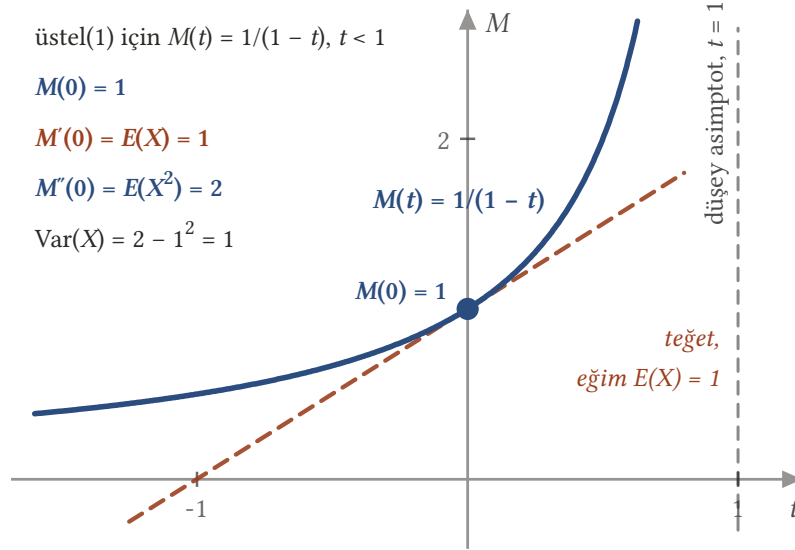
$t < 1$ ise $1-t > 0$ ve integral $\left[\frac{e^{-(1-t)x}}{-(1-t)} \right]_0^\infty = \frac{1}{1-t}$ olur. $t \geq 1$ ise integrand her $x > 0$ için $e^{-(1-t)x} \geq 1$ olduğundan integral sonsuzdur. Demek ki $h = 1$ ile tanımın koşulu sağlanır ve

$$M_X(t) = \frac{1}{1-t} = (1-t)^{-1}, \quad t < 1.$$

Momentler. $M_X'(t) = (1-t)^{-2}$, $M_X''(t) = 2(1-t)^{-3}$ ve tümevarımla $M_X^{(n)}(t) = n!(1-t)^{-(n+1)}$ olur. **Teorem 17.4** ile

$$E(X^n) = M_X^{(n)}(0) = n!, \quad n = 1, 2, \dots$$

Özel olarak $E(X) = 1$, $E(X^2) = 2$ ve $\text{Var}(X) = 2 - 1^2 = 1$. Seri açılımıyla da görülür: $|t| < 1$ için $\frac{1}{1-t} = \sum_{n=0}^\infty t^n = \sum_{n=0}^\infty \frac{n!}{n!} t^n$, yani t^n 'nin katsayısı $\frac{E(X^n)}{n!} = 1$ 'dir.



Şekil 17.2: Üstel dağılımın moment üreten fonksiyonu $M(t) = 1/(1 - t)$, yalnız $t < 1$ için tanımlıdır ve $t = 1$ 'de sonsuza gider. Her moment üreten fonksiyon gibi $M(0) = 1$ 'dir; $t = 0$ 'daki teğetin eğimi $M'(0) = E(X) = 1$, eğrilik ise $M''(0) = E(X^2) = 2$ verir. Momentler, tek bir fonksiyonun sıfırdaki türevlerinden okunur.

Önerme 17.1 gereği karakteristik fonksiyon $\varphi_X(t) = M_X(it) = \frac{1}{1-it}$ 'dir.

Örnek 17.5 (Poisson Dağılımının Moment Üreten Fonksiyonu). $\lambda > 0$ olmak üzere X rastgele değişkeninin olasılık fonksiyonu

$$f_X(x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, 2, \dots$$

olsun (Poisson dağılımı). M_X 'i bulunuz; $E(X)$, $E(X^2)$ ve $\text{Var}(X)$ 'i hesaplayınız.

Çözüm.

Önce f_X 'in bir olasılık fonksiyonu olduğunu not edelim: $\sum_{x=0}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^{-\lambda} e^{\lambda} = 1$. Her $t \in \mathbb{R}$ için üstel seri yardımıyla

$$M_X(t) = \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda} \lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} = e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t-1)}.$$

Seri her t için yakınsadığından moment üreten fonksiyon bütün $\mathbb{R}$ üzerinde vardır (h istenildiği kadar büyük seçilebilir).

Momentler. Zincir kuralıyla

$$M_X'(t) = \lambda e^t e^{\lambda(e^t-1)}, \quad M_X''(t) = \lambda e^t e^{\lambda(e^t-1)} + (\lambda e^t)^2 e^{\lambda(e^t-1)}.$$

$t = 0$ 'da $e^{\lambda(e^0-1)} = e^0 = 1$ olduğundan

$$E(X) = M_X'(0) = \lambda, \quad E(X^2) = M_X''(0) = \lambda + \lambda^2.$$

Dolayısıyla $\text{Var}(X) = E(X^2) - (E(X))^2 = \lambda + \lambda^2 - \lambda^2 = \lambda$: Poisson dağılımında beklenen değer ile varyans eşittir. Karakteristik fonksiyon ise $\varphi_X(t) = M_X(it) = e^{\lambda(e^{it}-1)}$ 'dir.

Örnek 17.6 (Moment Üreten Fonksiyondan Dağılımı Okumak). Aşağıdaki her durumda X rastgele değişkeninin dağılımını (olasılık ya da yoğunluk fonksiyonunu) bulunuz.

- (a) $M_X(t) = e^{e^t-1}$, $t \in \mathbb{R}$.
 (b) $M_X(t) = \frac{1}{4} + \frac{1}{2}e^t + \frac{1}{4}e^{3t}$, $t \in \mathbb{R}$.
 (c) $M_X(t) = \frac{2}{2-t}$, $t < 2$.

Çözüm.

Her üç şıkta da yöntem aynıdır: verilen fonksiyonu moment üreten fonksiyonu bilinen bir dağılımla eşleştirmek ve Teorem 17.5'i uygulamak.

(a) Örnek 17.5'de λ parametrelili Poisson dağılımının moment üreten fonksiyonunu $e^{\lambda(e^t-1)}$ bulduk. $\lambda = 1$ için bu tam olarak e^{e^t-1} 'dir. Teklik teoremi gereği X , $\lambda = 1$ parametrelili Poisson dağılımına sahiptir:

$$f_X(x) = \frac{e^{-1}}{x!}, \quad x = 0, 1, 2, \dots$$

(b) Kesikli bir Y rastgele değişkeninin moment üreten fonksiyonu $\sum_y e^{ty} f_Y(y)$ biçimindedir; yani e^{ty} teriminin katsayısı $P(Y = y)$ 'dir. Verilen fonksiyonu $\frac{1}{4}e^{t \cdot 0} + \frac{1}{2}e^{t \cdot 1} + \frac{1}{4}e^{t \cdot 3}$ olarak okursak, olasılık fonksiyonu

y	0	1	3
$f_Y(y)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

olan Y için $M_Y = M_X$ olur (katsayılar negatif değildir ve toplamı 1'dir; dolayısıyla bu gerçekten bir olasılık fonksiyonudur). Teklik teoremi gereği X de bu dağılıma sahiptir: $P(X = 0) = \frac{1}{4}$, $P(X = 1) = \frac{1}{2}$, $P(X = 3) = \frac{1}{4}$.

(c) $\frac{2}{2-t} = \frac{1}{1-t/2}$ yazalım. Örnek 17.4'de yoğunluğu e^{-z} ($z > 0$) olan değişkenin moment üreten fonksiyonunu $\frac{1}{1-t}$ bulmuştuk; bu değişkene burada Z diyelim, yani $M_Z(t) = \frac{1}{1-t}$. $W = \frac{Z}{2}$ alınırsa $M_W(t) = (E^{tZ/2}) = M(\frac{t}{2}) = \frac{1}{1-t/2}$, $t < 2$ olur; bu verilen fonksiyondur. W 'nin yoğunluğu, $g(z) = z/2$ dönüşümü için Teorem 12.2 ile $f_W(w) = f_Z(2w) \cdot 2 = 2e^{-2w}$, $w > 0$ 'dır. Teklik teoremi gereği

$$f_X(x) = \begin{cases} 2e^{-2x}, & x > 0 \\ 0, & \text{diğer durumlarda.} \end{cases}$$

Doğrulama: $\int_0^\infty e^{tx} 2e^{-2x} dx = 2 \int_0^\infty e^{-(2-t)x} dx = \frac{2}{2-t}$, $t < 2$.

■

17.3 Çarpımsal Moment Üreten Fonksiyon

Beklenen Değer, Momentler ve Varyans bölümünde X 'in k . çarpımsal momentini $E[X(X-1)\cdots(X-k+1)]$ olarak tanımlamıştık (Tanım 16.2). Değer kümesi $\{0, 1, 2, \dots\}$ olan rastgele değişkenlerde bu momentler çoğu zaman sıradan momentlerden daha kolay hesaplanır; onları üreten fonksiyon e^{tX} yerine t^X 'i kullanır.

Tanım 17.3 (Çarpımsal Moment Üreten Fonksiyon). X bir rastgele değişken olsun. (E^X) beklenen değerinin (sonlu olarak) var olduğu t değerleri kümesinde tanımlı

$$N_X(t) = (E^X)$$

fonksiyonuna X rastgele değişkeninin **çarpımsal moment üreten fonksiyonu** denir. $t > 0$ için $t^X = e^{X \ln t}$ anlamındadır; X 'in değer kümesi $\{0, 1, 2, \dots\}$ içinde ise t^X her gerçel t için anlamlıdır ($0^0 = 1$ uzlaşımıyla) ve

$$N_X(t) = \sum_{x=0}^{\infty} t^x f_X(x)$$

olur.

Tanımla ilgili üç gözlem:

- $N_X(1) = E(1) = 1$ 'dir; $t = 1$ noktası, moment üreten fonksiyondaki $t = 0$ noktasının rolünü oynar.
- $t > 0$ için $t^X = e^{(\ln t)^X}$ olduğundan $N_X(t) = M_X(\ln t)$, eşdeğer olarak $M_X(t) = N_X(e^t)$ 'dir. Böylece $M_X, (-h, h)$ üzerinde varsa N_X en azından (e^{-h}, e^h) aralığında vardır; iki fonksiyon aynı bilgiyi taşır ve Teorem 17.5 teklik sonucunu N_X 'e de aktarır.
- X 'in değer kümesi $\{0, 1, 2, \dots\}$ içindeyse $N_X(t) = \sum_x t^x f_X(x)$ bir kuvvet serisidir ve $\sum_x f_X(x) = 1$ olduğundan en azından $|t| \leq 1$ için mutlak yakınsar. Katsayılar olasılıkların kendisidir; bu yüzden N_X 'e **olasılık üreten fonksiyon** da denir. Bir kuvvet serisinin katsayıları, serinin topladığı fonksiyon tarafından tek biçimde belirlendiğinden, N_X bu durumda dağılımı analitiklik tartışmasına gerek kalmadan belirler: $f_X(x) = \frac{N_X^{(x)}(0)}{x!}$.

Çarpımsal moment üreten fonksiyonun türevleri, $t = 1$ 'de, çarpımsal momentleri verir. Formül doğrudan t^x 'in türevinin biçiminden çıkar: $\frac{d^n}{dt^n} t^x = x(x-1)\cdots(x-n+1)t^{x-n}$.

Teorem 17.6 (Çarpımsal Moment Üreten Fonksiyondan Çarpımsal Momentler). X 'in çarpımsal moment üreten fonksiyonu $t = 1$ 'i içeren bir açık aralıkta tanımlı olsun ve $k = 1, 2, \dots, n$ için bu aralıkta türev ile beklenen değerini yer değiştirilebilsin:

$$\frac{d^k}{dt^k}(E^X) = \left(E \frac{d^k}{dt^k} t^X \right).$$

O zaman

$$N_X^{(n)}(1) = E [X(X-1)\cdots(X-n+1)],$$

yani N_X 'in $t = 1$ 'deki n . türevi X 'in n . çarpımsal momentidir. Özel olarak

$$E(X) = N_X'(1), \quad (E^2) = N_X''(1) + N_X'(1), \quad \text{Var}(X) = N_X''(1) + N_X'(1) - (N_X'(1))^2.$$

İspat.

$t > 0$ sabit ve x gerçel olsun. $t^x = e^{x \ln t}$ 'nin t 'ye göre türevi $x t^{x-1}$, ikinci türevi $x(x-1) t^{x-2}$ ve tümevarımla n . türevi $x(x-1)\cdots(x-n+1) t^{x-n}$ 'dir. (x negatif olmayan bir tam sayı ve $x < n$ ise çarpımda bir sıfır çarpan bulunur ve türev sıfırdır; bu, t^x 'in derecesi x olan bir polinom olmasıyla tutarlıdır.) Varsayım gereği türev beklenen değerini içine alınabilir:

$$N_X^{(n)}(t) = \left(E \frac{d^n}{dt^n} t^X \right) = E [X(X-1)\cdots(X-n+1) t^{X-n}].$$

$t = 1$ konursa $t^{X-n} = 1$ olur ve ilk iddia çıkar. $E(X) = N_X'(1)$ bunun $n = 1$ halidir. $X^2 = X(X-1) + X$ özdeşliğinden ve beklenen değer doğrusallığından $E(X^2) = E[X(X-1)] + E(X) = N_X''(1) + N_X'(1)$; varyans formülü Önerme 16.3 ile tamamlanır.

■

Türev ile beklenen değerini yer değiştirme varsayımı, sık karşılaşılan iki durumda kendiliğinden sağlanır. X 'in değer kümesi $\{0, 1, 2, \dots\}$ içindeyse ve N_X bir $\delta > 0$ için $|t| < 1 + \delta$ aralığında yakınsıyorsa, N_X bir kuvvet serisidir ve yakınsaklık aralığının içinde terim terim türevlenebilir; varsayım sağlanır. N_X, M_X üzerinden $(N_X(t) = M_X(\ln t))$ elde ediliyorsa Teorem 17.4'in ispatındaki baskın yakınsaklık tartışması aynen geçerlidir.

Örnek 17.7 (Poisson Dağılımının Çarpımsal Moment Üreten Fonksiyonu). $\lambda > 0$ olmak üzere X 'in olasılık fonksiyonu $f_X(x) = \frac{e^{-\lambda} \lambda^x}{x!}$, $x = 0, 1, 2, \dots$ olsun. N_X 'i bulunuz ve bütün çarpımsal momentleri hesaplayınız.

Çözüm.

Her $t \in \mathbb{R}$ için

$$N_X(t) = \sum_{x=0}^{\infty} t^x \frac{e^{-\lambda} \lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda t)^x}{x!} = e^{-\lambda} e^{\lambda t} = e^{\lambda(t-1)}.$$

Seri bütün $\mathbb{R}'$ 'de yakınsadığından N_X her yerde tanımlıdır ve terim terim türevlenebilir. Beklendiği gibi $N_X(t) = M_X(\ln t) = e^{\lambda(e^{\ln t}-1)} = e^{\lambda(t-1)}$ 'dir.

Çarpımsal momentler. $N_X^{(n)}(t) = \lambda^n e^{\lambda(t-1)}$ olduğundan [Teorem 17.6](#) ile

$$E[X(X-1)\cdots(X-n+1)] = N_X^{(n)}(1) = \lambda^n, \quad n = 1, 2, \dots$$

Özel olarak $E(X) = \lambda$, $E[X(X-1)] = \lambda^2$; buradan $E(X^2) = \lambda^2 + \lambda$ ve $\text{Var}(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$. Bu, [Örnek 17.5](#) ile uyuyor; fakat çarpımsal momentlerin ne kadar temiz çıktığına dikkat ediniz: Poisson dağılımında n . çarpımsal moment λ^n 'dir, oysa sıradan momentler n büyüdükçe karmaşıklaşır. Son olarak olasılık üreten fonksiyon bakış açısını doğrulayalım: $N_X^{(x)}(0) = \lambda^x e^{-\lambda}$ olduğundan $\frac{N_X^{(x)}(0)}{x!} = \frac{e^{-\lambda} \lambda^x}{x!} = f_X(x)$.

■

Örnek 17.8 (Üstel Dağılımın Çarpımsal Moment Üreten Fonksiyonu). X 'in yoğunluk fonksiyonu $f_X(x) = e^{-x}$ ($x > 0$), diğer durumlarda 0 olsun. N_X 'i tanım kümesiyle birlikte bulunuz; $E(X)$ ve $E[X(X-1)]$ 'i hesaplayınız.

Çözüm.

$t > 0$ için $t^x = e^{x \ln t}$ olduğundan

$$N_X(t) = \int_0^\infty t^x e^{-x} dx = \int_0^\infty e^{x \ln t} e^{-x} dx = \int_0^\infty e^{(\ln t - 1)x} dx.$$

Bu integral ancak $\ln t - 1 < 0$, yani $0 < t < e$ iken sonludur ve o zaman

$$N_X(t) = \left[\frac{e^{(\ln t - 1)x}}{\ln t - 1} \right]_0^\infty = \frac{0 - 1}{\ln t - 1} = \frac{1}{1 - \ln t}, \quad 0 < t < e.$$

$t \geq e$ için integrand 1'den küçük olmadığından integral sonsuzdur; demek ki N_X 'in tanım kümesi $(0, e)$ aralığıdır. Bu aralık $t = 1$ 'i içerir. Sonuç $N_X(t) = M_X(\ln t) = \frac{1}{1 - \ln t}$ ilişkisiyle de uyumludur ([Örnek 17.4](#)).

Türevler. Zincir kuralıyla

$$N_X'(t) = \frac{d}{dt}(1 - \ln t)^{-1} = (1 - \ln t)^{-2} \cdot \frac{1}{t} = \frac{1}{t(1 - \ln t)^2}.$$

$t = 1$ 'de $\ln 1 = 0$ olduğundan $E(X) = N_X'(1) = \frac{1}{1 \cdot 1} = 1$.

İkinci türev için $N_X'(t) = t^{-1}(1 - \ln t)^{-2}$ çarpımını türevleyelim:

$$N_X''(t) = -t^{-2}(1 - \ln t)^{-2} + t^{-1} \cdot 2(1 - \ln t)^{-3} \cdot \frac{1}{t} = \frac{2(1 - \ln t)^{-3} - (1 - \ln t)^{-2}}{t^2}.$$

$t = 1$ 'de $E[X(X-1)] = N_X''(1) = 2 - 1 = 1$ bulunur. Buradan $E(X^2) = E[X(X-1)] + E(X) = 2$ ve $\text{Var}(X) = 2 - 1 = 1$; sonuçlar [Örnek 17.4](#) ile aynıdır.

■

17.4 Bir Dönüşümün Karakteristik Fonksiyonu

$Y = g(X)$ biçiminde bir rastgele değişkenin karakteristik fonksiyonu için önce Y 'nin dağılımını bulmak gerekmez. $e^{itY} = e^{itg(X)}$, X 'in bir fonksiyonu olduğundan [Tanım 16.1](#), $\cos(tg(x))$ ve $\sin(tg(x))$ fonksiyonlarına ayrı ayrı uygulanır:

$$\varphi_Y(t) = (E^{itg(X)}) = \int_{-\infty}^{\infty} e^{itg(x)} f_X(x) dx \quad (\text{sürekli durum}), \quad \varphi_Y(t) = \sum_{x \in D_X} e^{itg(x)} f_X(x) \quad (\text{kesikli durum}).$$

$g(x) = ax + b$ için bu, [Teorem 17.1](#) (d)'yi verir. Doğrusal olmayan bir dönüşümde ise integrali g 'nin davranışına göre parçalara ayırmak gerekir.

Örnek 17.9 (Mutlak Değerin Karakteristik Fonksiyonu). X rastgele değişkeninin yoğunluk fonksiyonu

$$f_X(x) = \begin{cases} \frac{1}{3}, & -1 \leq x \leq 2 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun. $Y = |X|$ rastgele değişkeninin karakteristik fonksiyonunu bulunuz ve buradan $E(Y)$ ile $E(Y^2)$ 'yi okuyunuz.

Çözüm.

Doğrudan hesap. $|x|$, $[-1, 0]$ üzerinde $-x$, $[0, 2]$ üzerinde x 'e eşittir. $t \neq 0$ için

$$\varphi_Y(t) = \int_{-1}^2 e^{it|x|} \frac{1}{3} dx = \frac{1}{3} \int_{-1}^0 e^{-itx} dx + \frac{1}{3} \int_0^2 e^{itx} dx.$$

İlk integral: $\frac{d}{dx} e^{-itx} = -it e^{-itx}$ olduğundan

$$\frac{1}{3} \int_{-1}^0 e^{-itx} dx = \frac{1}{3} \left[\frac{e^{-itx}}{-it} \right]_{-1}^0 = \frac{1}{3} \cdot \frac{1 - e^{it}}{-it} = \frac{e^{it} - 1}{3it}.$$

İkinci integral:

$$\frac{1}{3} \int_0^2 e^{itx} dx = \frac{1}{3} \left[\frac{e^{itx}}{it} \right]_0^2 = \frac{e^{2it} - 1}{3it}.$$

Toplayarak

$$\varphi_Y(t) = \frac{e^{it} - 1 + e^{2it} - 1}{3it} = \frac{e^{2it} + e^{it} - 2}{3it}, \quad t \neq 0; \quad \varphi_Y(0) = 1.$$

Yoğunluk üzerinden doğrulama. Y 'nin dağılımını bulup aynı sonuca ulaşalım. $0 \leq y \leq 1$ için $F_Y(y) = P(-y \leq X \leq y) = \frac{2y}{3}$; $1 < y \leq 2$ için $F_Y(y) = P(-1 \leq X \leq y) = \frac{y+1}{3}$. Türev alınırsa

$$f_Y(y) = \begin{cases} \frac{2}{3}, & 0 \leq y \leq 1 \\ \frac{1}{3}, & 1 < y \leq 2 \\ 0, & \text{diğer durumlarda.} \end{cases}$$

$(\int_{-\infty}^{\infty} f_Y(y) dy = \frac{2}{3} + \frac{1}{3} = 1.)$ Buradan

$$\varphi_Y(t) = \frac{2}{3} \int_0^1 e^{ity} dy + \frac{1}{3} \int_1^2 e^{ity} dy = \frac{2(e^{it} - 1)}{3it} + \frac{e^{2it} - e^{it}}{3it} = \frac{e^{2it} + e^{it} - 2}{3it}.$$

Momentler. Payı seriye açalım: $e^{2it} = 1 + 2it - 2t^2 - \frac{4it^3}{3} + \dots$ ve $e^{it} = 1 + it - \frac{t^2}{2} - \frac{it^3}{6} + \dots$ olduğundan

$$e^{2it} + e^{it} - 2 = 3it - \frac{5t^2}{2} - \frac{3it^3}{2} + \dots$$

$3it$ 'ye bölünce

$$\varphi_Y(t) = 1 + \frac{5}{6} it - \frac{1}{2} t^2 + \dots$$

$1 + it E(Y) - \frac{t^2}{2} E(Y^2) + \dots$ ile karşılaştırılırsa $E(Y) = \frac{5}{6}$ ve $E(Y^2) = 1$ bulunur. $Y = |X|$ ve $Y^2 = X^2$ olduğundan doğrudan hesap bunu doğrular: $E(Y) = E|X| = \int_{-1}^2 \frac{|x|}{3} dx = \frac{1}{3}(\frac{1}{2} + 2) = \frac{5}{6}$ ve $E(Y^2) = E(X^2) = \int_{-1}^2 \frac{x^2}{3} dx = \frac{8+1}{9} = 1$.

■

17.5 Alıştırmalar

Alıştırma 17.1 (Karakteristik, Moment Üreten ve Çarpımsal Moment Üreten Fonksiyonlar). (a) X 'in yoğunluk fonksiyonu $f_X(x) = 2x$ ($0 < x < 1$), diğer durumlarda 0 olsun. φ_X 'i bulunuz ve buradan $E(X)$ ile $\text{Var}(X)$ 'i okuyunuz.

(b) $P(X=0) = P(X=1) = \frac{1}{2}$ olsun. φ_X 'i bulunuz; $Y = 2X - 1$ için φ_Y 'yi Teorem 17.1 (d) ile hesaplayıp Örnek 17.2 ile karşılaştırınız.

(c) $\varphi_X(t) = \frac{\sin t}{t}$ olsun. $Z = 3X + 2$ rastgele değişkeninin karakteristik fonksiyonunu bulunuz ve Z 'nin dağılımını belirleyiniz.

(d) X 'in olasılık fonksiyonu $f_X(x) = \left(\frac{1}{2}\right)^x$, $x = 1, 2, 3, \dots$ olsun. M_X ve N_X 'i tanım kümeleriyle birlikte bulunuz; $E(X)$ ve $\text{Var}(X)$ 'i her ikisiyle de hesaplayınız.

(e) $M_X(t) = \frac{1}{3}e^{-t} + \frac{1}{2} + \frac{1}{6}e^{2t}$ olsun. X 'in dağılımını, $E(X)$ 'i ve $\text{Var}(X)$ 'i bulunuz.

(f) X 'in yoğunluk fonksiyonu $f_X(x) = xe^{-x}$ ($x > 0$), diğer durumlarda 0 olsun. M_X 'i bulunuz; $E(X)$ ve $\text{Var}(X)$ 'i hesaplayınız.

(g) X , $(-1, 1)$ aralığında düzgün dağılımlı olsun. $Y = |X|$ 'in karakteristik fonksiyonunu bulunuz ve Y 'nin dağılımını belirleyiniz.

Çözüm.

(a) $t \neq 0$ için kısmi integrasyonla ($u = 2x$, $dv = e^{itx} dx$, $v = \frac{e^{itx}}{it}$):

$$\varphi_X(t) = \int_0^1 2x e^{itx} dx = \left[\frac{2x e^{itx}}{it} \right]_0^1 - \int_0^1 \frac{2e^{itx}}{it} dx = \frac{2e^{it}}{it} - \frac{2}{it} \cdot \frac{e^{it} - 1}{it} = \frac{2e^{it}}{it} + \frac{2(e^{it} - 1)}{t^2},$$

çünkü $(it)^2 = -t^2$. $\frac{1}{i} = -i$ olduğundan

$$\varphi_X(t) = \frac{2(e^{it} - 1)}{t^2} - \frac{2ie^{it}}{t}, \quad t \neq 0; \quad \varphi_X(0) = 1.$$

Momentler için seriye açalım. $e^{it} - 1 = it - \frac{t^2}{2} - \frac{it^3}{6} + \frac{t^4}{24} - \dots$ olduğundan

$$\frac{2(e^{it} - 1)}{t^2} = \frac{2i}{t} - 1 - \frac{it}{3} + \frac{t^2}{12} - \dots, \quad -\frac{2ie^{it}}{t} = -\frac{2i}{t} + 2 + it - \frac{t^2}{3} + \dots$$

Toplam: $\varphi_X(t) = 1 + \frac{2}{3}it - \frac{1}{4}t^2 + \dots$. Buradan $E(X) = \frac{2}{3}$, $-\frac{E(X^2)}{2} = -\frac{1}{4}$, yani $E(X^2) = \frac{1}{2}$ ve $\text{Var}(X) = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}$. Doğrudan hesap: $E(X) = \int_0^1 2x^2 dx = \frac{2}{3}$, $E(X^2) = \int_0^1 2x^3 dx = \frac{1}{2}$.

(b) $\varphi_X(t) = \frac{1}{2}e^{it \cdot 0} + \frac{1}{2}e^{it} = \frac{1+e^{it}}{2}$. $Y = 2X - 1$ için $a = 2$, $b = -1$ ile

$$\varphi_Y(t) = e^{-it} \varphi_X(2t) = e^{-it} \cdot \frac{1+e^{2it}}{2} = \frac{e^{-it} + e^{it}}{2} = \cos t.$$

Y , -1 ve 1 değerlerini $\frac{1}{2}$ olasılıkla alır; Örnek 17.2'te bu dağılım için $\cos t$ bulmuştuk. Sonuçlar uyuyor.

(c) $a = 3$, $b = 2$ ile $\varphi_Z(t) = e^{2it} \varphi_X(3t) = e^{2it} \frac{\sin 3t}{3t}$. Örnek 17.1 gereği X , $(-1, 1)$ aralığında düzgün dağılımlıdır (teklik teoremi); $Z = 3X + 2$ ise $(-1, 5)$ aralığında düzgün dağılımlıdır. Doğrulama: $(a, b) = (-1, 5)$ için düzgün dağılımın karakteristik fonksiyonu

$$\frac{e^{5it} - e^{-it}}{6it} = e^{2it} \cdot \frac{e^{3it} - e^{-3it}}{6it} = e^{2it} \cdot \frac{2i \sin 3t}{6it} = e^{2it} \frac{\sin 3t}{3t}.$$

(d) Önce $\sum_{x=1}^{\infty} \left(\frac{1}{2}\right)^x = 1$ olduğunu not edelim. Geometrik seri ile, $\frac{e^t}{2} < 1$ yani $t < \ln 2$ için

$$M_X(t) = \sum_{x=1}^{\infty} e^{tx} \left(\frac{1}{2}\right)^x = \sum_{x=1}^{\infty} \left(\frac{e^t}{2}\right)^x = \frac{e^t/2}{1 - e^t/2} = \frac{e^t}{2 - e^t}, \quad t < \ln 2;$$

$t \geq \ln 2$ için seri iraksar. Benzer biçimde $|t| < 2$ için

$$N_X(t) = \sum_{x=1}^{\infty} t^x \left(\frac{1}{2}\right)^x = \frac{t/2}{1 - t/2} = \frac{t}{2 - t}, \quad |t| < 2.$$

($N_X(t) = M_X(\ln t)$ olduğu görülür: $\frac{e^{\ln t}}{2 - e^{\ln t}} = \frac{t}{2 - t}$.)

M_X ile: bölüm kuralıyla $M'_X(t) = \frac{e^t(2-e^t)+e^t \cdot e^t}{(2-e^t)^2} = \frac{2e^t}{(2-e^t)^2}$, dolayısıyla $E(X) = M'_X(0) = 2$. İkinci türev: $M''_X(t) = \frac{2e^t(2-e^t)^2+2e^t \cdot 2(2-e^t)e^t}{(2-e^t)^4} = \frac{2e^t(2+e^t)}{(2-e^t)^3}$, dolayısıyla $E(X^2) = M''_X(0) = \frac{2 \cdot 3}{1} = 6$ ve $\text{Var}(X) = 6 - 4 = 2$.

N_X ile: $N'_X(t) = \frac{(2-t)+t}{(2-t)^2} = \frac{2}{(2-t)^2}$ ve $N''_X(t) = \frac{4}{(2-t)^3}$. Teorem 17.6 ile $E(X) = N'_X(1) = 2$, $E[X(X-1)] = N''_X(1) = 4$, $E(X^2) = 4 + 2 = 6$, $\text{Var}(X) = 6 - 4 = 2$. İki yol aynı sonucu verir.

(e) Örnek 17.6 (b)'deki gibi, e^{tx} terimlerinin katsayıları olasılıklardır:

x	-1	0	2
$f_X(x)$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$

Katsayılar negatif değildir ve toplamları 1'dir; teklik teoremi gereği X 'in dağılımı budur. $E(X) = -\frac{1}{3} + 0 + \frac{2}{6} = 0$ ve $E(X^2) = \frac{1}{3} + 0 + \frac{4}{6} = 1$, dolayısıyla $\text{Var}(X) = 1$. Türevle doğrulama: $M'_X(0) = -\frac{1}{3} + \frac{2}{6} = 0$ ve $M''_X(0) = \frac{1}{3} + \frac{4}{6} = 1$.

(f) $\int_0^\infty x e^{-cx} dx = 1$ olduğundan f_X bir yoğunluktur. $t < 1$ için $c = 1 - t > 0$ diyelim; kısmi integrasyonla

$$M_X(t) = \int_0^\infty x e^{-cx} dx = \left[-\frac{x e^{-cx}}{c} \right]_0^\infty + \frac{1}{c} \int_0^\infty e^{-cx} dx = 0 + \frac{1}{c^2} = \frac{1}{(1-t)^2}, \quad t < 1.$$

$t \geq 1$ için integral ıraksar. $M'_X(t) = 2(1-t)^{-3}$ ve $M''_X(t) = 6(1-t)^{-4}$ olduğundan $E(X) = 2$, $E(X^2) = 6$ ve $\text{Var}(X) = 6 - 4 = 2$.

(g) $f_X(x) = \frac{1}{2} (-1 < x < 1)$. $t \neq 0$ için, Örnek 17.9'teki gibi integrali ikiye bölerek

$$\varphi_Y(t) = \frac{1}{2} \int_{-1}^0 e^{-itx} dx + \frac{1}{2} \int_0^1 e^{itx} dx = \frac{1}{2} \cdot \frac{e^{it} - 1}{it} + \frac{1}{2} \cdot \frac{e^{-it} - 1}{-it} = \frac{e^{it} - 1}{it}.$$

Örnek 17.1'e göre bu, $(0, 1)$ aralığında düzgün dağılımın karakteristik fonksiyonudur ($a = 0$, $b = 1$). Teorem 17.3 gereği $Y = |X|$, $(0, 1)$ aralığında düzgün dağılımlıdır. Doğrudan doğrulama: $0 \leq y \leq 1$ için $P(|X| \leq y) = P(-y \leq X \leq y) = \frac{2y}{2} = y$.

■

Bu bölümde bir rastgele değişkenin bütün momentlerini ve dağılımını tek bir fonksiyona sıkıştırmayı öğrendik. Bu araçların asıl gücü, bağımsız rastgele değişkenlerin toplamlarında ortaya çıkar: toplamın karakteristik fonksiyonu, toplananların karakteristik fonksiyonlarının çarpımıdır. Bunu görebilmek için önce birden çok rastgele değişkenin beklenen değerlerini, çarpımlarının beklenen değerini ve aralarındaki doğrusal bağı ölçen kovaryans ile korelasyonu incelemek gerekir: [Rastgele Vektörlerde Beklenen Değer, Kovaryans ve Korelasyon](#).

18. Rastgele Vektörlerde Beklenen Değer, Kovaryans ve Korelasyon

Karakteristik ve Moment Üreten Fonksiyonlar bölümü de içinde olmak üzere, şimdiye kadar beklenen değeri hep **tek** bir rastgele değişken için inceledik: $E(X)$, $E(X^2)$, $\text{Var}(X)$, $M_X(t)$. Oysa uygulamada rastgele değişkenler çoğunlukla birlikte gelir: bir kişinin boyu ile kilosunu, iki zarın gösterdiği sayılar, bir cihazın iki parçasının ömürleri. İki değişkenin **birlikte** nasıl davrandığını anlamak için, ortak dağılımdan hesaplanan beklenen değerlere ihtiyacımız vardır.

Bu bölümde önce $g(X_1, X_2)$ biçimindeki bir rastgele değişkenin beklenen değerini ortak olasılık fonksiyonu ya da ortak yoğunluk fonksiyonu üzerinden tanımlayacağız. Bu tanım bize iki yeni araç verecek: beklenen değerin farklı değişkenlerin toplamı üzerinde de lineer olması ve bağımsız değişkenlerde çarpımın beklenen değerinin beklenen değerlerin çarpımına ayrılması. Ardından iki değişken arasındaki lineer ilişkiyi ölçen iki sayıyı, **kovaryans** ve **korelasyon katsayısını** tanımlayacağız; korelasyon katsayısının her zaman -1 ile 1 arasında kaldığını Cauchy-Schwarz eşitsizliğiyle ispatlayacağız ve bağımsızlık ile kovaryansın sıfır olması arasındaki ilişkinin **tek yönlü** olduğunu bir örnekle göreceğiz.

Bölümün son iki kısmı hesaba ayrılmıştır: bir lineer birleşimin varyansı için genel formül ve aynı beklenen değeri birden çok yolla hesaplayan çözümlü örnekler. Bu örneklerde ortak yoğunluk, marjinal yoğunluk, ortak moment üreten fonksiyon ve dönüştürülmüş değişkenin kendi dağılımı gibi farklı yolların hep aynı sayıya ulaştığını görmek, sonraki bölümlerde en kısa yolu seçebilmemizi sağlayacaktır.

18.1 Rastgele Vektörün Fonksiyonunun Beklenen Değeri

Tek değişkenli beklenen değeri (**Tanım 16.1**) tanımlarken X 'in olasılık fonksiyonunu ya da yoğunluk fonksiyonunu kullandık. (X_1, X_2) bir rastgele vektörse elimizdeki nesne ortak olasılık fonksiyonu f_{X_1, X_2} (**Tanım 13.4**) ya da ortak yoğunluk fonksiyonu f_{X_1, X_2} 'dir (**Tanım 13.6**). $g(X_1, X_2)$ biçimindeki bir rastgele değişkenin beklenen değeri de aynı düşünceyle, g 'nin aldığı değerlerin ortak olasılıkla ağırlıklı ortalaması olarak tanımlanır.

Tanım 18.1 (Rastgele Vektörün Fonksiyonunun Beklenen Değeri). (X_1, X_2) bir rastgele vektör ve $g : \mathbb{R}^2 \rightarrow \mathbb{R}$, her $B \in \mathcal{B}(\mathbb{R})$ için

$$\{(x_1, x_2) \in \mathbb{R}^2 : g(x_1, x_2) \in B\} \in \mathcal{B}(\mathbb{R}^2)$$

koşulunu sağlayan (Borel ölçülebilir) bir fonksiyon olsun.

(i) (X_1, X_2) kesikli bir rastgele vektör ve ortak olasılık fonksiyonu f_{X_1, X_2} ise,

$$\sum_{x_1 \in D_{X_1}} \sum_{x_2 \in D_{X_2}} |g(x_1, x_2)| f_{X_1, X_2}(x_1, x_2) < \infty$$

olduğunda

$$E[g(X_1, X_2)] = \sum_{x_1 \in D_{X_1}} \sum_{x_2 \in D_{X_2}} g(x_1, x_2) f_{X_1, X_2}(x_1, x_2),$$

(ii) (X_1, X_2) sürekli bir rastgele vektör ve ortak yoğunluk fonksiyonu f_{X_1, X_2} ise,

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |g(x_1, x_2)| f_{X_1, X_2}(x_1, x_2) dx_1 dx_2 < \infty$$

olduğunda

$$E[g(X_1, X_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x_1, x_2) f_{X_1, X_2}(x_1, x_2) dx_1 dx_2$$

sayısına $g(X_1, X_2)$ rastgele değişkeninin **beklenen değeri** denir. Mutlak yakınsaklık koşulu sağlanmıyorsa beklenen değer tanımsızdır.

Tanımdaki ölçülebilirlik koşulu, $g(X_1, X_2)$ 'nin gerçekten bir rastgele değişken olmasını güvence altına alır; sürekli fonksiyonlar, polinomlar, parçalı sürekli fonksiyonlar bu koşulu sağlar ve bu bölümde karşımıza çıkacak bütün g 'ler bu türdendir.

Tanımın ilk sınaması, tek değişkenli tanımla çelişmemesidir. $g(x_1, x_2) = x_1$ alalım. Sürekli durumda iç integral marjinal yoğunluğu verir (Tanım 13.8):

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 f_{X_1, X_2}(x_1, x_2) dx_2 dx_1 = \int_{-\infty}^{\infty} x_1 \left(\int_{-\infty}^{\infty} f_{X_1, X_2}(x_1, x_2) dx_2 \right) dx_1 = \int_{-\infty}^{\infty} x_1 f_{X_1}(x_1) dx_1 = E(X_1).$$

Kesikli durumda aynı hesap toplamalarla yapılır. Demek ki $E(X_1)$ 'i ister ortak dağılımdan ister marjinal dağılımdan hesaplayalım, aynı sayıyı buluruz. Daha genel olarak $g(X_1, X_2)$ rastgele değişkeninin dağılımı bulunup (bunun için Teorem 15.1 kullanılabilir) beklenen değeri **kendi** dağılımından hesaplınsa da yine yukarıdaki tanımla aynı sayı elde edilir; bu, tek değişkenli dönüşüm formülünün (Teorem 12.2) iki boyutlu karşılığıdır. Kesikli durumda gerekçesi kısadır: $Z = g(X_1, X_2)$ için $(Z = z)$ olayı, $g(x_1, x_2) = z$ eşitliğini sağlayan hücrelerin ayrık birleşimidir, dolayısıyla

$$E(Z) = \sum_{z \in D_Z} z P(Z = z) = \sum_{z \in D_Z} z \sum_{(x_1, x_2): g(x_1, x_2) = z} f_{X_1, X_2}(x_1, x_2) = \sum_{x_1 \in D_{X_1}} \sum_{x_2 \in D_{X_2}} g(x_1, x_2) f_{X_1, X_2}(x_1, x_2).$$

Son adımda hücreler g 'nin değerine göre öbeklenmiş olduğundan her hücre tam bir kez sayılır; mutlak yakınsaklık da toplama sırasının değiştirilmesini haklı kılar. Sürekli durumda aynı akıl yürütme integrallerle yapılır. Bu bölümdeki örneklerde her iki yolu da izleyip aynı sonuca ulaştığımızı defalarca göreceğiz. Tek değişkende $E(X^k)$ momentlerini (Tanım 16.2) tanımlamıştık. İki değişkende, iki değişkenin kuvvetlerinin çarpımının beklenen değeri “ortak moment” adını alır.

Tanım 18.2 (Ortak Momentler ve Merkezi Ortak Momentler). (X_1, X_2) bir rastgele vektör ve $k_1, k_2 \in \{0, 1, 2, \dots\}$ negatif olmayan tam sayılar olsun. İlgili beklenen değerler var olduğunda,

- (a) $E(X_1^{k_1} X_2^{k_2})$ sayısına (X_1, X_2) rastgele vektörünün **(k_1, k_2) -ortak momenti**,
- (b) $E\left[(X_1 - E(X_1))^{k_1} (X_2 - E(X_2))^{k_2}\right]$ sayısına (X_1, X_2) rastgele vektörünün **(k_1, k_2) -merkezi ortak momenti** denir.

$k_2 = 0$ alınırsa $(k_1, 0)$ -ortak momenti $E(X_1^{k_1})$, yani X_1 'in kendi k_1 . momentidir; merkezi ortak moment de X_1 'in merkezi momentine döner. Bu bölümün ana kavramı olan kovaryans, $(1, 1)$ -merkezi ortak momentten başka bir şey değildir.

Tek değişkenli moment üreten fonksiyonun (Tanım 17.2) iki değişkenli karşılığı, üstel ifadeye her iki değişkeni de kendi parametresiyle koyarak elde edilir.

Tanım 18.3 (Ortak Moment Üreten Fonksiyon). (X_1, X_2) bir rastgele vektör olsun. Her $t_1 \in (-h_1, h_1)$ ve $t_2 \in (-h_2, h_2)$ için $E(e^{t_1 X_1 + t_2 X_2})$ beklenen değerinin sonlu olduğu $h_1, h_2 > 0$ sayıları varsa,

$$M_{X_1, X_2}(t_1, t_2) = E(e^{t_1 X_1 + t_2 X_2}), \quad |t_1| < h_1, |t_2| < h_2$$

fonksiyonuna (X_1, X_2) rastgele vektörünün **ortak moment üreten fonksiyonu** (ortak moment çıkaran fonksiyonu) denir.

Tek değişkenli moment üreten fonksiyonun türevleri momentleri veriyordu (Teorem 17.4). Ortak moment üreten fonksiyonun kısmi türevleri de ortak momentleri verir; üstelik marjinal moment üreten fonksiyonlar, ortak olanın bir değişkeni sıfırlanarak elde edilir. Bu iki özellik, örneklerde $E(X_1 X_2)$ gibi beklenen değerleri integral almadan bulmamızı sağlayacak.

Önerme 18.1 (Ortak Moment Üreten Fonksiyonun Özellikleri). (X_1, X_2) rastgele vektörünün ortak moment üreten fonksiyonu M_{X_1, X_2} , $|t_1| < h_1, |t_2| < h_2$ için var olsun.

(a) $M_{X_1, X_2}(t_1, 0) = M_{X_1}(t_1)$ ve $M_{X_1, X_2}(0, t_2) = M_{X_2}(t_2)$; yani marjinal moment üreten fonksiyonlar ortak olanın kesitleridir.

(b) Her $k_1, k_2 \in \{0, 1, 2, \dots\}$ için

$$\frac{\partial^{k_1+k_2}}{\partial t_1^{k_1} \partial t_2^{k_2}} M_{X_1, X_2}(t_1, t_2) \Big|_{t_1=t_2=0} = E(X_1^{k_1} X_2^{k_2}).$$

Özel olarak

$$\frac{\partial M_{X_1, X_2}}{\partial t_1} \Big|_{t_1=t_2=0} = E(X_1), \quad \frac{\partial^2 M_{X_1, X_2}}{\partial t_1 \partial t_2} \Big|_{t_1=t_2=0} = E(X_1 X_2).$$

İspat.

(a) $t_2 = 0$ konulduğunda $e^{t_1 X_1 + 0 \cdot X_2} = e^{t_1 X_1}$ olur; dolayısıyla $M_{X_1, X_2}(t_1, 0) = E(e^{t_1 X_1}) = M_{X_1}(t_1)$. Buradaki $E(e^{t_1 X_1})$ ortak dağılımdan hesaplanan beklenen değerdir; Tanım 18.1'den hemen sonra gösterdiğimiz gibi bu, marjinal dağılımdan hesaplananla aynıdır. Öteki eşitlik aynı biçimde çıkar.

(b) Sürekli durumu yazalım. $|t_1| < h_1, |t_2| < h_2$ için

$$M_{X_1, X_2}(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{t_1 x_1 + t_2 x_2} f_{X_1, X_2}(x_1, x_2) dx_1 dx_2.$$

Teorem 17.4'in ispatındaki gerekçeyle türev integralin altına alınabilir; t_1 'e göre k_1 kez ve t_2 'ye göre k_2 kez türev almak integrandı $x_1^{k_1} x_2^{k_2}$ ile çarpar:

$$\frac{\partial^{k_1+k_2}}{\partial t_1^{k_1} \partial t_2^{k_2}} M_{X_1, X_2}(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1^{k_1} x_2^{k_2} e^{t_1 x_1 + t_2 x_2} f_{X_1, X_2}(x_1, x_2) dx_1 dx_2.$$

$t_1 = t_2 = 0$ konulduğunda üstel çarpan 1 olur ve sağ taraf, Tanım 18.1 gereği $E(X_1^{k_1} X_2^{k_2})$ sayısına eşittir. Kesikli durumda integrallerin yerini toplamalar alır, hesap aynıdır.

■

Beklenen değerin lineerliğini tek değişken için Teorem 16.1 ile göstermiştik; ama o sonuç yalnızca aynı değişkenin fonksiyonlarını toplayabiliyordu. $E(X_1 + X_2) = E(X_1) + E(X_2)$ gibi bir eşitlik için iki değişkenin ortak dağılımına ihtiyaç vardır. Aşağıdaki önerme bunu sağlar ve dikkat edilirse **bağımsızlık gerektirmez**.

Önerme 18.2 (Beklenen Değerin Lineerliği). (X_1, X_2) bir rastgele vektör, $g_1, g_2: \mathbb{R}^2 \rightarrow \mathbb{R}$ Borel ölçülebilir fonksiyonlar ve $E[g_1(X_1, X_2)]$ ile $E[g_2(X_1, X_2)]$ var olsun. $c_1, c_2 \in \mathbb{R}$ sabitleri için $E[c_1 g_1(X_1, X_2) + c_2 g_2(X_1, X_2)]$ vardır ve

$$E \left[\sum_{i=1}^2 c_i g_i(X_1, X_2) \right] = \sum_{i=1}^2 c_i E [g_i(X_1, X_2)]$$

olur. Özel olarak, $E(X_1)$ ve $E(X_2)$ var olduğunda

$$E(X_1 + X_2) = E(X_1) + E(X_2), \quad E(c) = c \quad (c \in \mathbb{R}).$$

İspat.

Sürekli durumu yazalım; kesikli durumda integrallerin yerini toplamalar alır. Önce beklenen değer varlığı: her (x_1, x_2) için üçgen eşitsizliğinden

$$|c_1 g_1(x_1, x_2) + c_2 g_2(x_1, x_2)| f_{X_1, X_2}(x_1, x_2) \leq |c_1| |g_1(x_1, x_2)| f_{X_1, X_2}(x_1, x_2) + |c_2| |g_2(x_1, x_2)| f_{X_1, X_2}(x_1, x_2)$$

olur; sağ tarafın integrali varsayım gereği sonludur, dolayısıyla sol tarafinki de sonludur. Şimdi integralin lineerliğinden

$$E [c_1 g_1(X_1, X_2) + c_2 g_2(X_1, X_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (c_1 g_1 + c_2 g_2)(x_1, x_2) f_{X_1, X_2}(x_1, x_2) dx_1 dx_2 = c_1 E [g_1(X_1, X_2)] + c_2 E [g_2(X_1, X_2)]$$

$g_1(x_1, x_2) = x_1$, $g_2(x_1, x_2) = x_2$ ve $c_1 = c_2 = 1$ alınırsa $E(X_1 + X_2) = E(X_1) + E(X_2)$ çıkar; burada $E(X_1)$ ve $E(X_2)$ 'nin ortak dağılımdan ya da marjinallerden hesaplanmasının fark etmediğini yukarıda görmüştük. Sabit için $g(x_1, x_2) = c$ alınır: $E(c) = c \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X_1, X_2}(x_1, x_2) dx_1 dx_2 = c \cdot 1 = c$.

■

Tanım 18.1 ile **Önerme 18.2**'i iki boyutta yazdık; ikisi de n boyutlu rastgele vektörler için sözcüğü sözcüğüne geçerlidir. $(X_1, \dots, X_n)$ vektörü ve $g : \mathbb{R}^n \rightarrow \mathbb{R}$ için beklenen değer n katlı toplam ya da n katlı integrallerle tanımlanır, lineerlik ispatı da aynen tekrarlanır. Aşağıda üç değişkenli ve n değişkenli kullanımlarda bu genel biçime dayanacağız.

Toplamanın beklenen değeri her zaman beklenen değerlerin toplamıdır; ama **çarpımın** beklenen değeri genel olarak beklenen değerlerin çarpımı **değildir**. Bu eşitliğin sağlandığı en önemli durum bağımsızlıktır. Aşağıdaki teorem, bağımsız değişkenlerin ayrı ayrı fonksiyonlarının çarpımı için bunu söyler; hem $E(X_1 X_2)$ hem de ortak moment üreten fonksiyonun çarpanlara ayrılması bu teoremin özel halleridir.

Teorem 18.1 (Bağımsız Değişkenlerde Çarpımın Beklenen Değeri). X_1 ve X_2 bağımsız rastgele değişkenler; $g_1, g_2 : \mathbb{R} \rightarrow \mathbb{R}$, her $B \in \mathcal{B}(\mathbb{R})$ için $\{x : g_i(x) \in B\} \in \mathcal{B}(\mathbb{R})$ koşulunu sağlayan fonksiyonlar olsun. $E[g_1(X_1)]$ ve $E[g_2(X_2)]$ varsa $E[g_1(X_1) g_2(X_2)]$ de vardır ve

$$E \left[\prod_{i=1}^2 g_i(X_i) \right] = \prod_{i=1}^2 E [g_i(X_i)]$$

olur.

İspat.

Kesikli durum. X_1 ile X_2 bağımsız olduğundan her (x_1, x_2) için ortak olasılık fonksiyonu marjinallerin çarpımıdır (**Tanım 14.1**):

$$f_{X_1, X_2}(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2).$$

Önce varlık: terimleri negatif olmayan iki serinin çarpımı, terimlerin ikiyeşerli çarpımlarından oluşan çift toplama eşittir, dolayısıyla

$$\sum_{x_1 \in D_{X_1}} \sum_{x_2 \in D_{X_2}} |g_1(x_1) g_2(x_2)| f_{X_1}(x_1) f_{X_2}(x_2) = \left(\sum_{x_1 \in D_{X_1}} |g_1(x_1)| f_{X_1}(x_1) \right) \left(\sum_{x_2 \in D_{X_2}} |g_2(x_2)| f_{X_2}(x_2) \right) < \infty.$$

Mutlak yakınsaklık sağlandığından toplama sırası serbestçe değiştirilebilir ve

$$\begin{aligned} E [g_1(X_1) g_2(X_2)] &= \sum_{x_1 \in D_{X_1}} \sum_{x_2 \in D_{X_2}} g_1(x_1) g_2(x_2) f_{X_1, X_2}(x_1, x_2) \\ &= \sum_{x_1 \in D_{X_1}} g_1(x_1) f_{X_1}(x_1) \left(\sum_{x_2 \in D_{X_2}} g_2(x_2) f_{X_2}(x_2) \right) \\ &= E [g_2(X_2)] \sum_{x_1 \in D_{X_1}} g_1(x_1) f_{X_1}(x_1) = E [g_1(X_1)] E [g_2(X_2)] \end{aligned}$$

bulunur.

Süreklilik durumu. Bağımsızlık gereği $f_{X_1, X_2}(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2)$ 'dir. Mutlak integrallenebilirlik, kesikli durumdaki gibi iki tek katlı integralin çarpımından okunur:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |g_1(x_1) g_2(x_2)| f_{X_1}(x_1) f_{X_2}(x_2) dx_1 dx_2 = \left(\int_{-\infty}^{\infty} |g_1(x_1)| f_{X_1}(x_1) dx_1 \right) \left(\int_{-\infty}^{\infty} |g_2(x_2)| f_{X_2}(x_2) dx_2 \right) < \infty.$$

Böylece iki katlı integral ardışık integrallere ayrılabilir:

$$E [g_1(X_1) g_2(X_2)] = \int_{-\infty}^{\infty} g_1(x_1) f_{X_1}(x_1) \left(\int_{-\infty}^{\infty} g_2(x_2) f_{X_2}(x_2) dx_2 \right) dx_1 = E [g_1(X_1)] E [g_2(X_2)] .$$

■

Teoremda g_1 ve g_2 için özel seçimler yapınca sık kullanılan üç sonuç ortaya çıkar.

Sonuç 18.1 (Bağımsızlığın Beklenen Değere Yansımaları). X_1 ve X_2 bağımsız rastgele değişkenler olsun. İlgili beklenen değerler var olduğunda:

(1) $E(X_1 X_2) = E(X_1) E(X_2)$; daha genel olarak her $k_1, k_2 \in \{0, 1, 2, \dots\}$ için $E(X_1^{k_1} X_2^{k_2}) = E(X_1^{k_1}) E(X_2^{k_2})$, yani ortak momentler momentlerin çarpımıdır.

(2) $X_1, X_2 \geq 0$ ise $E(\sqrt{X_1 X_2}) = E(\sqrt{X_1}) E(\sqrt{X_2})$.

(3) Marjinal moment üreten fonksiyonlar $|t_1| < h_1, |t_2| < h_2$ için varsa ortak moment üreten fonksiyon da vardır ve

$$M_{X_1, X_2}(t_1, t_2) = M_{X_1}(t_1) M_{X_2}(t_2).$$

İspat.

Üçü de Teorem 18.1'in doğrudan uygulamasıdır. (1) için $g_1(x) = x^{k_1}, g_2(x) = x^{k_2}$ alınır; $k_1 = k_2 = 1$ özel hâli ilk eşitliği verir. (2) için $g_1(x) = g_2(x) = \sqrt{x}$ alınır; $\sqrt{X_1 X_2} = \sqrt{X_1} \sqrt{X_2}$ olduğu için teorem uygulanabilir. (3) için t_1, t_2 sabitlenip $g_1(x) = e^{t_1 x}, g_2(x) = e^{t_2 x}$ alınır:

$$M_{X_1, X_2}(t_1, t_2) = E(e^{t_1 X_1} e^{t_2 X_2}) = E(e^{t_1 X_1}) E(e^{t_2 X_2}) = M_{X_1}(t_1) M_{X_2}(t_2).$$

■

Sonucun (3) maddesi, ortak moment üreten fonksiyon hesaplanmış bir vektörün bağımsız **olmadığını** göstermek için de kullanılabilir: $M_{X_1, X_2}(t_1, t_2) \neq M_{X_1, X_2}(t_1, 0) M_{X_1, X_2}(0, t_2)$ ise, (3) bozulduğundan X_1 ile X_2 bağımsız olamaz.

18.2 Kovaryans ve Korelasyon Katsayısı

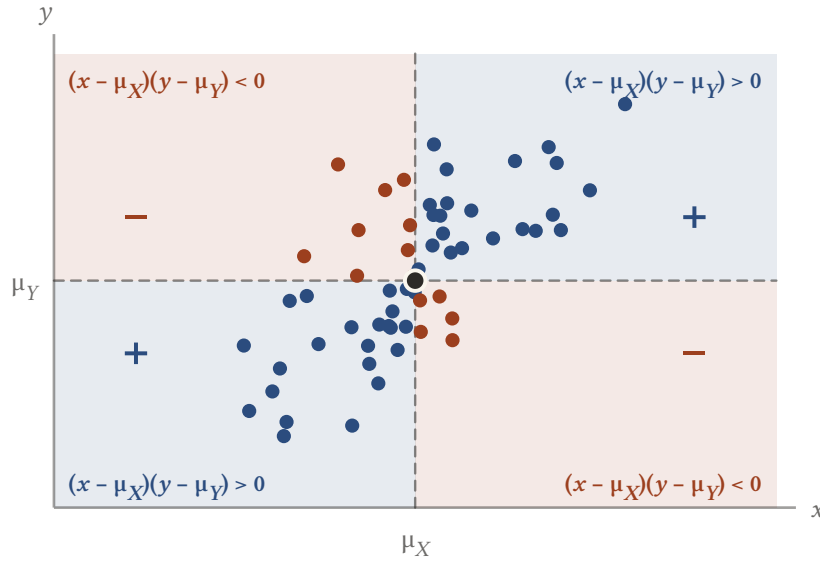
$E(X_1 X_2)$ ile $E(X_1) E(X_2)$ arasındaki fark, bağımsızlıktan ne kadar uzaklaşıldığının bir ölçüsü gibi görünür; bağımsız değişkenlerde bu fark sıfırdır. Bu farkı merkezi ortak moment biçiminde yazarak adlandırıyoruz. Tanımın anlamlı olması için ikinci momentlerin sonlu olması yeter: her x, y reel sayısı için $|xy| \leq \frac{1}{2}(x^2 + y^2)$ olduğundan, $E(X^2)$ ve $E(Y^2)$ sonluya $E|XY|$ de sonludur ve $E(XY)$ tanımlıdır.

Tanım 18.4 (Kovaryans). (X, Y) , $E(X^2) < \infty$ ve $E(Y^2) < \infty$ olan bir rastgele vektör olsun.

$$\text{Cov}(X, Y) = E[(X - E(X))(Y - E(Y))]$$

sayısına X ile Y rastgele değişkenlerinin **kovaryansı** denir. Kovaryans, (X, Y) vektörünün $(1, 1)$ -merkezi ortak momentidir.

Kovaryansın işareti ilişkinin **yönünü** söyler: X ortalamasının üstündeyken Y de çoğunlukla ortalamasının üstündeyse çarpım $(X - E(X))(Y - E(Y))$ çoğunlukla pozitif ve kovaryans pozitif çıkar; X büyükken Y küçük olma eğilimindeyse kovaryans negatiftir. Ancak kovaryansın **büyüklüğü** birimlere bağlıdır: X metre yerine santimetreyle ölçülürse kovaryans 100 katına çıkar (bunu birazdan ispatlayacağız). Birimden bağımsız bir ölçü elde etmek için kovaryans, standart sapmaların çarpımına bölünür.



Şekil 18.1: Kovaryansın işareti. Kesikli eksenler ortalama noktası (μ_X, μ_Y) 'de kesişir; sağ üst ve sol alt çeyreklerde $(x - \mu_X)(y - \mu_Y)$ çarpımı pozitif (mavi), öteki iki çeyrekte negatiftir (turuncu). Noktaların çoğu mavi çeyreklerde toplandığından çarpımların ortalaması, yani $\text{Cov}(X, Y)$, pozitif çıkar. X büyükken Y küçük olma eğiliminde olsaydı turuncu çeyrekler ağır basar ve kovaryans negatif olurdu.

Tanım 18.5 (Korelasyon Katsayısı). (X, Y) , $E(X^2) < \infty$, $E(Y^2) < \infty$, $\text{Var}(X) > 0$ ve $\text{Var}(Y) > 0$ olan bir rastgele vektör olsun. $\sigma_X = \sqrt{\text{Var}(X)}$ ve $\sigma_Y = \sqrt{\text{Var}(Y)}$ olmak üzere

$$\rho_{X,Y} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \text{Var}(Y)}}$$

sayısına X ile Y arasındaki **korelasyon katsayısı** (Pearson korelasyon katsayısı) denir. $\text{Cov}(X, Y) = 0$, dolayısıyla $\rho_{X,Y} = 0$ ise X ile Y 'ye **ilişkisiz** (korelasyonsuz) denir.

Korelasyon katsayısı, X ile Y arasındaki **lineer** ilişkinin ölçüsüdür. Bu bölümde şunları göstereceğiz: $\rho_{X,Y}$ her zaman $[-1, 1]$ aralığındadır; $\rho_{X,Y} = 1$ ancak ve ancak Y , olasılığı 1 olan bir olay üzerinde X 'in artan bir lineer fonksiyonuysa ($Y = aX + b$, $a > 0$); $\rho_{X,Y} = -1$ ancak ve ancak azalan bir lineer fonksiyonuysa ($a < 0$). $\rho_{X,Y} = 0$ olması ise X ile Y arasında **lineer** bir ilişki bulunmadığını söyler; bu, aralarında hiçbir ilişki olmadığı anlamına **gelmez**, ilişki lineer olmayan bir biçimde pekâlâ var olabilir. Bu ayrımı önce kovaryansın hesap kurallarını çıkararak, sonra somut bir örnekle netleştireceğiz.

Önerme 18.3 (Kovaryansın ve Korelasyon Katsayısının Özellikleri). X, Y, Z ikinci momentleri sonlu rastgele değişkenler ve $a, b, c, d \in \mathbb{R}$ olsun.

(a) $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$.

(b) $\text{Cov}(X, Y) = \text{Cov}(Y, X)$.

(c) $\text{Cov}(X, X) = \text{Var}(X)$.

(d) $\text{Cov}(X + a, Y + b) = \text{Cov}(X, Y)$: sabit eklemek kovaryansı değiştirmez.

(e) $\text{Cov}(aX, bY) = ab \text{Cov}(X, Y)$.

(f) $\text{Cov}(X, Y + Z) = \text{Cov}(X, Y) + \text{Cov}(X, Z)$; (b) ile birlikte kovaryans her iki girdisinde de toplamsaldır.

(g) $\text{Var}(X) > 0$, $\text{Var}(Y) > 0$ ve $ac \neq 0$ ise

$$\rho_{aX+b, cY+d} = \begin{cases} \rho_{X,Y}, & ac > 0 \\ -\rho_{X,Y}, & ac < 0. \end{cases}$$

Yani korelasyon katsayısı birim ve başlangıç noktası değişimlerinden etkilenmez; yalnızca yön ters çevrilirse işaret değiştirir.

İspat.

$\mu_X = E(X)$, $\mu_Y = E(Y)$, $\mu_Z = E(Z)$ yazalım.

(a) Çarpımı açıp **Önerme 18.2**'i uygulayalım; μ_X , μ_Y sabittir:

$$\begin{aligned}\text{Cov}(X, Y) &= E[XY - \mu_Y X - \mu_X Y + \mu_X \mu_Y] \\ &= E(XY) - \mu_Y E(X) - \mu_X E(Y) + \mu_X \mu_Y = E(XY) - \mu_X \mu_Y.\end{aligned}$$

(b) Tanımda çarpanların sırası değiştirilebilir: $(X - \mu_X)(Y - \mu_Y) = (Y - \mu_Y)(X - \mu_X)$.

(c) $Y = X$ alınırsa tanım $E[(X - \mu_X)^2] = \text{Var}(X)$ 'i verir (**Tanım 16.3**).

(d) $E(X + a) = \mu_X + a$ ve $E(Y + b) = \mu_Y + b$ olduğundan (**Teorem 16.2**)

$$(X + a) - E(X + a) = X - \mu_X, \quad (Y + b) - E(Y + b) = Y - \mu_Y;$$

sapmalar değişmediği için kovaryans da değişmez.

(e) $E(aX) = a\mu_X$ ve $E(bY) = b\mu_Y$ olduğundan

$$\text{Cov}(aX, bY) = E[(aX - a\mu_X)(bY - b\mu_Y)] = E[ab(X - \mu_X)(Y - \mu_Y)] = ab \text{Cov}(X, Y).$$

(f) $E(Y + Z) = \mu_Y + \mu_Z$ olduğundan

$$\text{Cov}(X, Y + Z) = E[(X - \mu_X)((Y - \mu_Y) + (Z - \mu_Z))] = E[(X - \mu_X)(Y - \mu_Y)] + E[(X - \mu_X)(Z - \mu_Z)],$$

son adımda **Önerme 18.2** kullanıldı.

(g) (d) ve (e) ile $\text{Cov}(aX + b, cY + d) = \text{Cov}(aX, cY) = ac \text{Cov}(X, Y)$. Öte yandan $\text{Var}(aX + b) = a^2 \text{Var}(X)$ ve $\text{Var}(cY + d) = c^2 \text{Var}(Y)$ olduğundan (**Teorem 16.2**) $\sigma_{aX+b} = |a| \sigma_X$, $\sigma_{cY+d} = |c| \sigma_Y$. Böylece

$$\rho_{aX+b, cY+d} = \frac{ac \text{Cov}(X, Y)}{|a| |c| \sigma_X \sigma_Y} = \frac{ac}{|ac|} \rho_{X, Y},$$

ve $\frac{ac}{|ac|}$ çarpanı $ac > 0$ iken 1, $ac < 0$ iken -1 'dir.

■

Önermenin (a) maddesi hesap için en kullanışlı biçimdir: kovaryans için ortak dağılımdan yalnızca $E(XY)$ 'yi, marjinallerden $E(X)$ ve $E(Y)$ 'yi hesaplamak yeter. Tek değişkenli karşılığı $\text{Var}(X) = E(X^2) - [E(X)]^2$ formülüdür (**Önerme 16.3**); (c) gereği bu formül (a)'nın $Y = X$ hâlidir.

Bağımsızlığın kovaryansa yansması artık tek satırdır.

Teorem 18.2 (Bağımsız Değişkenlerin Kovaryansı Sıfırdır). X ve Y ikinci momentleri sonlu, bağımsız rastgele değişkenler ise $\text{Cov}(X, Y) = 0$ 'dır. Ayrıca $\text{Var}(X) > 0$ ve $\text{Var}(Y) > 0$ ise $\rho_{X, Y} = 0$ 'dır; yani bağımsız değişkenler ilişkisizdir.

İspat.

Bağımsızlık gereği $E(XY) = E(X)E(Y)$ (**Sonuç 18.1 (1)**). **Önerme 18.3 (a)** ile $\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 0$ olur; korelasyon katsayısı kovaryansın pozitif bir sayıya bölümü olduğundan o da sıfırdır.

■

Teoremin tersi doğru **değildir**: kovaryansı sıfır olan iki değişken bağımsız olmayabilir. Aşağıdaki örnek bunu bir tablo ile gösterir; aynı tabloda $E(X^2 Y)$ 'yi iki yolla hesaplayıp **Tanım 18.1**'den sonra söylediğimiz “iki yol aynı sonucu verir” ilkesini de doğrulayacağız.

Örnek 18.1 (Kovaryansı Sıfır Olan Bağımlı Değişkenler). (X, Y) rastgele vektörünün ortak olasılık fonksiyonu aşağıdaki tablo ile verilsin; son satır ve son sütun marjinal olasılık fonksiyonlarıdır.

$x \setminus y$	0	1	2	$f_X(x)$
0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{3}{8}$
1	$\frac{1}{8}$	0	$\frac{1}{8}$	$\frac{2}{8}$
2	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{3}{8}$
$f_Y(y)$	$\frac{3}{8}$	$\frac{2}{8}$	$\frac{3}{8}$	1

(a) X ile Y bağımsız mıdır?

(b) $E(X)$, $E(Y)$, $\text{Var}(X)$, $\text{Var}(Y)$, $E(XY)$, $\text{Cov}(X, Y)$ ve $\rho_{X,Y}$ değerlerini hesaplayınız.

(c) $E(X^2Y)$ değerini iki yolla hesaplayınız.

(d) $\text{Var}(X^2Y)$ değerini hesaplayınız.

Çözüm.

(a) Marjinaller satır ve sütun toplamlarıdır: $f_X(0) = f_X(2) = \frac{3}{8}$, $f_X(1) = \frac{2}{8}$ ve f_Y de aynıdır. Bağımsızlık için her (x, y) çiftinde $f_{X,Y}(x, y) = f_X(x) f_Y(y)$ olmalıdır (Tanım 14.1). Oysa

$$f_{X,Y}(1, 1) = 0 \neq \frac{2}{8} \cdot \frac{2}{8} = f_X(1) f_Y(1).$$

Tek bir çiftte bozulma yeter: X ile Y bağımsız **değildir**.

(b) Marjinalden

$$E(X) = 0 \cdot \frac{3}{8} + 1 \cdot \frac{2}{8} + 2 \cdot \frac{3}{8} = 1, \quad E(X^2) = 0 \cdot \frac{3}{8} + 1 \cdot \frac{2}{8} + 4 \cdot \frac{3}{8} = \frac{14}{8} = \frac{7}{4},$$

$$\text{Var}(X) = E(X^2) - [E(X)]^2 = \frac{7}{4} - 1 = \frac{3}{4}.$$

Aynı $E(X)$ 'i ortak dağılımdan da hesaplayabiliriz (Tanım 18.1, $g(x, y) = x$): $x = 0$ satırı katkı vermez,

$$E(X) = \sum_{x=0}^2 \sum_{y=0}^2 x f_{X,Y}(x, y) = 1 \cdot \left(\frac{1}{8} + 0 + \frac{1}{8} \right) + 2 \cdot \left(\frac{1}{8} + \frac{1}{8} + \frac{1}{8} \right) = \frac{2}{8} + \frac{6}{8} = 1.$$

Tablo x ile y 'nin yer değiştirmesine göre bakışumlu olduğundan $E(Y) = 1$ ve $\text{Var}(Y) = \frac{3}{4}$ 'tür. $E(XY)$ için yalnızca $x \geq 1$ ve $y \geq 1$ olan hücreler katkı verir:

$$E(XY) = 1 \cdot 1 \cdot 0 + 1 \cdot 2 \cdot \frac{1}{8} + 2 \cdot 1 \cdot \frac{1}{8} + 2 \cdot 2 \cdot \frac{1}{8} = \frac{2+2+4}{8} = 1.$$

Böylece Önerme 18.3 (a) ile

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 1 - 1 \cdot 1 = 0, \quad \rho_{X,Y} = \frac{0}{\sqrt{\frac{3}{4} \cdot \frac{3}{4}}} = 0.$$

Demek ki X ile Y ilişkisizdir, ama (a) gereği bağımsız değildir.

(c) **Birinci yol: ortak dağılımdan.** $g(x, y) = x^2y$ alınır; yine yalnızca $x, y \geq 1$ hücreleri katkı verir:

$$E(X^2Y) = 1^2 \cdot 1 \cdot 0 + 1^2 \cdot 2 \cdot \frac{1}{8} + 2^2 \cdot 1 \cdot \frac{1}{8} + 2^2 \cdot 2 \cdot \frac{1}{8} = \frac{2+4+8}{8} = \frac{14}{8} = \frac{7}{4}.$$

İkinci yol: $Z = X^2Y$ 'nin kendi dağılımından. Z 'nin değer kümesi $D_Z = \{0, 1, 2, 4, 8\}$ 'dir. $Z = 0$ olayı $X = 0$ ya da $Y = 0$ olmasıdır:

$$P(Z = 0) = P(X = 0) + P(Y = 0) - P(X = 0, Y = 0) = \frac{3}{8} + \frac{3}{8} - \frac{1}{8} = \frac{5}{8}.$$

Öteki değerler tek bir hücreden gelir: $Z = 1$ yalnızca $(1, 1)$ 'den, $Z = 2$ yalnızca $(1, 2)$ 'den, $Z = 4$ yalnızca $(2, 1)$ 'den, $Z = 8$ yalnızca $(2, 2)$ 'den.

z	0	1	2	4	8
$P(Z = z)$	$\frac{5}{8}$	0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

Olasılıkların toplamı $\frac{5+0+1+1+1}{8} = 1$ 'dir. Tanım 16.1 ile

$$E(Z) = 0 \cdot \frac{5}{8} + 1 \cdot 0 + 2 \cdot \frac{1}{8} + 4 \cdot \frac{1}{8} + 8 \cdot \frac{1}{8} = \frac{14}{8} = \frac{7}{4}.$$

İki yol da aynı sonucu verdi.

(d) Z 'nin dağılımı elimizde olduğuna göre

$$E(Z^2) = 4 \cdot \frac{1}{8} + 16 \cdot \frac{1}{8} + 64 \cdot \frac{1}{8} = \frac{84}{8} = \frac{21}{2}, \quad \text{Var}(Z) = \frac{21}{2} - \left(\frac{7}{4}\right)^2 = \frac{168 - 49}{16} = \frac{119}{16}.$$

■

⚠ Kovaryansın sıfır olması bağımsızlık demek değildir

Gerektirme tek yönlüdür:

$$X, Y \text{ bağımsız} \Rightarrow \text{Cov}(X, Y) = 0, \quad \text{ama} \quad \text{Cov}(X, Y) = 0 \nRightarrow X, Y \text{ bağımsız}.$$

Yukarıdaki tabloda $X = 1$ olduğu bilinirse Y 'nin 1 olamayacağı kesindir; bu güçlü bir bağımlılıktır, ama lineer bir bağımlılık değildir ve korelasyon katsayısı onu görmez. Karşıt yönde kullanım ise geçerlidir: $\text{Cov}(X, Y) \neq 0$ bulunursa X ile Y 'nin bağımsız **olmadığı** kesinleşir.

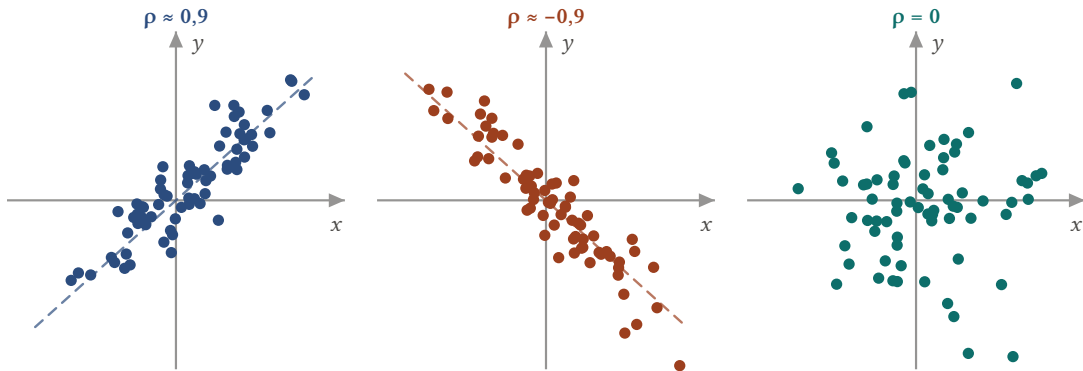
Şimdi korelasyon katsayısının neden $[-1, 1]$ aralığında kaldığını gösterelim. İspatın özü, olasılık kuramında sık kullanılan **Cauchy-Schwarz eşitsizliği**dir: ikinci momentleri sonlu U ve V için $[E(UV)]^2 \leq E(U^2)E(V^2)$. Eşitsizliği doğrudan ispatlayıp ardından sapmalara uygulayacağız; eşitlik hâlini incelemek de $\rho_{X,Y} = \pm 1$ olan durumların ne olduğunu tam olarak söyleyecek.

Teorem 18.3 (Korelasyon Katsayısının Sınırları (Cauchy-Schwarz Eşitsizliği)). X ve Y , ikinci momentleri sonlu, $\text{Var}(X) > 0$ ve $\text{Var}(Y) > 0$ olan rastgele değişkenler olsun.

(a) $[\text{Cov}(X, Y)]^2 \leq \text{Var}(X)\text{Var}(Y)$; yani

$$-1 \leq \rho_{X,Y} \leq 1.$$

(b) $|\rho_{X,Y}| = 1$ olması için gerek ve yeter koşul, $P(Y = aX + b) = 1$ olacak biçimde $a \neq 0$ ve b sabitlerinin bulunmasıdır. Bu durumda $a > 0$ ise $\rho_{X,Y} = 1$, $a < 0$ ise $\rho_{X,Y} = -1$ 'dir.



Şekil 18.2: Korelasyon katsayısı doğrusal ilişkinin yönünü ve gücünü ölçer. Solda ρ 'ya yakın pozitif: noktalar artan bir doğru çevresinde toplanır; ortada negatif: azalan doğru; sağda $\rho = 0$: doğrusal bir eğilim yok. $\rho = 0$ bağımsızlık demek değildir — yalnızca doğrusal bağımlı değildir.

İspat.

$\mu_X = E(X)$, $\mu_Y = E(Y)$ olsun.

Adım 1: Cauchy-Schwarz eşitsizliği. U ve V , $E(U^2) < \infty$, $0 < E(V^2) < \infty$ olan rastgele değişkenler olsun. $|uv| \leq \frac{1}{2}(u^2 + v^2)$ olduğundan $E(UV)$ tanımlıdır. Her $t \in \mathbb{R}$ için $(U - tV)^2 \geq 0$ olduğundan beklenen değeri de negatif değildir; kareyi açıp **Önerme 18.2**'i uygularsak

$$0 \leq E[(U - tV)^2] = E(U^2) - 2tE(UV) + t^2E(V^2).$$

Bu, her t için geçerlidir; özel olarak $t_0 = \frac{E(UV)}{E(V^2)}$ seçelim:

$$0 \leq E(U^2) - 2 \frac{[E(UV)]^2}{E(V^2)} + \frac{[E(UV)]^2}{E(V^2)} = E(U^2) - \frac{[E(UV)]^2}{E(V^2)}.$$

$E(V^2) > 0$ ile çarpınca

$$[E(UV)]^2 \leq E(U^2)E(V^2)$$

elde edilir. Ayrıca eşitliğin sağlanması, $E[(U - t_0V)^2] = 0$ olmasına denktir.

Adım 2: (a) maddesi. $U = X - \mu_X$ ve $V = Y - \mu_Y$ alalım. $E(U^2) = \text{Var}(X)$, $E(V^2) = \text{Var}(Y) > 0$ ve $E(UV) = \text{Cov}(X, Y)$ 'dir. Adım 1 ile

$$[\text{Cov}(X, Y)]^2 \leq \text{Var}(X)\text{Var}(Y).$$

Her iki taraf $\text{Var}(X)\text{Var}(Y) > 0$ 'a bölünürse $\rho_{X,Y}^2 \leq 1$, yani $-1 \leq \rho_{X,Y} \leq 1$ bulunur.

Adım 3: $|\rho_{X,Y}| = 1$ ise Y , X 'in lineer fonksiyonudur. $\rho_{X,Y}^2 = 1$ olsun. O zaman Adım 1'deki eşitsizlik eşitlikle sağlanır; dolayısıyla $t_0 = \frac{\text{Cov}(X,Y)}{\text{Var}(Y)}$ için

$$W = U - t_0V = (X - \mu_X) - t_0(Y - \mu_Y)$$

rastgele değişkeni $E(W^2) = 0$ sağlar. $E(W) = E(U) - t_0E(V) = 0$ olduğundan $\text{Var}(W) = E(W^2) = 0$ 'dır. Chebyshev eşitsizliği (**Teorem 16.4**) her $n \in \mathbb{N}$ için

$$P\left(|W| \geq \frac{1}{n}\right) \leq n^2 \text{Var}(W) = 0$$

verir. $(W \neq 0) = \bigcup_{n=1}^{\infty} (|W| \geq \frac{1}{n})$ olduğundan, olasılık ölçüsünün sayılabilir alt toplamsallığıyla (**Sonuç 3.1**)

$$P(W \neq 0) \leq \sum_{n=1}^{\infty} P\left(|W| \geq \frac{1}{n}\right) = 0,$$

yani $P(W = 0) = 1$. Demek ki olasılığı 1 olan bir olay üzerinde

$$X - \mu_X = t_0(Y - \mu_Y).$$

$\rho_{X,Y} \neq 0$ olduğundan $\text{Cov}(X, Y) \neq 0$, dolayısıyla $t_0 \neq 0$ 'dır; eşitlik Y için çözülebilir:

$$Y = \mu_Y + \frac{1}{t_0}(X - \mu_X) = aX + b, \quad a = \frac{1}{t_0} = \frac{\text{Var}(Y)}{\text{Cov}(X, Y)}, \quad b = \mu_Y - a\mu_X.$$

a 'nın işareti $\text{Cov}(X, Y)$ 'nin, yani $\rho_{X,Y}$ 'nin işaretidir: $\rho_{X,Y} = 1$ ise $a > 0$, $\rho_{X,Y} = -1$ ise $a < 0$.

Adım 4: Y , X 'in lineer fonksiyonuysa $|\rho_{X,Y}| = 1$. $a \neq 0$ ve $P(Y = aX + b) = 1$ olsun. Beklenen değer hesaplarında olasılığı sıfır olan bir olay sonucu değiştirmedikinden Y yerine $aX + b$ yazılabilir. **Önerme 18.3** (d), (e) ve (c) ile

$$\text{Cov}(X, Y) = \text{Cov}(X, aX + b) = \text{Cov}(X, aX) = a \text{Cov}(X, X) = a \text{Var}(X),$$

ve **Teorem 16.2** ile $\text{Var}(Y) = a^2 \text{Var}(X)$, yani $\sigma_Y = |a| \sigma_X$. Böylece

$$\rho_{X,Y} = \frac{a \text{Var}(X)}{\sigma_X \cdot |a| \sigma_X} = \frac{a}{|a|},$$

bu da $a > 0$ için 1, $a < 0$ için -1 'dir. Adım 3 ile birlikte (b) ispatlanmış olur.

■

Teorem, korelasyon katsayısının okunuşunu kesinleştirir: $\rho_{X,Y}$ 'nin 1'e ya da -1 'e yakın olması, (X, Y) değer çiftlerinin bir doğru çevresinde toplandığını; 0'a yakın olması, lineer bir eğilimin bulunmadığını söyler. Ancak $\rho_{X,Y} = 0$ olduğunda bile, Örnek 18.1'deki gibi, lineer olmayan güçlü bir bağımlılık var olabilir.

Sürekli bir örnekle kovaryans ve korelasyon hesabının bütün adımlarını görelim.

Örnek 18.2 (Sürekli Bir Vektörde Kovaryans ve Korelasyon). (X, Y) rastgele vektörünün ortak yoğunluk fonksiyonu

$$f_{X,Y}(x, y) = \begin{cases} \frac{3x^2+xy}{3}, & 0 < x < 1, 0 < y < 2 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun. $E(X)$, $E(Y)$, $E(XY)$, $\text{Cov}(X, Y)$, $\text{Var}(X)$, $\text{Var}(Y)$ ve $\rho_{X,Y}$ değerlerini hesaplayınız.

Çözüm.

Önce marjinal yoğunlukları bulalım (Tanım 13.8). $0 < x < 1$ için

$$f_X(x) = \int_0^2 \frac{3x^2+xy}{3} dy = \frac{1}{3} \left[3x^2y + \frac{xy^2}{2} \right]_{y=0}^{y=2} = \frac{6x^2+2x}{3} = 2x^2 + \frac{2x}{3},$$

ve $0 < y < 2$ için

$$f_Y(y) = \int_0^1 \frac{3x^2+xy}{3} dx = \frac{1}{3} \left[x^3 + \frac{x^2y}{2} \right]_{x=0}^{x=1} = \frac{1}{3} + \frac{y}{6}.$$

Sağlama: $\int_0^1 (2x^2 + \frac{2x}{3}) dx = \frac{2}{3} + \frac{1}{3} = 1$.

Beklenen değerler. Marjinallerden

$$E(X) = \int_0^1 x \left(2x^2 + \frac{2x}{3} \right) dx = \frac{2}{4} + \frac{2}{9} = \frac{13}{18}, \quad E(X^2) = \int_0^1 x^2 \left(2x^2 + \frac{2x}{3} \right) dx = \frac{2}{5} + \frac{2}{12} = \frac{17}{30},$$

$$E(Y) = \int_0^2 y \left(\frac{1}{3} + \frac{y}{6} \right) dy = \frac{1}{3} \cdot 2 + \frac{1}{6} \cdot \frac{8}{3} = \frac{2}{3} + \frac{4}{9} = \frac{10}{9}, \quad E(Y^2) = \int_0^2 y^2 \left(\frac{1}{3} + \frac{y}{6} \right) dy = \frac{8}{9} + \frac{16}{24} = \frac{14}{9}.$$

$E(XY)$ için ortak yoğunluk gerekir (Tanım 18.1):

$$E(XY) = \int_0^1 \int_0^2 xy \frac{3x^2+xy}{3} dy dx = \frac{1}{3} \int_0^1 \left[\frac{3x^3y^2}{2} + \frac{x^2y^3}{3} \right]_{y=0}^{y=2} dx = \frac{1}{3} \int_0^1 \left(6x^3 + \frac{8x^2}{3} \right) dx = \frac{1}{3} \left(\frac{3}{2} + \frac{8}{9} \right) = \frac{43}{54}.$$

Kovaryans ve varyanslar. Önerme 18.3 (a) ve Önerme 16.3 ile

$$\text{Cov}(X, Y) = \frac{43}{54} - \frac{13}{18} \cdot \frac{10}{9} = \frac{129}{162} - \frac{130}{162} = -\frac{1}{162},$$

$$\text{Var}(X) = \frac{17}{30} - \left(\frac{13}{18} \right)^2 = \frac{918}{1620} - \frac{845}{1620} = \frac{73}{1620}, \quad \text{Var}(Y) = \frac{14}{9} - \left(\frac{10}{9} \right)^2 = \frac{126}{81} - \frac{100}{81} = \frac{26}{81}.$$

Korelasyon katsayısı.

$$\rho_{X,Y} = \frac{-\frac{1}{162}}{\sqrt{\frac{73}{1620} \cdot \frac{26}{81}}} \approx \frac{-0,00617}{0,1203} \approx -0,051.$$

Kovaryans negatif ama sıfıra çok yakındır; korelasyon katsayısı da $[-1, 1]$ aralığında, Teorem 18.3'nin öngördüğü gibi, sıfıra çok yakın çıkmıştır. Yine de X ile Y bağımsız değildir: örneğin yoğunluğun tanımlı olduğu bölgenin iç noktalarından $x = \frac{1}{2}, y = \frac{3}{2}$ noktasında $f_X(\frac{1}{2}, \frac{3}{2}) = \frac{3/4+3/4}{3} = \frac{1}{2} = \frac{36}{72}$, oysa $f_X(\frac{1}{2})f_Y(\frac{3}{2}) = \frac{5}{6} \cdot \frac{7}{12} = \frac{35}{72} \neq \frac{36}{72}$. Bu, $\text{Cov}(X, Y) \neq 0$ bulgusuyla da tutarlıdır: kovaryansı sıfırdan farklı olan değişkenler bağımsız olamaz.

■

18.3 Lineer Birleşimin Varyansı

Beklenen değer toplam üzerinde lineerdir; varyans değildir. $\text{Var}(X_1 + X_2)$ 'yi hesaplamaya kalkınca, iki değişkenin sapmalarının çarpımı, yani kovaryans, kaçınılmaz olarak ortaya çıkar. Aşağıdaki teorem genel formülü verir; sonrasında bağımsız (daha doğrusu ilişkisiz) değişkenlerde formülün nasıl sadeleştiğini göreceğiz.

Teorem 18.4 (Lineer Birleşimin Varyansı). X_1 ve X_2 ikinci momentleri sonlu rastgele değişkenler ve $a_1, a_2 \in \mathbb{R}$ olsun.

$$\text{Var} \left(\sum_{i=1}^2 a_i X_i \right) = \sum_{i=1}^2 a_i^2 \text{Var}(X_i) + 2a_1 a_2 \text{Cov}(X_1, X_2).$$

Özel olarak

$$\text{Var}(X_1 + X_2) = \text{Var}(X_1) + \text{Var}(X_2) + 2 \text{Cov}(X_1, X_2), \quad \text{Var}(X_1 - X_2) = \text{Var}(X_1) + \text{Var}(X_2) - 2 \text{Cov}(X_1, X_2).$$

İspat.

$\mu_i = E(X_i)$ olsun ve $S = a_1 X_1 + a_2 X_2$ yazalım. Önerme 18.2 ile $E(S) = a_1 \mu_1 + a_2 \mu_2$; dolayısıyla

$$S - E(S) = a_1(X_1 - \mu_1) + a_2(X_2 - \mu_2).$$

Karesini alalım:

$$(S - E(S))^2 = a_1^2(X_1 - \mu_1)^2 + a_2^2(X_2 - \mu_2)^2 + 2a_1 a_2(X_1 - \mu_1)(X_2 - \mu_2).$$

Her üç terimin beklenen değeri vardır (ikinci momentler sonlu ve $|uv| \leq \frac{1}{2}(u^2 + v^2)$). Beklenen değer alıp lineerliği kullanırsak

$$\text{Var}(S) = a_1^2 E[(X_1 - \mu_1)^2] + a_2^2 E[(X_2 - \mu_2)^2] + 2a_1 a_2 E[(X_1 - \mu_1)(X_2 - \mu_2)],$$

ve sağdaki üç beklenen değer sırasıyla $\text{Var}(X_1)$, $\text{Var}(X_2)$ ve $\text{Cov}(X_1, X_2)$ 'dir. Özel durumlar $a_1 = 1, a_2 = \pm 1$ ile elde edilir.

■

Sonuç 18.2 (Bağımsız Değişkenlerde Toplamın Varyansı). X_1 ve X_2 ikinci momentleri sonlu, bağımsız rastgele değişkenler ise

$$\text{Var}(X_1 + X_2) = \text{Var}(X_1 - X_2) = \text{Var}(X_1) + \text{Var}(X_2)$$

ve daha genel olarak $\text{Var}(a_1 X_1 + a_2 X_2) = a_1^2 \text{Var}(X_1) + a_2^2 \text{Var}(X_2)$ olur.

İspat.

Bağımsızlık gereği $\text{Cov}(X_1, X_2) = 0$ 'dır (Teorem 18.2). Teorem 18.4'ndeki kovaryans terimi düşer; $a_1 = 1, a_2 = \pm 1$ alınırsa iki özel eşitlik çıkar. Farkın varyansında da varyansların toplanmasının nedeni, $a_2 = -1$ değerinin formüle yalnızca karesiyle girmesidir.

■

i Bağımsızlık değil, ilişkisizlik yeter

Sonucun ispatında bağımsızlığın tek kullanımı $\text{Cov}(X_1, X_2) = 0$ eşitliğidir. Dolayısıyla $\text{Var}(X_1 \pm X_2) = \text{Var}(X_1) + \text{Var}(X_2)$ eşitliği, bağımsız olmayan ama ilişkisiz değişkenler için de geçerlidir; **Örnek 18.1**'deki X ile Y buna örnektir.

Aynı hesap n değişken için de yapılır: $S = \sum_{i=1}^n a_i X_i$ için $(S - E(S))^2$ açıldığında $i = j$ terimleri varyansları, $i \neq j$ terimleri kovaryansları verir ve

$$\text{Var}\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 \text{Var}(X_i) + 2 \sum_{1 \leq i < j \leq n} a_i a_j \text{Cov}(X_i, X_j)$$

bulunur. Değişkenler ikişer ikişer ilişkisizse yalnızca ilk toplam kalır.

18.4 Çözümlü Örnekler

Bu kısımdaki örneklerin ortak amacı, aynı beklenen değere giden farklı yolları yan yana görmektir: ortak yoğunlukla doğrudan integral, marjinal yoğunluk, ortak moment üreten fonksiyonun türevleri, bağımsızlıkla çarpanlara ayırma ve dönüştürülmüş değişkenin kendi yoğunluğu. Hangi yolun kısa olduğu probleme göre değişir.

Örnek 18.3 (Bağımsız Üstel Değişkenlerde Ortak Moment Üreten Fonksiyon). X ve Y rastgele değişkenlerinin ortak yoğunluk fonksiyonu

$$f_{X,Y}(x,y) = \begin{cases} 6e^{-3x-2y}, & x > 0, y > 0 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun.

- (a) Marjinal yoğunluk fonksiyonlarını bulunuz; X ile Y bağımsız mıdır?
- (b) Ortak moment üreten fonksiyonu bulunuz.
- (c) Marjinal moment üreten fonksiyonları bulunuz ve ortak olanla karşılaştırınız.
- (d) $E(XY)$ değerini üç farklı yolla hesaplayınız.

Çözüm.

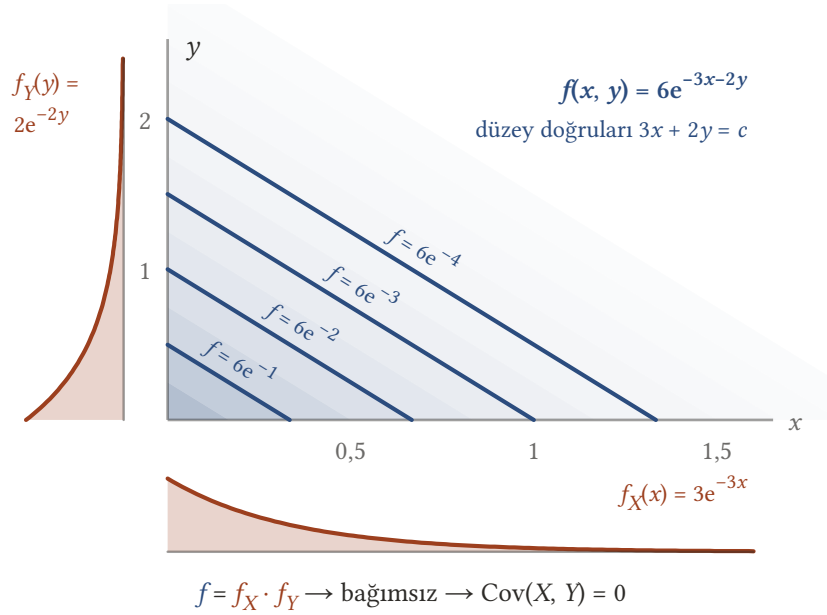
(a) $x > 0$ için

$$f_X(x) = \int_0^\infty 6e^{-3x-2y} dy = 6e^{-3x} \int_0^\infty e^{-2y} dy = 6e^{-3x} \cdot \frac{1}{2} = 3e^{-3x},$$

ve $y > 0$ için

$$f_Y(y) = \int_0^\infty 6e^{-3x-2y} dx = 6e^{-2y} \cdot \frac{1}{3} = 2e^{-2y}.$$

Her (x, y) için $f_X(x) f_Y(y) = 3e^{-3x} \cdot 2e^{-2y} = 6e^{-3x-2y} = f_{X,Y}(x, y)$ olduğundan ($x \leq 0$ ya da $y \leq 0$ olduğunda her iki taraf 0'dır) X ile Y **bağımsızdır** (Tanım 14.1).



Şekil 18.3: Birinci çeyrekte $f(x, y) = 6e^{-3x-2y}$ ortak yoğunluğu. Yoğunluk yalnızca $3x + 2y$ toplamına bağlıdır; bu yüzden düzey eğrileri paralel doğrulardır ve gölge orijinden uzaklaştıkça açılır. Kenarlardaki marjinaler $f_X(x) = 3e^{-3x}$ ve $f_Y(y) = 2e^{-2y}$ üstel yoğunluklardır; çarpımları ortak yoğunluğu verdiği için X ile Y bağımsızdır, dolayısıyla $\text{Cov}(X, Y) = 0$ olur.

(b) Tanım 18.3 ile

$$M_{X,Y}(t_1, t_2) = E(e^{t_1 X + t_2 Y}) = \int_0^\infty \int_0^\infty e^{t_1 x + t_2 y} 6e^{-3x-2y} dx dy = 6 \int_0^\infty e^{-(3-t_1)x} dx \int_0^\infty e^{-(2-t_2)y} dy.$$

İntegraller $3 - t_1 > 0$ ve $2 - t_2 > 0$ iken yakınsar ve sırasıyla $\frac{1}{3-t_1}$, $\frac{1}{2-t_2}$ değerini alır:

$$M_{X,Y}(t_1, t_2) = \frac{6}{(3-t_1)(2-t_2)}, \quad t_1 < 3, t_2 < 2.$$

(c) Tanım 17.2 ile

$$M_X(t_1) = \int_0^\infty e^{t_1 x} 3e^{-3x} dx = \frac{3}{3-t_1} \quad (t_1 < 3), \quad M_Y(t_2) = \int_0^\infty e^{t_2 y} 2e^{-2y} dy = \frac{2}{2-t_2} \quad (t_2 < 2).$$

Çarpımları $\frac{3}{3-t_1} \cdot \frac{2}{2-t_2} = \frac{6}{(3-t_1)(2-t_2)} = M_{X,Y}(t_1, t_2)$ 'dir; bu, bağımsız değişkenler için **Sonuç 18.1** (3)'ün söylediğidir. Ayrıca $M_{X,Y}(t_1, 0) = \frac{6}{(3-t_1) \cdot 2} = M_X(t_1)$ olması **Önerme 18.1** (a) ile uyumludur.

(d) **Birinci yol: ortak moment üreten fonksiyonun karma türevi.** **Önerme 18.1** (b) gereği

$$\frac{\partial^2 M_{X,Y}}{\partial t_1 \partial t_2} = \frac{\partial}{\partial t_1} \left(\frac{6}{(3-t_1)(2-t_2)^2} \right) = \frac{6}{(3-t_1)^2(2-t_2)^2}, \quad E(XY) = \frac{6}{(3-t_1)^2(2-t_2)^2} \Big|_{t_1=t_2=0} = \frac{6}{9 \cdot 4} = \frac{1}{6}.$$

İkinci yol: ortak yoğunlukla doğrudan integral. **Tanım 18.1** ile, $\int_0^\infty u e^{-cu} du = \frac{1}{c^2}$ ($c > 0$) olduğundan,

$$E(XY) = \int_0^\infty \int_0^\infty xy 6e^{-3x-2y} dx dy = 6 \int_0^\infty x e^{-3x} dx \int_0^\infty y e^{-2y} dy = 6 \cdot \frac{1}{9} \cdot \frac{1}{4} = \frac{1}{6}.$$

Üçüncü yol: bağımsızlık. X ile Y bağımsız olduğundan $E(XY) = E(X)E(Y)$ 'dir (**Sonuç 18.1** (1)). Marjinallerden $E(X) = \int_0^\infty 3x e^{-3x} dx = \frac{3}{9} = \frac{1}{3}$ ve $E(Y) = \int_0^\infty 2y e^{-2y} dy = \frac{2}{4} = \frac{1}{2}$; aynı değerler $M'_X(0) = \frac{3}{(3-t_1)^2} \Big|_0 = \frac{1}{3}$ ve $M'_Y(0) = \frac{2}{(2-t_2)^2} \Big|_0 = \frac{1}{2}$ türevlerinden de okunur (**Teorem 17.4**). Böylece

$$E(XY) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}.$$

Üç yol da $\frac{1}{6}$ verdi. Sonuç olarak $\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = \frac{1}{6} - \frac{1}{6} = 0$ 'dır; **Teorem 18.2'**in söylediği gibi.

■

Örnek 18.4 (Üçgen Bölge Üzerinde Beklenen Değerin İki Yolla Hesabı). (X, Y) rastgele vektörün ortak yoğunluk fonksiyonu

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{2}, & 0 < y < x < 2 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun.

(a) Marjinal yoğunluk fonksiyonlarını bulunuz.

(b) $E(X)$ değerini ortak yoğunlukla ve marjinal yoğunlukla hesaplayınız.

(c) $E(XY)$ değerini ortak yoğunlukla ve $Z = XY$ 'nin yoğunluk fonksiyonuyla hesaplayınız.

Çözüm.

Yoğunluk, köşeleri $(0, 0)$, $(2, 0)$, $(2, 2)$ olan, alanı 2 olan üçgen üzerinde sabittir; $\frac{1}{2} \cdot 2 = 1$ olduğundan toplam olasılık 1'dir.

(a) $0 < x < 2$ için y , 0'dan x 'e kadar değişir:

$$f_X(x) = \int_0^x \frac{1}{2} dy = \frac{x}{2}.$$

$0 < y < 2$ için x , y 'den 2'ye kadar değişir:

$$f_Y(y) = \int_y^2 \frac{1}{2} dx = \frac{2-y}{2}.$$

(b) **Ortak yoğunlukla.** **Tanım 18.1'**de $g(x, y) = x$:

$$E(X) = \int_0^2 \int_0^x x \cdot \frac{1}{2} dy dx = \int_0^2 \frac{x^2}{2} dx = \frac{8}{6} = \frac{4}{3}.$$

Marjinal yoğunlukla. **Tanım 16.1** ile

$$E(X) = \int_0^2 x \cdot \frac{x}{2} dx = \frac{8}{6} = \frac{4}{3}.$$

İki hesap aynıdır; aslında birincideki iç integral tam olarak $f_X(x)$ 'i üretmektedir.

(c) **Ortak yoğunlukla.**

$$E(XY) = \int_0^2 \int_0^x xy \cdot \frac{1}{2} dy dx = \int_0^2 \frac{x}{2} \cdot \frac{x^2}{2} dx = \int_0^2 \frac{x^3}{4} dx = \frac{16}{16} = 1.$$

$Z = XY$ 'nin yoğunluğuyla. Önce Z 'nin yoğunluğunu **Teorem 15.1** ile bulalım. $Z = XY$, $U = X$ dönüşümünün tersi $x = u$, $y = \frac{z}{u}$ 'dur ve

$$\left| \frac{\partial(x, y)}{\partial(u, z)} \right| = \left| \det \begin{pmatrix} 1 & 0 \\ -\frac{z}{u^2} & \frac{1}{u} \end{pmatrix} \right| = \frac{1}{u}.$$

$0 < y < x < 2$ koşulu $0 < \frac{z}{u} < u < 2$ 'ye, yani $z > 0$, $u > \sqrt{z}$ ve $u < 2$ 'ye dönüşür; bu bölge boş olmasın diye $0 < z < 4$ olmalıdır. Dolayısıyla $f_{U,Z}(u, z) = \frac{1}{2} \cdot \frac{1}{u} (\sqrt{z} < u < 2)$ ve

$$f_Z(z) = \int_{\sqrt{z}}^2 \frac{1}{2u} du = \frac{1}{2} (\ln 2 - \ln \sqrt{z}) = \frac{1}{2} \ln 2 - \frac{1}{4} \ln z = \frac{1}{4} \ln \frac{4}{z}, \quad 0 < z < 4.$$

Hesaplarda $z = 4w$ dönüşümü ve

$$\int_0^1 w^n (-\ln w) dw = \frac{1}{(n+1)^2} \quad (n = 0, 1, 2, \dots)$$

formülü işimize yarayacak. Formül kısmi integrasyonla çıkar:

$$\int_0^1 w^n (-\ln w) dw = \left[-\ln w \cdot \frac{w^{n+1}}{n+1} \right]_0^1 + \int_0^1 \frac{w^n}{n+1} dw = 0 + \frac{1}{(n+1)^2}.$$

Sınır terimi $w \rightarrow 0^+$ iken $w^{n+1} \ln w \rightarrow 0$ olduğu için düşer. Buna göre

$$\int_0^4 f_Z(z) dz = \int_0^1 \frac{1}{4} \ln \frac{1}{w} \cdot 4 dw = \int_0^1 (-\ln w) dw = 1,$$

yani f_Z gerçekten bir yoğunluktur, ve

$$E(Z) = \int_0^4 z \cdot \frac{1}{4} \ln \frac{4}{z} dz = \int_0^1 4w \cdot \frac{1}{4} \ln \frac{1}{w} \cdot 4 dw = 4 \int_0^1 w (-\ln w) dw = 4 \cdot \frac{1}{4} = 1.$$

Bu da $E(XY) = 1$ sonucunu doğrular. Marjinalden $E(Y) = \int_0^2 y \cdot \frac{2-y}{2} dy = \frac{2}{3}$ olduğundan $\text{Cov}(X, Y) = 1 - \frac{4}{3} \cdot \frac{2}{3} = \frac{1}{9} > 0$ 'dır: $Y < X$ kısıtı yüzünden X büyükken Y de büyük olma eğilimindedir.

■

Örnek 18.5 (Çarpımın Varyansının İki Yolla Hesabı). (X, Y) rastgele vektörünün ortak yoğunluk fonksiyonu

$$f_{X,Y}(x, y) = \begin{cases} 2, & 0 < x < y < 1 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun. $Z = XY$ rastgele değişkeninin varyansını iki yolla hesaplayınız.

Çözüm.

Birinci yol: ortak yoğunlukla. $\text{Var}(Z) = E(Z^2) - [E(Z)]^2$ formülünü (Önerme 16.3) kullanacağız; iki beklenen değer de Tanım 18.1 ile ortak yoğunluktan hesaplanır. Bölge $0 < x < y < 1$ olduğundan iç integral y üzerinden x 'ten 1'e alınır:

$$E(XY) = \int_0^1 \int_x^1 2xy dy dx = \int_0^1 x(1-x^2) dx = \frac{1}{2} - \frac{1}{4} = \frac{1}{4},$$

$$E[(XY)^2] = \int_0^1 \int_x^1 2x^2 y^2 dy dx = \int_0^1 \frac{2x^2}{3} (1-x^3) dx = \frac{2}{3} \left(\frac{1}{3} - \frac{1}{6} \right) = \frac{1}{9}.$$

Böylece

$$\text{Var}(XY) = \frac{1}{9} - \frac{1}{16} = \frac{16-9}{144} = \frac{7}{144}.$$

Aynı sonuç varyansın tanımıyla da bulunur:

$$\begin{aligned}
 E \left[\left(XY - \frac{1}{4} \right)^2 \right] &= E[(XY)^2] - \frac{1}{2}E(XY) + \frac{1}{16} \\
 &= \frac{1}{9} - \frac{1}{8} + \frac{1}{16} = \frac{16 - 18 + 9}{144} = \frac{7}{144}.
 \end{aligned}$$

İkinci yol: Z 'nin yoğunluğuyla. $Z = XY$, $U = X$ dönüşümü için ters dönüşüm $x = u$, $y = \frac{z}{u}$ ve Jacobian mutlak değeri $\frac{1}{u}$ 'dur (Örnek 18.4'daki hesabın aynısı). $0 < x < y < 1$ koşulu $0 < u < \frac{z}{u} < 1$ 'e dönüşür: $u > 0$, $u^2 < z$ (yani $u < \sqrt{z}$) ve $z < u$. Demek ki $z < u < \sqrt{z}$, bu aralığın boş olmaması için $0 < z < 1$ gerekir. Teorem 15.1 ile $f_{U,Z}(u, z) = 2 \cdot \frac{1}{u}$ ($z < u < \sqrt{z}$) ve

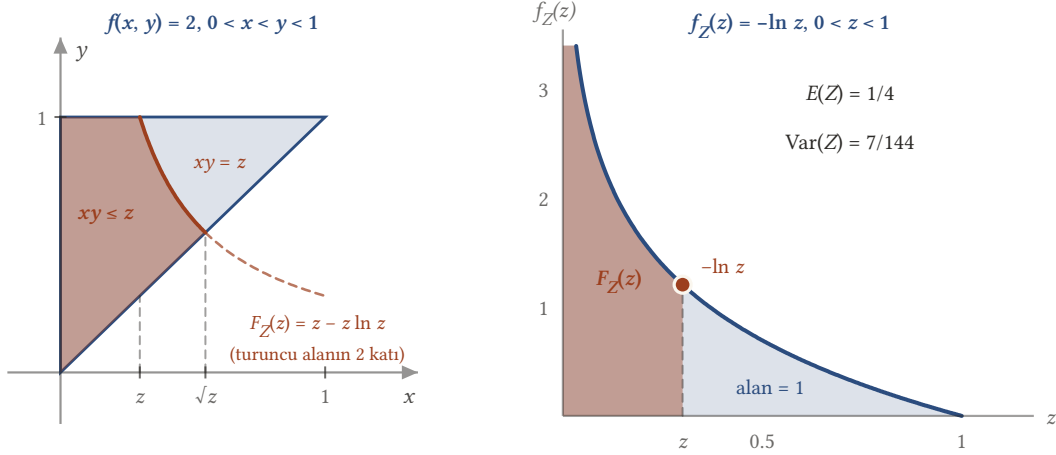
$$f_Z(z) = \int_z^{\sqrt{z}} \frac{2}{u} du = 2 (\ln \sqrt{z} - \ln z) = 2 \left(\frac{1}{2} \ln z - \ln z \right) = -\ln z, \quad 0 < z < 1.$$

$\int_0^1 (-\ln z) dz = 1$ olduğundan bu bir yoğunluktur. Örnek 18.4'da ispatladığımız $\int_0^1 z^n (-\ln z) dz = \frac{1}{(n+1)^2}$ formülüyle

$$E(Z) = \int_0^1 z (-\ln z) dz = \frac{1}{4}, \quad E(Z^2) = \int_0^1 z^2 (-\ln z) dz = \frac{1}{9},$$

$$\text{Var}(Z) = \frac{1}{9} - \frac{1}{16} = \frac{7}{144}.$$

İki yol aynı sonucu verdi. Birinci yol Z 'nin dağılımını hiç gerektirmez; ikinci yol ise Z hakkında daha fazlasını, örneğin $P(Z \leq \frac{1}{2})$ 'yi de hesaplamaya izin verir.



Şekil 18.4: $Z = XY$ çarpımının dağılımı. Solda mavi üçgen, ortak yoğunluğun 2 olduğu bölge; $xy = z$ hiperbolünün altında kalan turuncu parça $\{XY \leq z\}$ olayıdır ve alanının 2 katı $F_Z(z) = z - z \ln z$ verir. Sağda Z 'nin yoğunluğu $-\ln z$: $z \rightarrow 0$ iken sınırsız büyür ama altındaki toplam alan 1'dir; turuncu alan yine $F_Z(z)$ 'dir. Beklenen değer $E(Z) = 1/4$, varyans $7/144$ 'tür.

■

Örnek 18.6 (Bağımsız Üstel Değişkenlerin Ölçekli Farkının Moment Üreten Fonksiyonu). X_1 ve X_2 rastgele değişkenlerinin ortak yoğunluk fonksiyonu

$$f_{X_1, X_2}(x_1, x_2) = \begin{cases} \frac{1}{4} e^{-\frac{x_1+x_2}{2}}, & x_1 > 0, x_2 > 0 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun. $Y = \frac{X_1 - X_2}{2}$ rastgele değişkeninin moment üreten fonksiyonunu bulunuz; buradan $E(Y)$ ve $\text{Var}(Y)$ 'yi okuyup Teorem 18.4 ile doğrulayınız.

Çözüm.

Moment üreten fonksiyon. $Y = g(X_1, X_2)$ olduğundan $M_Y(t) = E(e^{tY})$ ortak yoğunlukla hesaplanır (Tanım 18.1):

$$M_Y(t) = E\left(e^{\frac{t}{2}X_1 - \frac{t}{2}X_2}\right) = \int_0^\infty \int_0^\infty e^{\frac{t}{2}x_1} e^{-\frac{t}{2}x_2} \frac{1}{4} e^{-\frac{x_1}{2}} e^{-\frac{x_2}{2}} dx_1 dx_2 = \left(\frac{1}{2} \int_0^\infty e^{-\frac{(1-t)x_1}{2}} dx_1\right) \left(\frac{1}{2} \int_0^\infty e^{-\frac{(1+t)x_2}{2}} dx_2\right).$$

Birinci integral $1 - t > 0$ iken $\frac{2}{1-t}$, ikincisi $1 + t > 0$ iken $\frac{2}{1+t}$ değerini alır. Böylece

$$M_Y(t) = \frac{1}{1-t} \cdot \frac{1}{1+t} = \frac{1}{1-t^2}, \quad -1 < t < 1.$$

Bu çarpanlara ayrılma rastlantı değildir. Ortak yoğunluk $f_{X_1, X_2}(x_1, x_2) = \frac{1}{2}e^{-x_1/2} \cdot \frac{1}{2}e^{-x_2/2}$ biçiminde marjinalerin çarpımıdır, dolayısıyla X_1 ile X_2 bağımsızdır ve her birinin moment üreten fonksiyonu $M_{X_i}(s) = \frac{1}{2} \int_0^\infty e^{sx} e^{-x/2} dx = \frac{1}{1-2s}$ ($s < \frac{1}{2}$)'dir. Teorem 18.1'i $g_1(x) = e^{tx/2}$, $g_2(x) = e^{-tx/2}$ ile uygularsak

$$M_Y(t) = E\left(e^{\frac{t}{2}X_1}\right) E\left(e^{-\frac{t}{2}X_2}\right) = M_{X_1}\left(\frac{t}{2}\right) M_{X_2}\left(-\frac{t}{2}\right) = \frac{1}{1-t} \cdot \frac{1}{1+t},$$

aynı sonuç.

Momentler. $|t| < 1$ için $\frac{1}{1-t^2} = 1 + t^2 + t^4 + \dots$ açılımında t 'nin katsayısı 0, t^2 'nin katsayısı 1'dir. Teorem 17.4 gereği $E(Y) = M_Y'(0) = 0$ ve $E(Y^2) = M_Y''(0) = 2! \cdot 1 = 2$; dolayısıyla

$$\text{Var}(Y) = E(Y^2) - [E(Y)]^2 = 2.$$

Doğrulama. $M_{X_i}(s) = (1 - 2s)^{-1}$ 'den $M'_{X_i}(s) = 2(1 - 2s)^{-2}$ ve $M''_{X_i}(s) = 8(1 - 2s)^{-3}$; böylece $E(X_i) = 2$, $E(X_i^2) = 8$ ve $\text{Var}(X_i) = 8 - 4 = 4$. $Y = \frac{1}{2}X_1 - \frac{1}{2}X_2$ için Önerme 18.2 ile $E(Y) = \frac{1}{2} \cdot 2 - \frac{1}{2} \cdot 2 = 0$, ve X_1 ile X_2 bağımsız olduğundan Sonuç 18.2 ile

$$\text{Var}(Y) = \left(\frac{1}{2}\right)^2 \text{Var}(X_1) + \left(-\frac{1}{2}\right)^2 \text{Var}(X_2) = \frac{4}{4} + \frac{4}{4} = 2.$$

Her iki yol da $E(Y) = 0$, $\text{Var}(Y) = 2$ verir.

■

18.5 Alıştırmalar

Alıştırma 18.1 (Rastgele Vektörlerde Beklenen Değer, Kovaryans ve Korelasyon). (a) (X, Y) rastgele vektörünün ortak olasılık fonksiyonu aşağıdaki tabloyla verilsin.

$x \setminus y$	0	1	2
0	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{8}$
1	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$

$\text{Cov}(X, Y)$ ve $\rho_{X,Y}$ değerlerini hesaplayınız. X ile Y bağımsız mıdır?

(b) İkinci momentleri sonlu X ve Y için $\text{Cov}(X + Y, X - Y) = \text{Var}(X) - \text{Var}(Y)$ olduğunu gösteriniz. $\text{Var}(X) = \text{Var}(Y)$ ise $X + Y$ ile $X - Y$ 'nin ilişkisiz olduğu sonucunu çıkarınız.

(c) (X, Y) rastgele vektörünün ortak yoğunluk fonksiyonu $0 < x < 1$, $0 < y < 1$ için $f_{X,Y}(x, y) = x + y$, diğer durumlarda 0 olsun. $\text{Cov}(X, Y)$ ve $\rho_{X,Y}$ değerlerini hesaplayınız.

(d) X ve Y bağımsız, $E(X) = 1$, $\text{Var}(X) = 2$, $E(Y) = -1$, $\text{Var}(Y) = 3$ olsun. $E(XY)$, $\text{Var}(2X - 3Y)$, $\text{Cov}(X + Y, X - Y)$ ve $E[(X + Y)^2]$ değerlerini hesaplayınız.

(e) (X_1, X_2) rastgele vektörünün ortak moment üreten fonksiyonu her $(t_1, t_2) \in \mathbb{R}^2$ için

$$M_{X_1, X_2}(t_1, t_2) = \exp\left(t_1 + 2t_2 + \frac{t_1^2 + 2t_1t_2 + 4t_2^2}{2}\right)$$

olsun. $E(X_1)$, $E(X_2)$, $\text{Var}(X_1)$, $\text{Var}(X_2)$, $\text{Cov}(X_1, X_2)$ ve ρ_{X_1, X_2} değerlerini bulunuz. X_1 ile X_2 bağımsız mıdır?

(f) X , $(-1, 1)$ aralığında düzgün dağılımlı olsun: $f_X(x) = \frac{1}{2}$ ($-1 < x < 1$). $Y = X^2$ için $\text{Cov}(X, Y) = 0$ olduğunu, ama X ile Y 'nin bağımsız olmadığını gösteriniz.

Çözüm.

(a) Marjinaler: $f_X(0) = \frac{1}{8} + \frac{1}{4} + \frac{1}{8} = \frac{1}{2}$, $f_X(1) = \frac{1}{2}$; $f_Y(0) = \frac{3}{8}$, $f_Y(1) = \frac{3}{8}$, $f_Y(2) = \frac{1}{4}$. Buradan

$$E(X) = \frac{1}{2}, \quad E(X^2) = \frac{1}{2}, \quad \text{Var}(X) = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}; \quad E(Y) = \frac{3}{8} + 2 \cdot \frac{1}{4} = \frac{7}{8}, \quad E(Y^2) = \frac{3}{8} + 4 \cdot \frac{1}{4} = \frac{11}{8}, \quad \text{Var}(Y) = \frac{11}{8} - \frac{49}{64} = \frac{39}{64}.$$

Ortak dağılımdan, yalnızca $x = 1$ satırı katkı verir: $E(XY) = 1 \cdot 1 \cdot \frac{1}{8} + 1 \cdot 2 \cdot \frac{1}{8} = \frac{3}{8}$. Böylece

$$\text{Cov}(X, Y) = \frac{3}{8} - \frac{1}{2} \cdot \frac{7}{8} = \frac{6-7}{16} = -\frac{1}{16}, \quad \rho_{X,Y} = \frac{-\frac{1}{16}}{\sqrt{\frac{1}{4} \cdot \frac{39}{64}}} = \frac{-\frac{1}{16}}{\frac{\sqrt{39}}{16}} = -\frac{1}{\sqrt{39}} \approx -0,16.$$

$\text{Cov}(X, Y) \neq 0$ olduğundan Teorem 18.2 gereği X ile Y bağımsız olamaz. Doğrudan da görülür: $f_{X,Y}(0,0) = \frac{1}{8}$ ama $f_X(0)f_Y(0) = \frac{1}{2} \cdot \frac{3}{8} = \frac{3}{16}$.

(b) Önerme 18.3'nin (f) maddesi (iki girdide de toplamsallık), (e) maddesi ($-Y = (-1)Y$), (b) ve (c) maddeleriyle

$$\text{Cov}(X+Y, X-Y) = \text{Cov}(X, X) - \text{Cov}(X, Y) + \text{Cov}(Y, X) - \text{Cov}(Y, Y) = \text{Var}(X) - \text{Var}(Y).$$

$\text{Var}(X) = \text{Var}(Y)$ ise bu sıfırdır; Tanım 18.5 gereği $X+Y$ ile $X-Y$ ilişkisizdir. Bu, bağımsız oldukları anlamına gelmez.

(c) Yoğunluk x ile y 'de bakışlıdır, dolayısıyla $E(X) = E(Y)$ ve $\text{Var}(X) = \text{Var}(Y)$.

$$E(X) = \int_0^1 \int_0^1 x(x+y) dy dx = \int_0^1 \left(x^2 + \frac{x}{2}\right) dx = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}, \quad E(X^2) = \int_0^1 \left(x^3 + \frac{x^2}{2}\right) dx = \frac{1}{4} + \frac{1}{6} = \frac{5}{12},$$

$$E(XY) = \int_0^1 \int_0^1 xy(x+y) dy dx = \int_0^1 \left(\frac{x^2}{2} + \frac{x}{3}\right) dx = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}.$$

Böylece

$$\text{Cov}(X, Y) = \frac{1}{3} - \left(\frac{7}{12}\right)^2 = \frac{48-49}{144} = -\frac{1}{144}, \quad \text{Var}(X) = \text{Var}(Y) = \frac{5}{12} - \frac{49}{144} = \frac{60-49}{144} = \frac{11}{144},$$

$$\rho_{X,Y} = \frac{-\frac{1}{144}}{\frac{11}{144}} = -\frac{1}{11}.$$

(d) Bağımsızlıktan $E(XY) = E(X)E(Y) = -1$ (Sonuç 18.1 (1)) ve $\text{Cov}(X, Y) = 0$. Teorem 18.4 ile

$$\text{Var}(2X - 3Y) = 4 \cdot 2 + 9 \cdot 3 + 2 \cdot 2 \cdot (-3) \cdot 0 = 35.$$

Alıştırmanın (b) maddesinden $\text{Cov}(X+Y, X-Y) = \text{Var}(X) - \text{Var}(Y) = 2 - 3 = -1$. Son olarak Önerme 16.3'nü $X+Y$ 'ye uygularsak, $E(X+Y) = 0$ ve $\text{Var}(X+Y) = 2 + 3 = 5$ olduğundan

$$E[(X+Y)^2] = \text{Var}(X+Y) + [E(X+Y)]^2 = 5 + 0 = 5.$$

(e) $\varphi(t_1, t_2) = t_1 + 2t_2 + \frac{1}{2}(t_1^2 + 2t_1t_2 + 4t_2^2)$ yazalım; $M = e^\varphi$, $\varphi(0,0) = 0$, $M(0,0) = 1$. Kısmi türevler

$$\frac{\partial \varphi}{\partial t_1} = 1 + t_1 + t_2, \quad \frac{\partial \varphi}{\partial t_2} = 2 + t_1 + 4t_2.$$

Önerme 18.1 (b) ile

$$E(X_1) = \frac{\partial M}{\partial t_1} \Big|_{t_1=t_2=0} = (1 + t_1 + t_2) M \Big|_{t_1=t_2=0} = 1, \quad E(X_2) = (2 + t_1 + 4t_2) M \Big|_{t_1=t_2=0} = 2,$$

$$E(X_1^2) = \left[(1 + t_1 + t_2)^2 + 1 \right] M \Big|_{t_1=t_2=0} = 2, \quad E(X_2^2) = \left[(2 + t_1 + 4t_2)^2 + 4 \right] M \Big|_{t_1=t_2=0} = 8,$$

$$E(X_1 X_2) = \frac{\partial^2 M}{\partial t_1 \partial t_2} \Big|_{t_1=t_2=0} = \left[(1 + t_1 + t_2)(2 + t_1 + 4t_2) + 1 \right] M \Big|_{t_1=t_2=0} = 2 + 1 = 3.$$

Buradan $\text{Var}(X_1) = 2 - 1 = 1$, $\text{Var}(X_2) = 8 - 4 = 4$, $\text{Cov}(X_1, X_2) = 3 - 1 \cdot 2 = 1$ ve

$$\rho_{X_1, X_2} = \frac{1}{\sqrt{1 \cdot 4}} = \frac{1}{2}.$$

$\text{Cov}(X_1, X_2) \neq 0$ olduğundan bağımsız değildirler. Aynı sonuç [Sonuç 18.1 \(3\)](#) ile de görülür: $M(t_1, 0) M(0, t_2) = \exp(t_1 + 2t_2 + \frac{1}{2}(t_1^2 + 4t_2^2))$, ortak fonksiyondan $e^{t_1 t_2}$ çarpanı kadar farklıdır.

(f) f_X çift fonksiyon olduğundan tek kuvvetlerin beklenen değeri sıfırdır: $E(X) = \int_{-1}^1 \frac{x}{2} dx = 0$ ve $E(XY) = E(X^3) = \int_{-1}^1 \frac{x^3}{2} dx = 0$. [Önerme 18.3 \(a\)](#) ile $\text{Cov}(X, Y) = 0 - 0 \cdot E(Y) = 0$. Varyanslar pozitif: $\text{Var}(X) = E(X^2) = \frac{1}{3}$ ve $\text{Var}(Y) = E(X^4) - [E(X^2)]^2 = \frac{1}{5} - \frac{1}{9} = \frac{4}{45}$; dolayısıyla $\rho_{X,Y} = 0$ da tanımlıdır.

Bağımsızlık için dağılım fonksiyonlarına bakalım ([Tanım 14.1](#)). $(Y \leq \frac{1}{4}) = (|X| \leq \frac{1}{2})$ olduğundan

$$P\left(X \leq \frac{1}{2}, Y \leq \frac{1}{4}\right) = P\left(-\frac{1}{2} \leq X \leq \frac{1}{2}\right) = \frac{1}{2}, \quad P\left(X \leq \frac{1}{2}\right)P\left(Y \leq \frac{1}{4}\right) = \frac{3}{4} \cdot \frac{1}{2} = \frac{3}{8}.$$

İki taraf eşit olmadığından X ile Y bağımsız değildir. Oysa Y , X 'in bir fonksiyonudur; aralarındaki bağımlılık tam ve deterministiktir, ama lineer olmadığı için korelasyon katsayısı onu göremez.

■

Bir değişkenin değeri bilindiğinde ötekinin beklenen değerinin nasıl güncellendiğini, yani koşullu dağılım üzerinden alınan beklenen değeri bir sonraki bölümde inceliyoruz: [Koşullu Beklenen Değer](#).

19. Koşullu Beklenen Değer

Rastgele Vektörlerde Beklenen Değer, Kovaryans ve Korelasyon bölümünde bir rastgele vektörün fonksiyonunun beklenen değerini tanımladık; kovaryans ve korelasyon katsayısıyla iki rastgele değişken arasındaki doğrusal ilişkiyi ölçtük. Bu araçların hepsi (X, Y) vektörünün **ortak** dağılımına bakar. Oysa uygulamada sık sık şu durumla karşılaşırız: X 'in değeri öğrenilmiştir ve artık yalnızca Y merak edilmektedir. Bu durumu betimleyen nesneyi, $X = x$ verilmişken Y 'nin **koşullu dağılımını**, **Bağımsız Rastgele Değişkenler ve Koşullu Dağılımlar** bölümünde tanımlamıştık. Bu bölümde o koşullu dağılımın beklenen değerini, yani **koşullu beklenen değeri** inceleyeceğiz.

Koşullu beklenen değer $E(Y | X = x)$, “ X 'in x olduğu bilindiğinde Y 'den ortalama olarak ne beklenir?” sorusunun yanıtıdır. Bu sayı x 'e bağlıdır; x yerine X yazınca $E(Y | X)$ adlı yeni bir rastgele değişken ortaya çıkar. Bölümün ana teoremi olan **toplam beklenen değer kuralı** (kule kuralı), bu rastgele değişkenin ortalamasının tam olarak $E(Y)$ olduğunu söyler: $E[E(Y | X)] = E(Y)$. Bu, toplam olasılık formülünün beklenen değerler için olan karşılığıdır ve bir beklenen değeri “önce X 'i sabit tut, sonra X üzerinden ortalama” biçiminde iki adıma bölmemizi sağlar. Aynı fikrin varyans için karşılığı **toplam varyans kuralıdır**.

Bölümü iki uygulamayla kapatacağız. İlki, terim sayısı da rastgele olan bir toplamın $(X_1 + \dots + X_N)$ beklenen değeri ve varyansdır; sigorta hasarları, kuyruk uzunlukları gibi pek çok modelde bu yapı görülür. İkincisi, koşullu beklenen değerin X 'e bakarak Y 'yi tahmin etmenin **en iyi** yolu olmasıdır: ortalama karesel hatayı en küçük yapan tahmin $E(Y | X)$ 'tir. Bu son gözlem, istatistikteki regresyon fikrinin olasılık kuramındaki kaynağıdır.

19.1 Koşullu Beklenen Değerin Tanımı

Önce gereken nesneleri hatırlayalım. (X, Y) kesikli bir rastgele vektör ve x , $f_X(x) = P(X = x) > 0$ olan bir nokta olsun. **Teorem 14.1**'na göre $X = x$ verilmişken Y 'nin koşullu olasılık fonksiyonu

$$f_{Y|X=x}(y) = P(Y = y | X = x) = \frac{f_{X,Y}(x, y)}{f_X(x)}, \quad y \in D_{Y|X=x}$$

dir; burada $D_{Y|X=x} = \{y \in D_Y : f_{X,Y}(x, y) > 0\}$, koşullu dağılımın değer kümesidir. (X, Y) sürekli bir rastgele vektör ve $f_X(x) > 0$ ise **Tanım 14.3** gereği koşullu olasılık yoğunluk fonksiyonu

$$f_{Y|X=x}(y) = \frac{f_{X,Y}(x, y)}{f_X(x)}$$

dir. Her iki durumda da koşullu olasılık (yoğunluk) fonksiyonu gerçek bir olasılık (yoğunluk) fonksiyonudur: negatif değildir ve y üzerinden toplamı (integrali) 1'dir. Öyleyse ona göre bir beklenen değer alınabilir. Bu beklenen değer, aşağıdaki tanımın konusudur.

Tanım 19.1 (Koşullu Beklenen Değer). (X, Y) bir rastgele vektör ve x , $f_X(x) > 0$ olan bir nokta olsun.

(i) (X, Y) kesikli ve $\sum_{y \in D_{Y|X=x}} |y| f_{Y|X=x}(y) < \infty$ ise

$$E(Y | X = x) = \sum_{y \in D_{Y|X=x}} y f_{Y|X=x}(y)$$

(ii) (X, Y) sürekli ve $\int_{-\infty}^{\infty} |y| f_{Y|X=x}(y) dy < \infty$ ise

$$E(Y | X = x) = \int_{-\infty}^{\infty} y f_{Y|X=x}(y) dy$$

sayısına, $X = x$ verilmişken Y rastgele değişkeninin **koşullu beklenen değeri** denir.

Daha genel olarak, $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ Borel ölçülebilir bir fonksiyon olmak üzere, $X = x$ verilmişken $h(X, Y)$ 'nin koşullu beklenen değeri (ilgili toplam ya da integral mutlak yakınsak olduğunda)

$$E(h(X, Y) | X = x) = \sum_{y \in D_{Y|X=x}} h(x, y) f_{Y|X=x}(y) \quad \text{ya da} \quad E(h(X, Y) | X = x) = \int_{-\infty}^{\infty} h(x, y) f_{Y|X=x}(y) dy$$

ile tanımlanır. $f_X(x) = 0$ olan x noktalarında $E(Y | X = x) = 0$ alınır.

Tanımda üç noktaya dikkat edelim.

- Genel formülde $h(X, Y)$ 'nin ilk bileşenine X değil x yazılır. Nedeni açıktır: $X = x$ koşulu altında X artık rastgele değildir, x sayısına eşittir; rastgele kalan tek şey Y 'dir. $h(x, y) = y$ seçimi $E(Y | X = x)$ 'i, $h(x, y) = g(y)$ seçimi $E(g(Y) | X = x)$ 'i verir.
- $E(Y | X = x)$ bir **sayıdır** ve bu sayı x 'e bağlıdır. $x \mapsto E(Y | X = x)$ fonksiyonuna Y 'nin X üzerine **regresyon fonksiyonu** da denir.
- $f_X(x) = 0$ olan noktalardaki değer önemsizdir; bu noktalar ileride göreceğimiz bütün toplamlara ve integrallere sıfır katkı verir.

i Koşullu beklenen değer sıradan bir beklenen değerdir

$E(Y | X = x)$, olasılık fonksiyonu $f_{Y|X=x}$ olan bir dağılıma göre alınmış, **Tanım 16.1** anlamında sıradan bir beklenen değerdir. Değişen tek şey, kullanılan dağılımın Y 'nin marjinal dağılımı f_Y değil, $X = x$ bilgisiyle güncellenmiş koşullu dağılımı $f_{Y|X=x}$ olmasıdır. Bu yüzden beklenen değerini bildiğimiz bütün özellikleri (lineerlik, monotonluk, sabitin beklenen değerinin kendisi olması) koşullu beklenen değer için de geçerlidir. Bunları bir sonraki alt bölümde açıkça yazacağız.

İlk örneğimiz, koşullu dağılımlarını **Örnek 14.2**'de incelediğimiz üç para atışı deneyidir.

Örnek 19.1 (Üç Para Atışında Koşullu Beklenen Değer). Düzgün bir para üç kez atılıyor. X_1 üç atışta gelen tura sayısı, X_2 ilk iki atışta gelen tura sayısı olsun. (X_1, X_2) rastgele vektörünün olasılık fonksiyonu tablosu şudur:

$x_1 \setminus x_2$	0	1	2	$f_{X_1}(x_1)$
0	$\frac{1}{8}$	0	0	$\frac{1}{8}$
1	$\frac{1}{8}$	$\frac{2}{8}$	0	$\frac{3}{8}$
2	0	$\frac{2}{8}$	$\frac{1}{8}$	$\frac{3}{8}$
3	0	0	$\frac{1}{8}$	$\frac{1}{8}$
$f_{X_2}(x_2)$	$\frac{2}{8}$	$\frac{4}{8}$	$\frac{2}{8}$	1

Her $x_2 \in \{0, 1, 2\}$ için $E(X_1 | X_2 = x_2)$ ve her $x_1 \in \{0, 1, 2, 3\}$ için $E(X_2 | X_1 = x_1)$ koşullu beklenen değerlerini hesaplayınız.

Çözüm.

Adım 1: $X_2 = x_2$ verilmişken X_1 'in koşullu dağılımı. **Örnek 14.2**'de olduğu gibi, tablonun x_2 sütununu o sütunun toplamı $f_{X_2}(x_2)$ 'ye bölelim.

$x_2 = 0$ için sütun toplamı $\frac{2}{8}$ 'dir ve

$$f_{X_1|X_2=0}(0) = \frac{1/8}{2/8} = \frac{1}{2}, \quad f_{X_1|X_2=0}(1) = \frac{1/8}{2/8} = \frac{1}{2}.$$

$x_2 = 1$ için sütun toplamı $\frac{4}{8}$ 'dir ve

$$f_{X_1|X_2=1}(1) = \frac{2/8}{4/8} = \frac{1}{2}, \quad f_{X_1|X_2=1}(2) = \frac{2/8}{4/8} = \frac{1}{2}.$$

$x_2 = 2$ için sütun toplamı $\frac{2}{8}$ 'dir ve

$$f_{X_1|X_2=2}(2) = \frac{1}{2}, \quad f_{X_1|X_2=2}(3) = \frac{1}{2}.$$

Sonuç sezgiyle uyumludur: ilk iki atışta x_2 tura geldiyse, üçüncü atışın sonucuna göre toplam tura sayısı x_2 ya da $x_2 + 1$ olur ve ikisi eşit olasılıklıdır.

Adım 2: $E(X_1 | X_2 = x_2)$. **Tanım 19.1'e** göre

$$E(X_1 | X_2 = 0) = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{1}{2},$$

$$E(X_1 | X_2 = 1) = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{2} = \frac{3}{2},$$

$$E(X_1 | X_2 = 2) = 2 \cdot \frac{1}{2} + 3 \cdot \frac{1}{2} = \frac{5}{2}.$$

Üç sonuç tek bir formülde toplanır:

$$E(X_1 | X_2 = x_2) = x_2 + \frac{1}{2}, \quad x_2 \in \{0, 1, 2\}.$$

Sözle: ilk iki atıştaki tura sayısına, üçüncü atıştan beklenen $\frac{1}{2}$ tura eklenir.

Adım 3: $X_1 = x_1$ verilmişken X_2 'nin koşullu dağılımı. Bu kez tablonun x_1 satırını satır toplamı $f_{X_1}(x_1)$ 'e böleriz.

- $x_1 = 0$: tek olası değer $x_2 = 0$; $f_{X_2|X_1=0}(0) = 1$.
- $x_1 = 1$: satır toplamı $\frac{3}{8}$; $f_{X_2|X_1=1}(0) = \frac{1/8}{3/8} = \frac{1}{3}$, $f_{X_2|X_1=1}(1) = \frac{2/8}{3/8} = \frac{2}{3}$.
- $x_1 = 2$: satır toplamı $\frac{3}{8}$; $f_{X_2|X_1=2}(1) = \frac{2}{3}$, $f_{X_2|X_1=2}(2) = \frac{1}{3}$.
- $x_1 = 3$: tek olası değer $x_2 = 2$; $f_{X_2|X_1=3}(2) = 1$.

Adım 4: $E(X_2 | X_1 = x_1)$.

$$E(X_2 | X_1 = 0) = 0,$$

$$E(X_2 | X_1 = 1) = 0 \cdot \frac{1}{3} + 1 \cdot \frac{2}{3} = \frac{2}{3},$$

$$E(X_2 | X_1 = 2) = 1 \cdot \frac{2}{3} + 2 \cdot \frac{1}{3} = \frac{4}{3},$$

$$E(X_2 | X_1 = 3) = 2.$$

Dört sonuç yine tek bir formüle sığar:

$$E(X_2 | X_1 = x_1) = \frac{2x_1}{3}, \quad x_1 \in \{0, 1, 2, 3\}.$$

Bunun da sezgisel bir okuması vardır: üç atışta x_1 tura geldiği biliniyorsa, bu turalar üç atış arasında eşit olasılıkla dağılmıştır; ilk iki atışa düşen pay $\frac{2}{3}x_1$ 'dir.

Karşılaştırma için marjinal beklenen değerleri de yazalım:

$$E(X_1) = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{12}{8} = \frac{3}{2}, \quad E(X_2) = 0 \cdot \frac{2}{8} + 1 \cdot \frac{4}{8} + 2 \cdot \frac{2}{8} = 1.$$

Koşullu beklenen değerler, koşul değiştikçe bu marjinal değerlerin etrafında dolaşır: $E(X_1 | X_2 = x_2)$ değerleri $\frac{1}{2}, \frac{3}{2}, \frac{5}{2}$ iken $E(X_1) = \frac{3}{2}$ 'dir. Bu "etrafında dolaşma"nın kesin ifadesi, toplam beklenen değer kuralı olacaktır.

■

Sürekli durumda hesap, integrallerle aynı adımları izler.

Örnek 19.2 (Üçgen Üzerinde Düzgün Dağılımda Koşullu Beklenen Değer). (X, Y) rastgele vektörünün ortak olasılık yoğunluk fonksiyonu

$$f_{X,Y}(x,y) = \begin{cases} 2, & 0 < x < y < 1 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olsun. $0 < x < 1$ için $E(Y | X = x)$ ve $0 < y < 1$ için $E(X | Y = y)$ koşullu beklenen değerlerini bulunuz.

Çözüm.

Adım 1: Marjinal yoğunluklar. Tanım bölgesi, köşeleri $(0, 0)$, $(0, 1)$ ve $(1, 1)$ olan üçgendir. Tanım 13.8'a göre, $0 < x < 1$ için

$$f_X(x) = \int_x^1 2 dy = 2(1 - x),$$

ve $0 < y < 1$ için

$$f_Y(y) = \int_0^y 2 dx = 2y.$$

Bu aralıkların dışında iki marjinal de sıfırdır.

Adım 2: $X = x$ verilmişken Y 'nin koşullu yoğunluğu. $0 < x < 1$ için $f_X(x) = 2(1 - x) > 0$ olduğundan Tanım 14.3 uygulanır:

$$f_{Y|X=x}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \begin{cases} \frac{2}{2(1-x)} = \frac{1}{1-x}, & x < y < 1 \\ 0, & \text{diğer durumlarda.} \end{cases}$$

Yani $X = x$ verilmişken Y , $(x, 1)$ aralığında düzgün dağılmıştır. Bunun beklenen değeri aralığın orta noktasıdır; doğrudan hesapla da görelim:

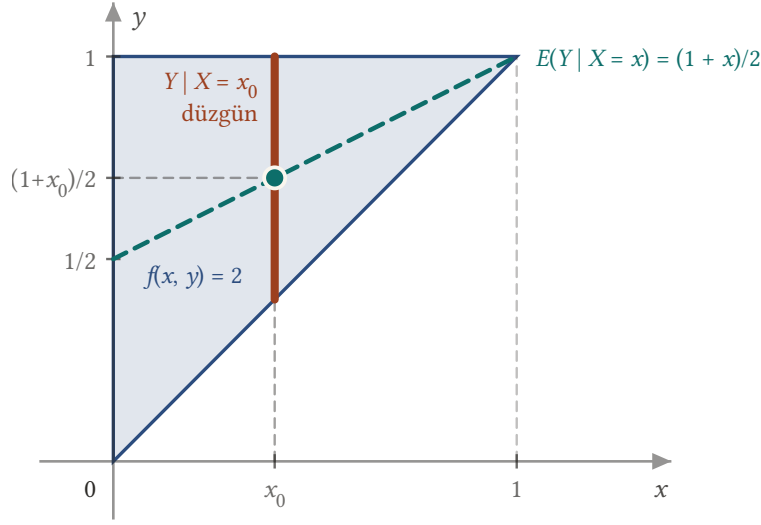
$$E(Y | X = x) = \int_x^1 y \cdot \frac{1}{1-x} dy = \frac{1}{1-x} \cdot \frac{y^2}{2} \Big|_x^1 = \frac{1-x^2}{2(1-x)} = \frac{1+x}{2}.$$

Adım 3: $Y = y$ verilmişken X 'in koşullu yoğunluğu. $0 < y < 1$ için $f_Y(y) = 2y > 0$ ve

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \begin{cases} \frac{2}{2y} = \frac{1}{y}, & 0 < x < y \\ 0, & \text{diğer durumlarda.} \end{cases}$$

Bu kez X , $(0, y)$ aralığında düzgün dağılmıştır:

$$E(X | Y = y) = \int_0^y x \cdot \frac{1}{y} dx = \frac{1}{y} \cdot \frac{y^2}{2} = \frac{y}{2}.$$



Şekil 19.1: Üçgen üzerinde düzgün dağılımın koşullu ortalaması. Taralı üçgen, $f(x, y) = 2$ yoğunluğunun tanım bölgesidir. $X = x_0$ verilmişken Y yalnızca kalın düşey kesit üzerinde, $(x_0, 1)$ aralığında düzgün dağılır; koşullu beklenen değer bu kesitin orta noktası $(1 + x_0)/2$ 'dir. x_0 değiştikçe orta noktalar kesikli $y = (1 + x)/2$ doğrusunu çizer: $(0, 1/2)$ 'den $(1, 1)$ 'e uzanan regresyon doğrusu.

Her iki regresyon fonksiyonu da doğrusaldır: $E(Y | X = x) = \frac{1+x}{2}$ doğrusu üçgenin düşey kesitlerinin orta noktalarını, $E(X | Y = y) = \frac{y}{2}$ doğrusu yatay kesitlerin orta noktalarını birleştirir. İki doğru farklıdır; genel olarak Y 'nin X üzerine regresyonu ile X 'in Y üzerine regresyonu aynı fonksiyonun tersi değildir.

■

19.2 Koşullu Beklenen Değerin Özellikleri

Koşullu beklenen değer bir dağılıma göre alınmış beklenen değer olduğundan, beklenen değer lineerlik gibi özelliklerini taşıması beklenir. Ancak koşullama, sıradan beklenen değerde bulunmayan iki özellik daha getirir: X 'e bağlı çarpanlar dışarı alınabilir; bağımsızlık ise koşullamayı etkisiz kılar. Bu dört özelliği bir arada yazıp ispatlayalım; bölümün geri kalanında sürekli kullanılacaklar.

Önerme 19.1 (Koşullu Beklenen Değerin Özellikleri). (X, Y) bir rastgele vektör, x noktası $f_X(x) > 0$ olacak biçimde sabit, $a, b, c \in \mathbb{R}$ ve g, h Borel ölçülebilir fonksiyonlar olsun. Aşağıdaki koşullu beklenen değerlerin var olduğu varsayalım.

(a) **Sabit:** $E(c | X = x) = c$.

(b) **Lineerlik:**

$$E(a g(X, Y) + b h(X, Y) | X = x) = a E(g(X, Y) | X = x) + b E(h(X, Y) | X = x).$$

(c) **Çarpanı dışarı alma:** $E(g(X) h(X, Y) | X = x) = g(x) E(h(X, Y) | X = x)$; özel olarak $E(g(X) h(Y) | X = x) = g(x) E(h(Y) | X = x)$ ve $E(g(X) | X = x) = g(x)$.

(d) **Bağımsızlık:** X ve Y bağımsız ise $E(h(Y) | X = x) = E(h(Y))$; özel olarak $E(Y | X = x) = E(Y)$.

İspat.

Önce kesikli durumu yazalım; her toplam $y \in D_{Y|X=x}$ üzerindedir ve bu toplamalar mutlak yakınsak olduğundan terimleri ayırıp yeniden düzenleyebiliriz.

(a) $f_{Y|X=x}$ bir olasılık fonksiyonu olduğundan toplamı 1'dir:

$$E(c | X = x) = \sum_y c f_{Y|X=x}(y) = c \sum_y f_{Y|X=x}(y) = c \cdot 1 = c.$$

(b) *Tanım 19.1'deki genel formülü $a g + b h$ fonksiyonuna uygulayalım:*

$$\begin{aligned} E (a g(X, Y) + b h(X, Y) \mid X = x) &= \sum_y [a g(x, y) + b h(x, y)] f_{Y|X=x}(y) \\ &= a \sum_y g(x, y) f_{Y|X=x}(y) + b \sum_y h(x, y) f_{Y|X=x}(y) \\ &= a E (g(X, Y) \mid X = x) + b E (h(X, Y) \mid X = x) . \end{aligned}$$

(c) Genel formülde h yerine $(x, y) \mapsto g(x) h(x, y)$ fonksiyonunu koyalım. Toplam y üzerinden alındığından $g(x)$ toplama göre bir sabittir:

$$E (g(X) h(X, Y) \mid X = x) = \sum_y g(x) h(x, y) f_{Y|X=x}(y) = g(x) \sum_y h(x, y) f_{Y|X=x}(y) = g(x) E (h(X, Y) \mid X = x) .$$

$h(x, y) = h(y)$ seçimi ilk özel durumu verir: $h \equiv 1$ seçimi ve (a) ise $E(g(X) \mid X = x) = g(x) \cdot 1 = g(x)$ verir.

(d) X ve Y bağımsız olsun. *Tanım 14.1'e göre her (x, y) için $f_{X,Y}(x, y) = f_X(x) f_Y(y)$ 'dir; dolayısıyla $f_X(x) > 0$ iken*

$$f_{Y|X=x}(y) = \frac{f_{X,Y}(x, y)}{f_X(x)} = \frac{f_X(x) f_Y(y)}{f_X(x)} = f_Y(y).$$

Koşullu dağılım marjinal dağılımla aynıdır; öyleyse

$$E (h(Y) \mid X = x) = \sum_y h(y) f_{Y|X=x}(y) = \sum_{y \in D_Y} h(y) f_Y(y) = E (h(Y))$$

olur; son eşitlik *Tanım 16.1'dir. ($D_{Y|X=x} = D_Y$ olduğuna dikkat edin: bağımsızlıkta $f_{X,Y}(x, y) > 0 \Leftrightarrow f_Y(y) > 0$ 'dır.)* $h(y) = y$ seçimi $E(Y \mid X = x) = E(Y)$ verir.

Şimdi sürekli durum. Bütün integraller $\mathbb{R}$ üzerinden ve mutlak yakınsaktır; adımlar birebir aynıdır, toplamın yerini integral alır.

(a) Koşullu yoğunluğun integrali 1 olduğundan sabit dışarı alınır:

$$E(c \mid X = x) = \int_{-\infty}^{\infty} c f_{Y|X=x}(y) dy = c \int_{-\infty}^{\infty} f_{Y|X=x}(y) dy = c.$$

(b) İntegralin lineerliğiyle

$$\int_{-\infty}^{\infty} [a g(x, y) + b h(x, y)] f_{Y|X=x}(y) dy = a \int_{-\infty}^{\infty} g(x, y) f_{Y|X=x}(y) dy + b \int_{-\infty}^{\infty} h(x, y) f_{Y|X=x}(y) dy.$$

(c) $g(x)$, y 'ye göre alınan integralde sabittir:

$$\int_{-\infty}^{\infty} g(x) h(x, y) f_{Y|X=x}(y) dy = g(x) \int_{-\infty}^{\infty} h(x, y) f_{Y|X=x}(y) dy.$$

(d) Bağımsızlıkta $f_{X,Y}(x, y) = f_X(x) f_Y(y)$ olduğundan $f_X(x) > 0$ iken $f_{Y|X=x}(y) = f_Y(y)$ 'dir ve

$$E (h(Y) \mid X = x) = \int_{-\infty}^{\infty} h(y) f_Y(y) dy = E (h(Y)) .$$

■

Özellik (c), koşullu beklenen değer en çok kullanılan özelliğidir. Sözle: **$X = x$ koşulu altında X 'in fonksiyonları sabit gibi davranır ve beklenen değer dışına çıkar.** Özellik (d) ise koşullamanın ne zaman işe yaramadığını söyler: X, Y hakkında bilgi taşıyorsa X 'i bilmek Y için beklentimizi değiştirmez.

⚠ Ters yön doğru değildir

Özellik (d)'nin tersi genel olarak yanlıştır: her x için $E(Y | X = x) = E(Y)$ olması X ile Y 'nin bağımsız olmasını gerektirmez. $Y, \{-1, 0, 1\}$ üzerinde düzgün dağılsın ve $X = Y^2$ olsun. $X = 1$ verilmişken $Y, \{-1, 1\}$ üzerinde düzgün dağılır ve $E(Y | X = 1) = 0$; $X = 0$ verilmişken $Y = 0$ kesindir ve $E(Y | X = 0) = 0$. Yani her x için $E(Y | X = x) = 0 = E(Y)$ 'dir. Oysa X ile Y bağımsız değildir: $P(Y = 0, X = 1) = 0$ ama $P(Y = 0)P(X = 1) = \frac{1}{3} \cdot \frac{2}{3} \neq 0$. Koşullu beklenen değer sabit olması, koşullu **dağılımın** sabit olmasından çok daha zayıf bir bilgidir.

19.3 Rastgele Değişken Olarak Koşullu Beklenen Değer

$E(Y | X = x)$ sayısı x 'e bağlıdır; yani elimizde $m(x) = E(Y | X = x)$ ile tanımlı bir $m : \mathbb{R} \rightarrow \mathbb{R}$ fonksiyonu vardır. Bu fonksiyonu x noktasında değil, rastgele değişken X 'te değerlendirirsek yeni bir rastgele değişken elde ederiz. Bu küçük gösterim değişikliği, bölümün ana teoremini ifade etmemizi sağlar.

Tanım 19.2 (Koşullu Beklenen Değer Rastgele Değişkeni). (X, Y) bir rastgele vektör ve $f_X(x) > 0$ olan her x için $E(Y | X = x)$ var olsun. $m(x) = E(Y | X = x)$ ile tanımlanan $m : \mathbb{R} \rightarrow \mathbb{R}$ fonksiyonu için

$$E(Y | X) = m(X)$$

rastgele değişkenine, X verilmişken Y 'nin **koşullu beklenen değeri** denir. Aynı biçimde, h Borel ölçülebilir ve $m_h(x) = E(h(X, Y) | X = x)$ olmak üzere $E(h(X, Y) | X) = m_h(X)$ yazılır.

$E(Y | X)$, X 'in bir fonksiyonu olduğundan bir rastgele değişkendir ve dağılımı X 'in dağılımıyla belirlenir. Örneklerimizde:

- Üç para atışında (Örnek 19.1) $E(X_1 | X_2 = x_2) = x_2 + \frac{1}{2}$ bulmuştuk; öyleyse $E(X_1 | X_2) = X_2 + \frac{1}{2}$ ve benzer biçimde $E(X_2 | X_1) = \frac{2}{3}X_1$ 'dir.
- Üçgen örneğinde (Örnek 19.2) $E(Y | X) = \frac{1+X}{2}$ ve $E(X | Y) = \frac{Y}{2}$ 'dir.

Bu rastgele değişkenin en önemli özelliği, ortalamasının Y 'nin ortalaması olmasıdır. Sezgisel gerekçe şudur: $E(Y | X = x)$, Y 'nin $X = x$ dilimi içindeki ortalamasıdır; bu dilim ortalamalarını dilimlerin olasılıklarıyla ağırlıklandırıp toplamak Y 'nin bütün ortalamasını verir. Toplam olasılık formülünde olayları X 'in değerlerine göre parçalayarak $P(A) = \sum_x P(A | X = x) P(X = x)$ yazmıştık; şimdi aynı parçalamayı beklenen değer için yapıyoruz.

Teorem 19.1 (Toplam Beklenen Değer Kuralı (Kule Kuralı)). (X, Y) bir rastgele vektör ve $h : \mathbb{R}^2 \rightarrow \mathbb{R}$, $E(|h(X, Y)|) < \infty$ olan Borel ölçülebilir bir fonksiyon olsun. O zaman

- (i) $E[E(h(X, Y) | X)] = E(h(X, Y))$; özel olarak $E[E(Y | X)] = E(Y)$.
- (ii) $g, E(|g(X)h(X, Y)|) < \infty$ olan bir fonksiyon ise $E[g(X)h(X, Y)] = E[g(X)E(h(X, Y) | X)]$; özel olarak $E[g(X)Y] = E[g(X)E(Y | X)]$.

İspat.

$m(x) = E(h(X, Y) | X = x)$ olsun; $f_X(x) = 0$ olan noktalarda $m(x) = 0$ 'dır. $E(h(X, Y) | X) = m(X)$, X 'in bir fonksiyonu olduğundan beklenen değeri Tanım 16.1 ile X 'in dağılımı üzerinden hesaplanır.

Kesikli durum, (i). $x \in D_X$ için $f_X(x) > 0$ 'dır ve

$$E[m(X)] = \sum_{x \in D_X} m(x) f_X(x) = \sum_{x \in D_X} \left(\sum_{y \in D_{Y|X=x}} h(x, y) f_{Y|X=x}(y) \right) f_X(x).$$

$f_X(x)$, iç toplama göre sabittir; içeri alıp $f_{Y|X=x}(y) f_X(x) = f_{X,Y}(x, y)$ eşitliğini (koşullu olasılık fonksiyonunun tanımı) kullanalım:

$$E[m(X)] = \sum_{x \in D_X} \sum_{y \in D_{Y|X=x}} h(x, y) f_{X,Y}(x, y).$$

$x \in D_X$ ve $y \in D_{Y|X=x}$ koşulu tam olarak $f_{X,Y}(x,y) > 0$, yani $(x,y) \in D_{(X,Y)}$ demektir. Öyleyse çift toplam, (X,Y) 'nin değer kümesi üzerinden alınan toplamdır ve Tanım 18.1'e göre

$$E[m(X)] = \sum_{(x,y) \in D_{(X,Y)}} h(x,y) f_{X,Y}(x,y) = E(h(X,Y))$$

olur. Toplamın sırasını serbestçe değiştirebildik; çünkü $E(|h(X,Y)|) < \infty$ varsayımı çift toplamın mutlak yakınsak olduğunu söyler. $h(x,y) = y$ seçimi $E[E(Y|X)] = E(Y)$ verir.

Kesikli durum, (ii). (i)'i $\tilde{h}(x,y) = g(x)h(x,y)$ fonksiyonuna uygulayalım. Önerme 19.1 (c)'ye göre

$$E(\tilde{h}(X,Y) | X = x) = E(g(X)h(X,Y) | X = x) = g(x)m(x),$$

yani $E(\tilde{h}(X,Y) | X) = g(X)m(X)$ 'tir. (i) gereği

$$E[g(X)h(X,Y)] = E(\tilde{h}(X,Y)) = E[E(\tilde{h}(X,Y) | X)] = E[g(X)m(X)] = E[g(X)E(h(X,Y) | X)].$$

Süreklilik durumu, (i). $f_X(x) > 0$ olan x noktaları için $m(x) = \int_{-\infty}^{\infty} h(x,y) f_{Y|X=x}(y) dy$ 'dir; $f_X(x) = 0$ olan noktalar aşağıdaki dış integrale katkı vermez. Tanım 16.1'e göre

$$E[m(X)] = \int_{-\infty}^{\infty} m(x) f_X(x) dx = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} h(x,y) f_{Y|X=x}(y) dy \right) f_X(x) dx.$$

$f_X(x)$ 'i iç integrale alıp $f_{Y|X=x}(y) f_X(x) = f_{X,Y}(x,y)$ yazalım; bu eşitlik $f_X(x) > 0$ iken koşullu yoğunluğun tanımıdır. $f_X(x) = 0$ olan x noktalarında ise $m(x) f_X(x) = 0$ 'dır ve $\int f_{X,Y}(x,y) dy = f_X(x) = 0$ olduğundan $\int h(x,y) f_{X,Y}(x,y) dy = 0$ 'dır; bu noktalar iki tarafa da katkı vermez. Böylece

$$E[m(X)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x,y) f_{X,Y}(x,y) dy dx = E(h(X,Y));$$

son eşitlik Tanım 18.1'dir. İntegrallerin sırasını değiştirmeyi $E(|h(X,Y)|) < \infty$ varsayımı, yani $\iint |h| f_{X,Y}$ integralinin sonlu olması sağlar (Fubini teoremi).

Süreklilik durumu, (ii). Kesikli durumdaki akıl yürütme aynen geçerlidir: Önerme 19.1 (c) ile $E(g(X)h(X,Y) | X) = g(X)m(X)$ ve (i) ile $E[g(X)h(X,Y)] = E[g(X)m(X)]$ olur.

■

Teoremin (i) kısmına “kule kuralı” denmesinin nedeni, iç içe iki beklenen değer bir kule gibi üst üste durmasıdır: içteki $E(\cdot | X)$ önce Y üzerinden, dıştaki E sonra X üzerinden ortalama alır. (ii) kısmı ise şunu söyler: bir beklenen değer içinde X 'in fonksiyonunun yanındaki Y , hiçbir şey değişmeden $E(Y | X)$ ile değiştirilebilir. Bu, bölüm sonunda en iyi tahmin sonucunu verecek olan araçtır.

💡 Kule kuralının hesap biçimi

Uygulamada kule kuralı çoğu zaman açık biçimiyle kullanılır:

$$E(Y) = \sum_{x \in D_X} E(Y | X = x) f_X(x) \quad \text{ya da} \quad E(Y) = \int_{-\infty}^{\infty} E(Y | X = x) f_X(x) dx.$$

Bu, toplam olasılık formülünün beklenen değer karşılığıdır: Y 'nin ortalaması, X 'in her değerine karşılık gelen dilim ortalamalarının X 'in dağılımıyla ağırlıklı ortalamasıdır. Kural özellikle Y 'nin marjinal dağılımı karmaşık, ama $X = x$ verildiğinde Y 'nin dağılımı basit olduğunda güçlüdür: $E(Y)$ marjinal hiç hesaplanmadan bulunur.

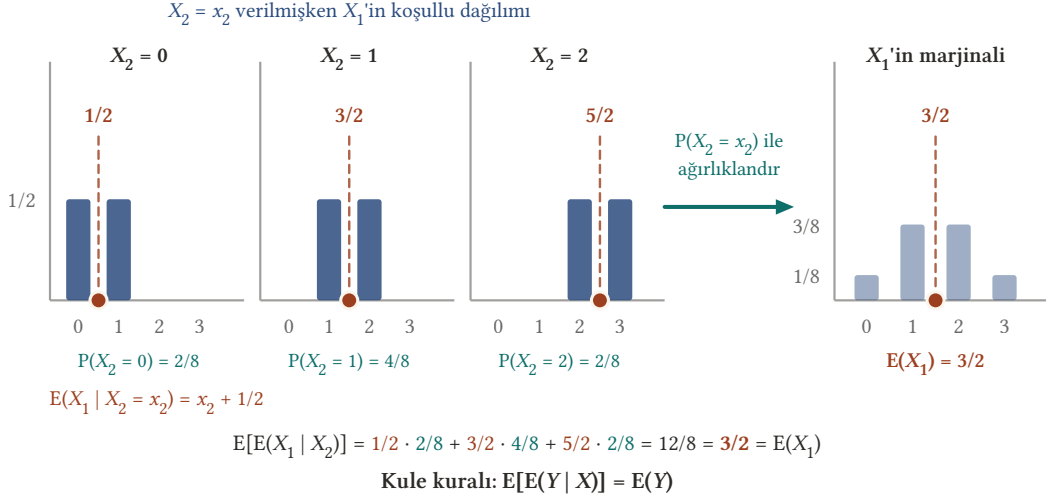
Örnek 19.3 (Kule Kuralının Doğrulanması). Örnek 19.1 ve Örnek 19.2'de bulunan koşullu beklenen değerlerden yola çıkarak kule kuralıyla $E(X_1)$, $E(X_2)$, $E(Y)$ ve $E(X)$ 'i hesaplayınız; sonuçları marjinal dağılımlardan bulunan değerlerle karşılaştırınız.

Çözüm.

Üç para atışı. $E(X_1 | X_2 = x_2) = x_2 + \frac{1}{2}$ ve f_{X_2} değerleri $\frac{2}{8}, \frac{4}{8}, \frac{2}{8}$ idi. Kule kuralına göre

$$E(X_1) = \sum_{x_2=0}^2 E(X_1 | X_2 = x_2) f_{X_2}(x_2) = \frac{1}{2} \cdot \frac{2}{8} + \frac{3}{2} \cdot \frac{4}{8} + \frac{5}{2} \cdot \frac{2}{8} = \frac{1+6+5}{8} = \frac{12}{8} = \frac{3}{2}.$$

Aynı sonuca daha kısa yoldan, rastgele değişken biçimini kullanarak da varılır: $E(X_1) = E[E(X_1 | X_2)] = E(X_2 + \frac{1}{2}) = E(X_2) + \frac{1}{2} = 1 + \frac{1}{2} = \frac{3}{2}$; burada Teorem 16.2 kullanıldı. Marjinalden bulduğumuz $E(X_1) = \frac{3}{2}$ ile uyuşur.



Şekil 19.2: Kule kuralı üç para atışı örneğinde. Solda, ilk iki atıştaki tura sayısı $X_2 = x_2$ verilmişken toplam tura sayısı X_1 , x_2 ve $x_2 + 1$ değerlerini eşit olasılıkla alır; dilim ortalaması $E(X_1 | X_2 = x_2) = x_2 + 1/2$ ekseninde turuncu noktayla gösterilmiştir. Bu üç dilim ortalaması $P(X_2 = x_2) = 2/8, 4/8, 2/8$ ağırlıklarıyla toplanınca sağdaki marjinal dağılımın ortalaması $E(X_1) = 3/2$ çıkar: $E[E(X_1 | X_2)] = E(X_1)$.

Öteki yönde $E(X_2 | X_1) = \frac{2}{3}X_1$ olduğundan

$$E(X_2) = E[E(X_2 | X_1)] = E\left(\frac{2}{3}X_1\right) = \frac{2}{3} \cdot \frac{3}{2} = 1,$$

marjinalden bulunan $E(X_2) = 1$ ile uyuşur.

Üçgen örneği. $E(Y | X = x) = \frac{1+x}{2}$ ve $f_X(x) = 2(1-x)$, $0 < x < 1$. Kule kuralı:

$$E(Y) = \int_0^1 \frac{1+x}{2} \cdot 2(1-x) dx = \int_0^1 (1-x^2) dx = 1 - \frac{1}{3} = \frac{2}{3}.$$

Doğrudan hesap: $E(Y) = \int_0^1 y \cdot 2y dy = \frac{2}{3}$. Uyuşur.

$E(X | Y = y) = \frac{y}{2}$ ve $f_Y(y) = 2y$, $0 < y < 1$ olduğundan

$$E(X) = \int_0^1 \frac{y}{2} \cdot 2y dy = \int_0^1 y^2 dy = \frac{1}{3}.$$

Doğrudan hesap: $E(X) = \int_0^1 x \cdot 2(1-x) dx = 1 - \frac{2}{3} = \frac{1}{3}$. Uyuşur.

■

19.4 Koşullu Varyans ve Toplam Varyans Kuralı

Koşullu dağılım gerçek bir dağılım olduğuna göre yalnızca beklenen değeri değil, varyansı da vardır. $X = x$ bilgisi Y 'nin ortalamasını $E(Y | X = x)$ 'e taşıdığı gibi, Y 'nin yayılımını da değiştirir; çoğu zaman azaltır.

Tanım 19.3 (Koşullu Varyans). (X, Y) bir rastgele vektör ve x , $f_X(x) > 0$ ve $E(Y^2 | X = x) < \infty$ olan bir nokta olsun.

$$\text{Var} (Y | X = x) = E \left([Y - E(Y | X = x)]^2 \middle| X = x \right)$$

sayısına, $X = x$ verilmişken Y 'nin **koşullu varyansı** denir. $v(x) = \text{Var} (Y | X = x)$ olmak üzere $\text{Var} (Y | X) = v(X)$ rastgele değişkenine, X verilmişken Y 'nin koşullu varyansı denir.

Koşullu varyans, tıpkı sıradan varyans gibi, ikinci momentle ortalamanın karesinin farkı olarak hesaplanır. $\mu(x) = E(Y | X = x)$ yazalım; bu, $X = x$ koşulu altında bir sabittir. **Önerme 19.1**'nin (a) ve (b) özellikleriyle

$$\begin{aligned} \text{Var} (Y | X = x) &= E (Y^2 - 2\mu(x)Y + \mu(x)^2 | X = x) \\ &= E(Y^2 | X = x) - 2\mu(x) E(Y | X = x) + \mu(x)^2 \\ &= E(Y^2 | X = x) - 2\mu(x)^2 + \mu(x)^2 = E(Y^2 | X = x) - [E(Y | X = x)]^2. \end{aligned}$$

Rastgele değişken biçiminde:

$$\text{Var} (Y | X) = E(Y^2 | X) - [E(Y | X)]^2.$$

Örneğin üç para atışında $X_2 = x_2$ verilmişken X_1 , x_2 ve $x_2 + 1$ değerlerini her biri $\frac{1}{2}$ olasılıkla alır; $E(X_1^2 | X_2 = x_2) = \frac{x_2^2 + (x_2+1)^2}{2} = x_2^2 + x_2 + \frac{1}{2}$ ve

$$\text{Var} (X_1 | X_2 = x_2) = x_2^2 + x_2 + \frac{1}{2} - \left(x_2 + \frac{1}{2}\right)^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

olur; koşullu varyans x_2 'ye bağlı değildir. Üçgen örneğinde ise $X = x$ verilmişken Y , uzunluğu $1 - x$ olan bir aralıkta düzgün dağılır; aşağıdaki örnekte $\text{Var} (Y | X = x) = \frac{(1-x)^2}{12}$ bulacağız: x büyüdükçe Y 'nin belirsizliği azalır.

Kule kuralı beklenen değeri iki aşamaya bölüyordu. Varyans için benzer bir bölme vardır, ama bir fark çıkar: Y 'nin toplam yayılımı, dilimlerin içindeki yayılımın ortalaması **artı** dilim ortalamalarının kendi aralarındaki yayılımıdır.

Teorem 19.2 (Toplam Varyans Kuralı). (X, Y) bir rastgele vektör ve $E(Y^2) < \infty$ olsun. O zaman

$$\text{Var} (Y) = E [\text{Var} (Y | X)] + \text{Var} [E(Y | X)].$$

İspat.

$\text{Var} (Y | X) = E(Y^2 | X) - [E(Y | X)]^2$ eşitliğinin iki tarafının beklenen değerini alalım. **Teorem 16.1** ve kule kuralı (**Teorem 19.1**, $h(x, y) = y^2$ ile) gereği

$$E [\text{Var} (Y | X)] = E [E(Y^2 | X)] - E ([E(Y | X)]^2) = E(Y^2) - E ([E(Y | X)]^2).$$

Öte yandan $E(Y | X)$ bir rastgele değişkendir ve varyansı **Önerme 16.3** ile, ardından kule kuralı ($E[E(Y | X)] = E(Y)$) ile

$$\text{Var} [E(Y | X)] = E ([E(Y | X)]^2) - (E [E(Y | X)])^2 = E ([E(Y | X)]^2) - [E(Y)]^2$$

olur. İki eşitliği taraf tarafa toplayalım; $E ([E(Y | X)]^2)$ terimleri birbirini götürür:

$$E [\text{Var} (Y | X)] + \text{Var} [E(Y | X)] = E(Y^2) - [E(Y)]^2 = \text{Var} (Y).$$

Son eşitlik yine **Önerme 16.3**'dür. $E(Y^2) < \infty$ varsayımı, bütün bu beklenen değerlerin sonlu olmasını sağlar. ■

Kuralın iki terimi ayrı ayrı anlamlıdır. $E[\text{Var} (Y | X)]$, **açıklanamayan** yayılımdır: X bilinse bile Y 'de kalan belirsizliğin ortalaması. $\text{Var} [E(Y | X)]$ ise **açıklanan** yayılımdır: Y 'nin ortalamasının X değişikçe ne kadar oynadığı. İki terim de negatif olmadığından iki sonuç hemen çıkar:

- $\text{Var} [E(Y | X)] \leq \text{Var} (Y)$: koşullu beklenen değer, Y 'den daha az yayılmış bir rastgele değişkendir.

- $E[\text{Var}(Y | X)] \leq \text{Var}(Y)$: X 'i bilmek, ortalamada, Y hakkındaki belirsizliği artırmaz.

Örnek 19.4 (Toplam Varyans Kuralının Doğrulanması). Üç para atışında $\text{Var}(X_1)$ 'i ve üçgen örneğinde $\text{Var}(Y)$ 'yi toplam varyans kuralıyla hesaplayınız; sonuçları marjinal dağılımlardan bulunan değerlerle karşılaştırınız.

Çözüm.

Üç para atışı. Yukarıda $\text{Var}(X_1 | X_2 = x_2) = \frac{1}{4}$ bulmuştuk; sabit olduğundan $E[\text{Var}(X_1 | X_2)] = \frac{1}{4}$. Öte yandan $E(X_1 | X_2) = X_2 + \frac{1}{2}$ olduğundan **Teorem 16.2** ile

$$\text{Var}[E(X_1 | X_2)] = \text{Var}\left(X_2 + \frac{1}{2}\right) = \text{Var}(X_2).$$

$E(X_2) = 1$ ve $E(X_2^2) = 0 \cdot \frac{2}{8} + 1 \cdot \frac{4}{8} + 4 \cdot \frac{2}{8} = \frac{12}{8} = \frac{3}{2}$ olduğundan $\text{Var}(X_2) = \frac{3}{2} - 1 = \frac{1}{2}$. Toplam varyans kuralı:

$$\text{Var}(X_1) = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}.$$

Doğrudan hesap: $E(X_1^2) = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 4 \cdot \frac{3}{8} + 9 \cdot \frac{1}{8} = \frac{24}{8} = 3$ ve $\text{Var}(X_1) = 3 - \left(\frac{3}{2}\right)^2 = 3 - \frac{9}{4} = \frac{3}{4}$. Uyuşur.

Üçgen örneği. Önce $\text{Var}(Y | X = x)$ 'i hesaplayalım. $X = x$ verilmişken Y 'nin yoğunluğu $(x, 1)$ üzerinde $\frac{1}{1-x}$ idi:

$$E(Y^2 | X = x) = \int_x^1 \frac{y^2}{1-x} dy = \frac{1-x^3}{3(1-x)} = \frac{1+x+x^2}{3}.$$

Böylece

$$\text{Var}(Y | X = x) = \frac{1+x+x^2}{3} - \left(\frac{1+x}{2}\right)^2 = \frac{4+4x+4x^2-3-6x-3x^2}{12} = \frac{1-2x+x^2}{12} = \frac{(1-x)^2}{12}.$$

Bunun X üzerinden ortalaması, $f_X(x) = 2(1-x)$ ile

$$E[\text{Var}(Y | X)] = \int_0^1 \frac{(1-x)^2}{12} \cdot 2(1-x) dx = \frac{1}{6} \int_0^1 (1-x)^3 dx = \frac{1}{6} \cdot \frac{1}{4} = \frac{1}{24}.$$

İkinci terim için $E(Y | X) = \frac{1+X}{2}$ olduğundan $\text{Var}[E(Y | X)] = \frac{1}{4} \text{Var}(X)$. $E(X) = \frac{1}{3}$ idi ve

$$E(X^2) = \int_0^1 x^2 \cdot 2(1-x) dx = 2\left(\frac{1}{3} - \frac{1}{4}\right) = \frac{1}{6}, \quad \text{Var}(X) = \frac{1}{6} - \frac{1}{9} = \frac{1}{18}.$$

Öyleyse $\text{Var}[E(Y | X)] = \frac{1}{72}$ ve toplam varyans kuralıyla

$$\text{Var}(Y) = \frac{1}{24} + \frac{1}{72} = \frac{3}{72} + \frac{1}{72} = \frac{4}{72} = \frac{1}{18}.$$

Doğrudan hesap: $E(Y^2) = \int_0^1 y^2 \cdot 2y dy = \frac{1}{2}$ ve $\text{Var}(Y) = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}$. Uyuşur.

■

19.5 Rastgele Sayıda Terimin Toplamı

Kule kuralı ve toplam varyans kuralının en tipik uygulaması, terim sayısının kendisi de rastgele olan toplamlardır. Bir mağazaya bir günde gelen müşteri sayısı ve her müşterinin harcaması; bir sigorta şirketine bir ayda gelen hasar talebi sayısı ve her talebin tutarı; bir sunucuya bir saatte gelen istek sayısı ve her isteğin işlem süresi hep bu yapıdadır. Toplamın marjinal dağılımı genellikle karmaşıktır; ama terim sayısı sabitlendiğinde toplam, bildiğimiz sabit sayıda terimli bir toplama dönüşür.

i Koşullayan değişken kesikli, koşullanan değişken herhangi bir dağılıma sahipse

Tanımlarımız ve kule kuralı, (X, Y) 'nin ikisi birden kesikli ya da ikisi birden sürekli olduğu durumu kapsıyor. Aşağıdaki örnekte ise koşullayan değişken N kesiklidir, koşullanan toplam S ise herhangi bir dağılıma sahip (örneğin sürekli) olabilir. Bu karma durumda tanım şöyledir: $P(X = x) > 0$ olan bir x için $P(\cdot | X = x)$ bir olasılık ölçüsüdür (Teorem 6.1); Y 'nin bu ölçüye göre dağılımına $X = x$ verilmişken Y 'nin **koşullu dağılımı**, bu dağılıma göre beklenen değerine de $E(Y | X = x)$ denir. Kule kuralı yine geçerlidir: $\{X = x\}$, $x \in D_X$, olayları Ω 'yı parçaladığından toplam olasılık formülü (Teorem 6.4) her $y \in \mathbb{R}$ için

$$P(Y \leq y) = \sum_{x \in D_X} P(Y \leq y | X = x) P(X = x)$$

verir; yani Y 'nin dağılım fonksiyonu, koşullu dağılım fonksiyonlarının f_X ağırlıklı toplamıdır. Aynı eşitlik, Y kesikli ise olasılık fonksiyonuna, sürekli ise (türev alarak) yoğunluk fonksiyonuna geçer: $f_Y(y) = \sum_{x \in D_X} f_{Y|X=x}(y) P(X = x)$. Şimdi $E(|Y|) < \infty$ olduğunu varsayalım; bu durumda terimleri negatif olmayan $\sum_{x \in D_X} \int |y| f_{Y|X=x}(y) P(X = x) dy$ ifadesi sonludur ve Tonelli-Fubini teoremi toplamla integralin (ya da iki toplamın) yerinin değiştirilmesine izin verir. Öyleyse iki tarafı y ile çarpıp y üzerinden toplayınca ya da integre edince

$$E(Y) = \sum_{x \in D_X} E(Y | X = x) P(X = x) = E[E(Y | X)]$$

çıkar. Toplam varyans kuralının ispatı yalnızca kule kuralını ve koşullu beklenen değer özelliklerini kullandığından o da karma durumda geçerli kalır.

Örnek 19.5 (Rastgele Sayıda Terimin Toplamı). N , değer kümesi $\{0, 1, 2, \dots\}$ olan ve $E(N^2) < \infty$ sağlayan bir rastgele değişken; $X_1, X_2, \dots$ ise bağımsız, aynı dağılımlı, $E(X_i) = \mu$ ve $\text{Var}(X_i) = \sigma^2$ olan rastgele değişkenler olsun. N 'nin $X_1, X_2, \dots$ dizisinden bağımsız olduğu varsayalım. **Rastgele toplam**

$$S = \sum_{i=1}^N X_i = X_1 + X_2 + \dots + X_N$$

ile tanımlansın ($N = 0$ ise $S = 0$).

(a) $E(S)$ 'yi hesaplayınız.

(b) $\text{Var}(S)$ 'yi hesaplayınız.

(c) N 'nin $\lambda > 0$ parametrelili Poisson dağılımına sahip olduğunu varsayınız; bu dağılımı **Poisson Dağılımı** bölümünde ayrıntılı inceleyeceğiz, burada yalnızca $E(N) = \text{Var}(N) = \lambda$ olduğunu kullanınız. $E(S)$ ve $\text{Var}(S)$ 'yi λ, μ, σ^2 cinsinden yazınız.

(d) Bir sigorta şirketine bir ayda gelen hasar talebi sayısı $\lambda = 20$ parametrelili Poisson dağılımlı, her talebin tutarı ortalaması 500 ve standart sapması 300 olsun. Aylık toplam hasarın beklenen değerini ve standart sapmasını bulunuz.

Çözüm.

(a) Doğrudan hesap zordur, çünkü S 'nin dağılımı N 'nin ve X_i 'lerin dağılımlarının karmaşık bir bileşimidir. Bunun yerine N 'ye göre koşullayalım.

$N = n$ verilmişken ($n \geq 1$) $S = X_1 + \dots + X_n$ 'dir. $N, (X_1, \dots, X_n)$ vektöründen bağımsız olduğundan, $N = n$ koşulu bu vektörün dağılımını değiştirmez: $(X_1, \dots, X_n)$ 'nin $N = n$ verilmişken koşullu dağılımı, koşulsuz dağılımıyla aynıdır. Dolayısıyla $X_1 + \dots + X_n$ 'nin $N = n$ verilmişken koşullu dağılımı da koşulsuz dağılımıyla aynıdır ve **Önerme 18.2**'in tümevarımıyla n değişkene genişletilmiş hâliyle

$$E(S | N = n) = E(X_1 + \dots + X_n) = n\mu.$$

$n = 0$ için $S = 0$ ve $E(S | N = 0) = 0 = 0 \cdot \mu$; formül $n = 0$ 'ı da kapsar. Rastgele değişken biçiminde $E(S | N) = \mu N$ 'dir. Kule kuralı (Teorem 19.1):

$$E(S) = E[E(S | N)] = E(\mu N) = \mu E(N).$$

Sonuç, sezgiyle tam uyumludur: ortalama terim sayısı çarpı ortalama terim.

(b) Aynı koşullamayla $N = n$ verilmişken S 'nin koşullu varyansı, bağımsız $X_1, \dots, X_n$ 'nin toplamının varyansıdır. Bağımsız rastgele değişkenlerin kovaryansları sıfır olduğundan (Teorem 18.2), Sonuç 18.2'nin tümevarımıyla n terime genişletilmiş hâli varyansların toplandığını söyler:

$$\text{Var} (S \mid N = n) = \text{Var} (X_1 + \dots + X_n) = n\sigma^2.$$

$n = 0$ için $\text{Var} (S \mid N = 0) = 0$; formül yine geçerlidir. Öyleyse $\text{Var} (S \mid N) = \sigma^2 N$. Toplam varyans kuralı (Teorem 19.2) ile, $\text{Var} (\mu N) = \mu^2 \text{Var} (N)$ olduğunu (Teorem 16.2) kullanarak

$$\text{Var} (S) = E [\text{Var} (S \mid N)] + \text{Var} [E(S \mid N)] = E(\sigma^2 N) + \text{Var} (\mu N) = \sigma^2 E(N) + \mu^2 \text{Var} (N).$$

İki terimin anlamı: birincisi, terim sayısı bilinse bile terimlerin kendi yayılımından gelen belirsizlik; ikincisi, terim sayısının belirsizliğinden gelen belirsizlik. N sabit ($N = n$, $\text{Var} (N) = 0$) olsaydı ikinci terim düşer ve bilinen $n\sigma^2$ formülü kalırdı.

(c) $E(N) = \text{Var} (N) = \lambda$ konursa

$$E(S) = \lambda\mu, \quad \text{Var} (S) = \lambda\sigma^2 + \lambda\mu^2 = \lambda (\sigma^2 + \mu^2) = \lambda E(X_1^2)$$

olur; son eşitlik Önerme 16.3'nden gelir ($E(X_1^2) = \sigma^2 + \mu^2$). Yani Poisson sayıda terimli bir toplamın varyansı, terim sayısının ortalaması ile bir terimin ikinci momentinin çarpımıdır.

(d) $\lambda = 20$, $\mu = 500$, $\sigma = 300$:

$$E(S) = 20 \cdot 500 = 10\,000, \quad \text{Var} (S) = 20 \cdot (300^2 + 500^2) = 20 \cdot 340\,000 = 6\,800\,000.$$

Standart sapma $\sigma_S = \sqrt{6\,800\,000} \approx 2607,7$ 'dir. Karşılaştırma için: talep sayısı rastgele olmayıp tam 20 olsaydı varyans $20 \cdot 300^2 = 1\,800\,000$, standart sapma yaklaşık 1341,6 olurdu. Toplam yayılımın büyük kısmı (5 000 000) talep sayısındaki belirsizlikten gelmektedir.

■

19.6 En İyi Tahmin Olarak Koşullu Beklenen Değer

X gözlenmiş, Y gözlenmemiş olsun; Y 'yi X 'in bir fonksiyonu $g(X)$ ile tahmin etmek istiyoruz. Hangi g en iyisidir? Bu soruyu sormak için önce "iyi"nin ölçüsünü seçmeliyiz. En yaygın ölçü **ortalama karesel hata** $E [(Y - g(X))^2]$ 'dir. Aşağıdaki örnekte, bu ölçüye göre en iyi tahminin koşullu beklenen değer olduğunu göstereceğiz. Kule kuralının (ii) kısmı ispatın kalbidir.

Örnek 19.6 (Koşullu Beklenen Değer En İyi Karesel Tahmindir). (X, Y) bir rastgele vektör, $E(Y^2) < \infty$ ve $m(x) = E(Y \mid X = x)$ olsun. $E [g(X)^2] < \infty$ olan her Borel ölçülebilir g fonksiyonu için

$$E [(Y - g(X))^2] \geq E [(Y - E(Y \mid X))^2] = E [\text{Var} (Y \mid X)]$$

olduğunu ve eşitliğin ancak $P (g(X) = E(Y \mid X)) = 1$ iken sağlandığını gösteriniz.

Çözüm.

Adım 1: Hatayı iki parçaya ayırma. $Y - g(X) = (Y - m(X)) + (m(X) - g(X))$ yazıp kare alalım:

$$(Y - g(X))^2 = (Y - m(X))^2 + (m(X) - g(X))^2 + 2(m(X) - g(X))(Y - m(X)).$$

Beklenen değer alalım (Önerme 18.2'in üç terime genişletilmiş hâliyle):

$$E [(Y - g(X))^2] = E [(Y - m(X))^2] + E [(m(X) - g(X))^2] + 2 E [(m(X) - g(X))(Y - m(X))].$$

Adım 2: Çarpaz terim sıfırdır. Çarpaz terimde X 'in fonksiyonu $m(X) - g(X)$ ile $Y - m(X)$ çarpılmıştır; bu çarpımın beklenen değerinin var olması $E(Y^2) < \infty$ ve $E[g(X)^2] < \infty$ varsayımlarından çıkar. Kule kuralının (ii) kısmına (Teorem 19.1, $g(X)$ yerine $m(X) - g(X)$ ve $h(X, Y) = Y - m(X)$ ile) başvuralım:

$$E [(m(X) - g(X))(Y - m(X))] = E [(m(X) - g(X)) E (Y - m(X) \mid X)].$$

İçteki koşullu beklenen değeri hesaplayalım. Sabit bir x için, Önerme 19.1 (b) ve (c) ile

$$E(Y - m(X) | X = x) = E(Y | X = x) - E(m(X) | X = x) = m(x) - m(x) = 0.$$

Yani $E(Y - m(X) | X) = 0$ özdeş olarak sıfırdır ve çapraz terim $E[(m(X) - g(X)) \cdot 0] = 0$ olur.

Adım 3: Eşitsizlik. Adım 1 ve 2'den

$$E[(Y - g(X))^2] = E[(Y - m(X))^2] + E[(m(X) - g(X))^2].$$

Sağdaki ikinci terim, negatif olmayan bir rastgele değişkenin beklenen değeri olduğundan ≥ 0 'dır; öyleyse

$$E[(Y - g(X))^2] \geq E[(Y - m(X))^2] = E[(Y - E(Y | X))^2].$$

Adım 4: En küçük hatanın değeri. $h(x, y) = (y - m(x))^2$ ile kule kuralının (i) kısmı

$$E[(Y - m(X))^2] = E[E[(Y - m(X))^2 | X]]$$

verir. Sabit x için $E[(Y - m(x))^2 | X = x]$, tam olarak Tanım 19.3'teki $\text{Var}(Y | X = x)$ 'tir. Öyleyse en küçük ortalama karesel hata $E[\text{Var}(Y | X)]$ 'tir.

Adım 5: Eşitlik durumu. Eşitlik ancak $E[(m(X) - g(X))^2] = 0$ iken sağlanır. $Z = m(X) - g(X)$ dersek Z^2 , değer kümesi $[0, \infty)$ içinde kalan ve beklenen değeri $E(Z^2) = 0$ olan bir rastgele değişkendir. Teorem 16.3 her $\varepsilon > 0$ için $P(Z^2 \geq \varepsilon) \leq E(Z^2)/\varepsilon = 0$ verir. $\{Z \neq 0\} = \bigcup_{k \geq 1} \{Z^2 \geq \frac{1}{k}\}$ sayılabilir birleşiminin olasılığı da sayılabilir alt toplamsallık (Sonuç 3.1) ile sıfırdır: $P(Z \neq 0) = 0$, yani $P(g(X) = m(X)) = 1$. Tersine, $P(g(X) = m(X)) = 1$ ise Z^2 olasılığı 1 olan bir olay üzerinde sıfırdır; beklenen değeri $E(Z^2) = 0$ olur ve eşitlik sağlanır. ■

Sonucu iki bakış açısıyla yorumlayalım.

Sabit tahminle karşılaştırma. g sabit bir c fonksiyonu olsun; X 'e hiç bakmadan Y 'yi c ile tahmin ediyoruz. $E[(Y - c)^2] = \text{Var}(Y) + (E(Y) - c)^2$ olduğundan en iyi sabit $c = E(Y)$ 'dir ve hatası $\text{Var}(Y)$ 'dir. X 'e bakarak tahmin yapınca hata $E[\text{Var}(Y | X)]$ 'e iner ve toplam varyans kuralına göre kazanç tam olarak $\text{Var}[E(Y | X)]$ 'tir: X 'in Y hakkında taşıdığı bilgi, Y 'nin ortalamasının X 'le ne kadar değiştiğiyle ölçülür. Üçgen örneğinde sabit tahminin hatası $\text{Var}(Y) = \frac{1}{18}$, en iyi tahmin $\frac{1+X}{2}$ 'nin hatası $\frac{1}{24}$ 'tür.

Regresyonla bağ. $m(x) = E(Y | X = x)$ fonksiyonuna regresyon fonksiyonu dememizin nedeni budur: istatistikte “ Y 'nin X üzerine regresyonu” denen iş, veriden bu fonksiyonu kestirmeye çalışmaktır. Üç para atışında ve üçgen örneğinde m doğrusaldı; genel olarak doğrusal olmak zorunda değildir ve koşullu beklenen değer, doğrusal olsun olmasın bütün tahminler arasında en iyisidir.

19.7 Alıştırmalar

Alıştırma 19.1 (Koşullu Beklenen Değer Alıştırmaları). (a) (X, Y) rastgele vektörünün olasılık fonksiyonu aşağıdaki tabloyla verilsin.

$x \setminus y$	0	1
0	$\frac{1}{8}$	$\frac{2}{8}$
1	$\frac{2}{8}$	$\frac{1}{8}$
2	$\frac{1}{8}$	$\frac{1}{8}$

Her x için $E(Y | X = x)$ 'i ve her y için $E(X | Y = y)$ 'yi bulunuz; kule kuralıyla $E(Y)$ 'yi hesaplayıp marjinalden bulunan değerle karşılaştırınız.

(b) (X, Y) 'nin ortak yoğunluğu $f_{X,Y}(x, y) = e^{-y}$ ($0 < x < y < \infty$), diğer durumlarda 0 olsun. $E(Y | X = x)$ ve $E(X | Y = y)$ 'yi bulunuz; kule kuralıyla $E(Y)$ 'yi, toplam varyans kuralıyla $\text{Var}(Y)$ 'yi hesaplayınız.

(c) Düzgün bir zar iki kez atılıyor; X ilk atıştaki sayı, Y iki atıştaki sayıların toplamı olsun. $E(Y | X = x)$ ve $E(X | Y = y)$ 'yi bulunuz; toplam varyans kuralıyla $\text{Var}(Y)$ 'yi hesaplayınız. (Bir zar atışının varyansının $\frac{35}{12}$ olduğunu kullanabilirsiniz.)

(d) $N, \{1, 2, 3\}$ üzerinde düzgün dağılmış olsun; X_1, X_2, X_3 bağımsız, aynı dağılımlı, $E(X_i) = 2$, $\text{Var}(X_i) = 1$ ve N 'den bağımsız olsun. $S = X_1 + \dots + X_N$ için $E(S)$ ve $\text{Var}(S)$ 'yi bulunuz.

(e) X ile Y 'nin ikinci momentleri sonlu ve $E(Y | X) = c$ sabit olsun. $E(Y) = c$ ve $\text{Cov}(X, Y) = 0$ olduğunu gösteriniz.

(f) $X, (0, 1)$ üzerinde düzgün dağılsın ve $X = x$ verilmişken $Y, (0, x)$ üzerinde düzgün dağılsın; yani $f_{Y|X=x}(y) = \frac{1}{x}, 0 < y < x$. Kule kuralı ve toplam varyans kuralıyla $E(Y)$ ve $\text{Var}(Y)$ 'yi bulunuz; sonra Y 'nin marjinal yoğunluğunu hesaplayarak doğrulayınız.

Çözüm.

(a) Satır toplamaları $f_X(0) = \frac{3}{8}, f_X(1) = \frac{3}{8}, f_X(2) = \frac{2}{8}$; sütun toplamaları $f_Y(0) = \frac{4}{8}, f_Y(1) = \frac{4}{8}$.

Y yalnızca 0 ve 1 değerlerini aldığından $E(Y | X = x) = P(Y = 1 | X = x) = \frac{f_{X,Y}(x,1)}{f_X(x)}$:

$$E(Y | X = 0) = \frac{2/8}{3/8} = \frac{2}{3}, \quad E(Y | X = 1) = \frac{1/8}{3/8} = \frac{1}{3}, \quad E(Y | X = 2) = \frac{1/8}{2/8} = \frac{1}{2}.$$

$$E(X | Y = y) = \sum_x x \frac{f_{X,Y}(x,y)}{f_Y(y)}.$$

$$E(X | Y = 0) = \frac{0 \cdot \frac{1}{8} + 1 \cdot \frac{2}{8} + 2 \cdot \frac{1}{8}}{4/8} = \frac{4/8}{4/8} = 1, \quad E(X | Y = 1) = \frac{0 \cdot \frac{2}{8} + 1 \cdot \frac{1}{8} + 2 \cdot \frac{1}{8}}{4/8} = \frac{3/8}{4/8} = \frac{3}{4}.$$

Kule kuralı:

$$E(Y) = \frac{2}{3} \cdot \frac{3}{8} + \frac{1}{3} \cdot \frac{3}{8} + \frac{1}{2} \cdot \frac{2}{8} = \frac{2}{8} + \frac{1}{8} + \frac{1}{8} = \frac{4}{8} = \frac{1}{2}.$$

Marjinalden: $E(Y) = 0 \cdot \frac{4}{8} + 1 \cdot \frac{4}{8} = \frac{1}{2}$. Uyuşur.

(b) Marjinaller: $x > 0$ için $f_X(x) = \int_0^\infty e^{-y} dy = e^{-x}$; $y > 0$ için $f_Y(y) = \int_0^y e^{-x} dx = ye^{-y}$.

$X = x$ verilmişken: $f_{Y|X=x}(y) = \frac{e^{-y}}{e^{-x}} = e^{-(y-x)}, y > x$. Bu, x kadar kaydırılmış bir üstel yoğundur. $u = y - x$ değişken değiştirmesiyle

$$E(Y | X = x) = \int_x^\infty y e^{-(y-x)} dy = \int_0^\infty (u+x) e^{-u} du = 1+x.$$

$Y = y$ verilmişken: $f_{X|Y=y}(x) = \frac{e^{-y}}{ye^{-y}} = \frac{1}{y}, 0 < x < y$; yani $(0, y)$ üzerinde düzgün dağılım ve $E(X | Y = y) = \frac{y}{2}$.

Kule kuralı: $E(Y) = E(1 + X) = 1 + E(X)$ ve $E(X) = \int_0^\infty xe^{-x} dx = 1$; öyleyse $E(Y) = 2$. Doğrulama: $E(Y) = \int_0^\infty y^2 e^{-y} dy = 2! = 2$.

Toplam varyans: $X = x$ verilmişken $Y - x$ 'in yoğunluğu $e^{-u}, u > 0$ 'dır; ikinci momenti $\int_0^\infty u^2 e^{-u} du = 2$ ve ortalaması 1 olduğundan $\text{Var}(Y | X = x) = 2 - 1 = 1$. Buradan $E[\text{Var}(Y | X)] = 1$. Öteki terim ise şu değeri alır:

$$\text{Var}[E(Y | X)] = \text{Var}(1 + X) = \text{Var}(X) = \int_0^\infty x^2 e^{-x} dx - 1 = 2 - 1 = 1.$$

Öyleyse $\text{Var}(Y) = 1 + 1 = 2$. Doğrulama: $E(Y^2) = \int_0^\infty y^3 e^{-y} dy = 3! = 6$ ve $\text{Var}(Y) = 6 - 4 = 2$.

(c) Z ikinci atıştaki sayı olsun; $Y = X + Z$ ve X ile Z bağımsız, her biri $\{1, \dots, 6\}$ üzerinde düzgün, $E(X) = E(Z) = \frac{7}{2}, \text{Var}(X) = \text{Var}(Z) = \frac{35}{12}$.

Önerme 19.1'nin (b) ve (c) maddeleri, bağımsızlık için de (d) maddesi şu sonucu verir:

$$E(Y | X = x) = E(X + Z | X = x) = x + E(Z | X = x) = x + E(Z) = x + \frac{7}{2}.$$

$E(X | Y = y)$ için bakışım kullanalım: (X, Z) 'nin ortak dağılımı bileşenlerin yer değiştirmesine göre bakışımıdır, dolayısıyla $Y = y$ verilmişken X ile Z 'nin koşullu dağılımları aynıdır ve $E(X | Y = y) = E(Z | Y = y)$. Lineerlikten ise şu çıkar:

$$E(X | Y = y) + E(Z | Y = y) = E(X + Z | Y = y) = E(Y | Y = y) = y.$$

Öyleyse $E(X | Y = y) = \frac{y}{2}$, $y \in \{2, \dots, 12\}$. (Doğrudan sayma ile de görülebilir: $Y = y$ verilmişken X , $\max(1, y - 6)$ ile $\min(6, y - 1)$ arasındaki tam sayılar üzerinde düzgün dağılır ve bu aralığın orta noktası $\frac{y}{2}$ 'dir.)

Toplam varyans: $X = x$ verilmişken $Y = x + Z$ 'dir ve Z , X 'ten bağımsız olduğundan $\text{Var}(Y | X = x) = \text{Var}(Z) = \frac{35}{12}$; $E[\text{Var}(Y | X)] = \frac{35}{12}$. $\text{Var}[E(Y | X)] = \text{Var}\left(X + \frac{7}{2}\right) = \text{Var}(X) = \frac{35}{12}$. Öyleyse $\text{Var}(Y) = \frac{35}{12} + \frac{35}{12} = \frac{35}{6}$; bu, bağımsız iki zarın toplamının varyansı olarak beklenen sonuçtur. **(d) Örnek 19.5**'daki formüllerle $E(S) = \mu E(N)$ ve $\text{Var}(S) = \sigma^2 E(N) + \mu^2 \text{Var}(N)$. $E(N) = \frac{1+2+3}{3} = 2$, $E(N^2) = \frac{1+4+9}{3} = \frac{14}{3}$, $\text{Var}(N) = \frac{14}{3} - 4 = \frac{2}{3}$. Öyleyse

$$E(S) = 2 \cdot 2 = 4, \quad \text{Var}(S) = 1 \cdot 2 + 4 \cdot \frac{2}{3} = 2 + \frac{8}{3} = \frac{14}{3}.$$

(e) Kule kuralıyla $E(Y) = E[E(Y | X)] = E(c) = c$. İkinci momentler sonlu olduğundan $|XY| \leq \frac{1}{2}(X^2 + Y^2)$ eşitsizliği $E|XY| < \infty$ verir; öyleyse **Teorem 19.1** (ii)'yi $g(X) = X$ ile uygulayabiliriz: $E(XY) = E[X E(Y | X)] = E(cX) = c E(X)$. **Önerme 18.3** (a)'ya göre

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = c E(X) - E(X) \cdot c = 0.$$

Yani $E(Y | X)$ 'in sabit olması, kovaryansın sıfır olmasından daha güçlü bir koşuldur; ama daha önce gördüğümüz gibi bağımsızlıktan zayıftır.

(f) $X = x$ verilmişken Y , $(0, x)$ üzerinde düzgün dağıldığından $E(Y | X = x) = \frac{x}{2}$ ve $\text{Var}(Y | X = x) = \frac{x^2}{12}$. X de $(0, 1)$ üzerinde düzgün olduğundan $E(X) = \frac{1}{2}$, $E(X^2) = \frac{1}{3}$, $\text{Var}(X) = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$. Kule kuralı: $E(Y) = E\left(\frac{X}{2}\right) = \frac{1}{4}$.

Toplam varyans: $E[\text{Var}(Y | X)] = E\left(\frac{X^2}{12}\right) = \frac{1}{36}$ ve $\text{Var}[E(Y | X)] = \text{Var}\left(\frac{X}{2}\right) = \frac{1}{4} \cdot \frac{1}{12} = \frac{1}{48}$. Öyleyse

$$\text{Var}(Y) = \frac{1}{36} + \frac{1}{48} = \frac{4}{144} + \frac{3}{144} = \frac{7}{144}.$$

Doğrulama: Ortak yoğunluk $f_{X,Y}(x, y) = f_X(x) f_{Y|X=x}(y) = \frac{1}{x}$, $0 < y < x < 1$. Marjinal: $0 < y < 1$ için

$$f_Y(y) = \int_y^1 \frac{1}{x} dx = -\ln y.$$

Buradaki momentler için gereken integrali kısmi integrasyonla hesaplayalım: $k \geq 0$ için $u = -\ln y$, $dv = y^k dy$ alınırsa $du = -\frac{dy}{y}$, $v = \frac{y^{k+1}}{k+1}$ olur ve

$$\int_0^1 y^k (-\ln y) dy = \left[-\frac{y^{k+1}}{k+1} \ln y \right]_0^1 + \frac{1}{k+1} \int_0^1 y^k dy = 0 + \frac{1}{k+1} \cdot \frac{1}{k+1} = \frac{1}{(k+1)^2};$$

sınır teriminde $y = 1$ için $\ln 1 = 0$, $y \rightarrow 0^+$ için ise $y^{k+1} \ln y \rightarrow 0$ 'dır. Öyleyse $E(Y) = \int_0^1 y(-\ln y) dy = \frac{1}{4}$ ve $E(Y^2) = \int_0^1 y^2(-\ln y) dy = \frac{1}{9}$; $\text{Var}(Y) = \frac{1}{9} - \frac{1}{16} = \frac{16-9}{144} = \frac{7}{144}$. Uyuşur.

■

Bu bölümle beklenen değer kuramının temel araçları tamamlanmış oldu. Bundan sonraki kısımda, buraya kadar geliştirilen kavramları en sık karşılaşılan somut dağılımlara uygulayacağız; ilk durak, tekrarlı denemelerin temel modeli olan Bernoulli ve binom dağılımlarıdır: **Bernoulli ve Binom Dağılımları**.

20. Bernoulli ve Binom Dağılımları

Koşullu Beklenen Değer bölümüyle birlikte olasılık kuramının genel yapısı tamamlandı: olasılık uzayı, rastgele değişkenler ve rastgele vektörler, dağılım fonksiyonu, beklenen değer ve varyans, moment üreten ve karakteristik fonksiyonlar. Bu araçların hepsini “herhangi bir” rastgele değişken için geliştirdik. Şimdi bakış açımızı değiştiriyoruz. Uygulamada tekrar tekrar karşımıza çıkan birkaç **somut dağılım ailesini** tek tek ele alacak; her biri için olasılık fonksiyonunu türetecek, beklenen değerini, varyansını ve moment üreten fonksiyonunu hesaplayacağız. Bu kısımda yeni bir aksiyom ya da yeni bir kavram gelmeyecek; her sonuç, önceki kısımlarda kurulan genel kuramın belirli bir dağılıma uygulanmasından ibaret olacak. En basit rastgele deney, yalnızca iki sonucu olan deneydir: bir para atılır, yazı ya da tura gelir; üretim bandından alınan bir parça sağlam ya da kusurludur; bir tohum çimlenir ya da çimlenmez. Bu türden deneylere **Bernoulli deneyi** denir. Bu bölümde önce tek bir Bernoulli deneyini betimleyen **Bernoulli dağılımını**, sonra da aynı deneyin bağımsız olarak n kez tekrarlanmasıyla ortaya çıkan **binom dağılımını** inceleyeceğiz. Binom dağılımı, “ n denemede kaç başarı olur?” sorusunun yanıtıdır ve kesikli dağılımların en temel olanıdır.

Bölümün sonunda şunları bileceğiz: binom olasılık fonksiyonu $\binom{n}{x}p^x(1-p)^{n-x}$ nereden gelir ve toplamı neden 1’dir; beklenen değer np , varyans $np(1-p)$ olduğu birbirinden bağımsız dört ayrı yolla nasıl gösterilir; moment üreten fonksiyon $(pe^t + 1 - p)^n$ ile karakteristik ve çarpımsal moment üreten fonksiyonlar momentleri nasıl üretir; aynı p ’li bağımsız binom değişkenlerinin toplamı neden yine binomdur; ve olasılık fonksiyonunun en büyük değerini aldığı **en olası değer** nasıl bulunur. Sonraki bölümde, n büyük ve p küçük olduğunda binom dağılımına yaklaşan Poisson dağılımı gelecek.

20.1 Bernoulli Deneyi ve Bernoulli Dağılımı

Bir rastgele deneyin sonuçlarını, ilgilendiğimiz bir olayın gerçekleşip gerçekleşmediğine göre ikiye ayıralım. Deneyin geri kalan ayrıntılarını unutup yalnızca bu ikili bilgiyi kaydettiğimizde elde ettiğimiz şey bir Bernoulli deneyidir.

Tanım 20.1 (Bernoulli Deneyi). $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı ve $A \in \mathcal{U}$, $0 < P(A) < 1$ koşulunu sağlayan bir olay olsun. Sonucun yalnızca “ A gerçekleşti” ya da “ A gerçekleşmedi” biçiminde kaydedildiği rastgele deneye bir **Bernoulli deneyi** denir. A olayına **başarı**, A^c olayına **başarısızlık**, $p = P(A)$ sayısına da deneyin **başarı olasılığı** adı verilir. Başarısızlık olasılığı $q = 1 - p = P(A^c)$ ile gösterilir.

“Başarı” sözcüğü bir değer yargısı taşımaz; yalnızca saydığımız sonucun adıdır. Kusurlu parçaları sayıyorsak “kusurlu çıkması” başarıdır. $p = 0$ ya da $p = 1$ durumlarında deneyin sonucu önceden bellidir; bu yozlaşmış durumları tanımın dışında bıraktık; beklenen değer, varyans ve üretici fonksiyon formülleri o durumlarda da doğru kalır (en olası değer sonucu ise $0 < p < 1$ varsayar).

Bir Bernoulli deneyini sayıya dönüştürmenin doğal yolu, A olayının **gösterge fonksiyonunu** almaktır: $\omega \in A$ ise $X(\omega) = 1$, değilse $X(\omega) = 0$. Bu X , **Tanım 10.1** anlamında kesikli bir rastgele değişkendir; olasılık fonksiyonunu (**Tanım 10.2**) tek formülle yazabiliriz.

Tanım 20.2 (Bernoulli Dağılımı). $0 < p < 1$ olsun. Değer kümesi $D_X = \{0, 1\}$ ve olasılık fonksiyonu

$$f_X(x) = \begin{cases} p^x(1-p)^{1-x}, & x \in \{0, 1\} \\ 0, & \text{diğer durumlarda} \end{cases}$$

olan kesikli rastgele değişken X ’e **p parametrelili Bernoulli dağılımına** sahiptir denir ve $X \sim B(1, p)$ yazılır. Açık biçimde $P(X = 1) = p$ ve $P(X = 0) = 1 - p$ ’dir.

Formüldeki üsler yalnızca iki değeri tek satırda toplamaya yarar: $x = 1$ için $p^1(1-p)^0 = p$, $x = 0$ için $p^0(1-p)^1 = 1 - p$ çıkar. Dağılım fonksiyonu (**Tanım 9.1**) iki basamaklı bir merdivendir:

$$F_X(x) = P(X \leq x) = \begin{cases} 0, & x < 0 \\ 1 - p, & 0 \leq x < 1 \\ 1, & x \geq 1. \end{cases}$$

Bernoulli dağılımı o kadar basittir ki bütün sayısal özellikleri tek bakışta hesaplanır. Yine de bunları bir önerme hâlinde kaydediyoruz, çünkü binom dağılımının bütün momentleri bu önermeden türeyecek.

Önerme 20.1 (Bernoulli Dağılımının Momentleri). $X \sim B(1, p)$ olsun. O zaman:

(a) Her $k \in \mathbb{N}$ için $E(X^k) = p$; özel olarak $E(X) = p$.

(b) $\text{Var}(X) = p(1 - p)$.

(c) Moment üreten fonksiyonu her $t \in \mathbb{R}$ için vardır ve $M_X(t) = 1 - p + pe^t$ 'dir.

(d) Karakteristik fonksiyonu $\varphi_X(t) = 1 - p + pe^{it}$, çarpımsal moment üreten fonksiyonu $N_X(t) = 1 - p + pt$ 'dir; ikisi de her $t \in \mathbb{R}$ için tanımlıdır.

İspat.

(a) X kesikli olduğundan Tanım 16.1 gereği $E(X^k) = \sum_{x \in D_X} x^k f_X(x)$ 'tir. $D_X = \{0, 1\}$ olduğundan

$$E(X^k) = 0^k \cdot (1 - p) + 1^k \cdot p = p.$$

(b) Önerme 16.3 gereği $\text{Var}(X) = E(X^2) - (E(X))^2$ 'dir. (a)'ya göre $E(X^2) = p$ ve $E(X) = p$; öyleyse

$$\text{Var}(X) = p - p^2 = p(1 - p).$$

(c) Tanım 17.2 gereği $M_X(t) = E(e^{tX})$ 'tir. Toplam yine iki terimlidir:

$$M_X(t) = e^{t \cdot 0}(1 - p) + e^{t \cdot 1}p = 1 - p + pe^t.$$

Toplam sonlu olduğundan her $t \in \mathbb{R}$ için tanımlıdır. Sağlama olarak Teorem 17.4'i uygulayalım: $M'_X(t) = pe^t$ ve $M''_X(t) = pe^t$ olduğundan $M'_X(0) = p = E(X)$, $M''_X(0) = p = E(X^2)$ bulunur; (a) ile uyuşur.

(d) Tanım 17.1 gereği $\varphi_X(t) = E(e^{itX})$ ve Tanım 17.3 gereği $N_X(t) = E(t^X)$ 'tir. Her ikisi de (c)'deki gibi iki terimli toplamlardır:

$$\varphi_X(t) = e^{it \cdot 0}(1 - p) + e^{it \cdot 1}p = 1 - p + pe^{it}, \quad N_X(t) = t^0(1 - p) + t^1p = 1 - p + pt.$$

$N'_X(t) = p$ olduğundan Teorem 17.6 ile $E(X) = N'_X(1) = p$; ikinci ve daha yüksek türevler sıfırdır, yani $E(X(X - 1)) = 0$. Bu, X 'in yalnızca 0 ve 1 değerlerini almasının doğrudan sonucudur: $X(X - 1) = 0$.

■

Varyans $p(1 - p)$ en büyük değerini $p = \frac{1}{2}$ 'de alır ve hiçbir zaman $\frac{1}{4}$ 'ü aşamaz (Alıştırma 20.1, (c) şıkkı): sonucu en az öngörülebilir olan Bernoulli deneyi, hilesiz para atışıdır.

20.2 Binom Dağılımı

Aynı Bernoulli deneyini, koşulları değiştirmeden ve her tekrarın sonucu ötekileri etkilemeyecek biçimde n kez yineleyelim. İlgilendiğimiz büyüklük, bu n denemedeki toplam başarı sayısıdır. Önce bu sayının dağılımını bir aile olarak tanımlayacak, sonra bağımsız tekrarlardaki başarı sayısının gerçekten bu dağılıma sahip olduğunu ispatlayacağız. Kısalık için baştan sona $q = 1 - p$ yazacağız.

Tanım 20.3 (Binom Dağılımı). $n \in \mathbb{N}$ ve $0 < p < 1$ olsun. Değer kümesi $D_X = \{0, 1, \dots, n\}$ ve olasılık fonksiyonu

$$f_X(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & x \in \{0, 1, \dots, n\} \\ 0, & \text{diğer durumlarda} \end{cases}$$

olan kesikli rastgele değişken X 'e **n ve p parametrelili binom dağılımına** sahiptir denir ve $X \sim B(n, p)$ yazılır. Burada $\binom{n}{x} = \frac{n!}{x!(n-x)!}$, Teorem 5.2 ile verilen kombinasyon sayısıdır. $n = 1$ için $B(1, p)$ tam olarak Bernoulli dağılımıdır.

Bu tanımın anlamlı olması için verilen f_X 'in gerçekten bir olasılık fonksiyonu olması gerekir. Değerlerin negatif olmadığı açıktır; toplamın 1 olduğu ise binom teoreminin (Teorem 5.4) doğrudan bir sonucudur:

$$\sum_{x=0}^n \binom{n}{x} p^x (1-p)^{n-x} = (p + (1-p))^n = 1^n = 1.$$

Dağılımın adı da buradan gelir: $f_X(0), f_X(1), \dots, f_X(n)$ olasılıkları, $(q+p)^n$ açılımının $\binom{n}{x} p^x q^{n-x}$ terimleridir; x sırasıyla $0, 1, \dots, n$ değerlerini alır.

Şimdi bu tanımın olasılık uzayına bağlanması gerekiyor. Aşağıdaki teorem, binom dağılımının “neden” bu biçimde olduğunu söyler: katsayı $\binom{n}{x}$ başarıların yerleşebileceği konum sayısını, $p^x q^{n-x}$ çarpanı ise belirli bir konum düzeninin olasılığını verir.

Teorem 20.1 (Bağımsız Denemelerde Başarı Sayısı). Başarı olasılığı p ($0 < p < 1$) olan bir Bernoulli deneyi, birbirinden bağımsız olarak ve aynı koşullar altında n kez tekrarlınsın. X , bu n denemedeki başarı sayısı olsun. O zaman $X \sim B(n, p)$ 'dir; yani

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \dots, n.$$

İspat.

Model. n denemenin sonucunu, i 'inci deneme başarılıysa $\omega_i = 1$, değilse $\omega_i = 0$ olmak üzere bir $\omega = (\omega_1, \dots, \omega_n)$ dizisiyle kaydedelim. Örnek uzay $\Omega = \{0, 1\}^n$, 2^n elemanlı sonlu bir kümedir; σ -cebiri olarak $\mathcal{U} = \mathcal{P}(\Omega)$ alınır. İçinde tam k tane 1 bulunan bir ω dizisine

$$P(\{\omega\}) = p^k q^{n-k}$$

olasılığını verelim; bir olayın olasılığı, elemanlarının olasılıklarının toplamıdır. Bu bir olasılık ölçüsüdür: k tane 1 içeren dizi sayısı $\binom{n}{k}$ olduğundan (Teorem 5.2) bütün tekil olasılıkların toplamı

$$\sum_{\omega \in \Omega} P(\{\omega\}) = \sum_{k=0}^n \binom{n}{k} p^k q^{n-k} = (p+q)^n = 1$$

olur. “Denemeler bağımsızdır ve her birinin başarı olasılığı p 'dir” cümlesinin matematiksel karşılığı da tam olarak bu ölçüdür. Gerçekten, $A_i = \{\omega \in \Omega : \omega_i = 1\}$ “ i 'inci deneme başarılı” olayı olsun. $i_1 < \dots < i_r$ indisleri için $A_{i_1} \cap \dots \cap A_{i_r}$ olayının olasılığını hesaplarken sabitlenen r koordinat p^r çarpanını verir; geriye kalan $n - r$ koordinat üzerinden toplam alındığında $(p+q)^{n-r} = 1$ çıkar. Öyleyse

$$P(A_{i_1} \cap \dots \cap A_{i_r}) = p^r = P(A_{i_1}) \dots P(A_{i_r}),$$

yani $A_1, \dots, A_n$ olayları Tanım 7.2 anlamında tam bağımsızdır ve her biri p olasılıklıdır.

Sayma. Başarı sayısı $X(\omega) = \omega_1 + \dots + \omega_n$ 'dir. $(X = x)$ olayı, içinde tam x tane 1 bulunan dizilerden oluşur. Böyle bir dizi, 1'lerin yerleşeceği x konumun n konum arasından seçilmesiyle belirlenir; Teorem 5.2 gereği $\binom{n}{x}$ tane vardır. Her birinin olasılığı $p^x q^{n-x}$ 'tir. Dolayısıyla

$$P(X = x) = \binom{n}{x} p^x q^{n-x}, \quad x = 0, 1, \dots, n.$$

X 'in başka değer alamayacağı açıktır; öyleyse $X \sim B(n, p)$ 'dir. ■

$p = \frac{1}{2}$ özel durumunda her dizinin olasılığı 2^{-n} 'dir; yani Ω bir Laplace uzayıdır (Tanım 4.1) ve formül, “uygun durum sayısı bölü bütün durum sayısı” biçimine indirgenir: $P(X = x) = \binom{n}{x} / 2^n$. Genel p için diziler eşit olasılıklı değildir; sayma yine yapılır ama her dizi kendi ağırlığı $p^x q^{n-x}$ ile sayılır.

İspatta ortaya çıkan $X(\omega) = \omega_1 + \dots + \omega_n$ yazılışı, tek bir formülden çok daha fazlasını söyler: binom değişkeni, her biri bir denemeyi izleyen Bernoulli değişkenlerinin toplamıdır. Bu gözlem, ileride beklenen değeri ve varyansı neredeyse hesapsız verecek.

Önerme 20.2 (Binom Değişkeni Bernoulli Değişkenlerinin Toplamıdır). $X_1, \dots, X_n$ bağımsız rastgele değişkenler ve her biri $B(1, p)$ dağılımlı olsun. O zaman

$$X = X_1 + X_2 + \dots + X_n \sim B(n, p)$$

olur. Tersine, $B(n, p)$ dağılımlı her rastgele değişken, dağılım bakımından böyle bir toplama eşittir.

İspat.

$(X_1, \dots, X_n)$ vektörü $\{0, 1\}^n$ kümesinde değer alır. Bağımsızlık (Tanım 14.1) gereği, içinde k tane 1 bulunan her $(\omega_1, \dots, \omega_n) \in \{0, 1\}^n$ için

$$P(X_1 = \omega_1, \dots, X_n = \omega_n) = \prod_{i=1}^n P(X_i = \omega_i) = p^k q^{n-k}$$

olur; çünkü çarpanların k tanesi $P(X_i = 1) = p$, geri kalan $n - k$ tanesi $P(X_i = 0) = q$ 'dur. Bu, Teorem 20.1 ispatındaki ölçünün ta kendisidir. $X = X_1 + \dots + X_n$, dizideki 1'lerin sayısını verir; öyleyse aynı sayma

$$P(X = x) = \sum_{\substack{\omega \in \{0,1\}^n \\ \omega_1 + \dots + \omega_n = x}} p^x q^{n-x} = \binom{n}{x} p^x q^{n-x}$$

eşitliğini verir, yani $X \sim B(n, p)$ 'dir.

Tersine $X \sim B(n, p)$ verilsin. Teorem 20.1 ispatındaki $(\Omega, \mathcal{P}(\Omega), P)$ uzayında $X_i(\omega) = \omega_i$ tanımlansın. Orada gösterildiği gibi $P(X_i = 1) = P(A_i) = p$, dolayısıyla $P(X_i = 0) = q$ 'dur; yani her $X_i \sim B(1, p)$ 'dir. İçinde k tane 1 bulunan her $\omega = (\omega_1, \dots, \omega_n) \in \{0, 1\}^n$ noktasında ortak olasılık fonksiyonu

$$P(X_1 = \omega_1, \dots, X_n = \omega_n) = P(\{\omega\}) = p^k q^{n-k} = \prod_{i=1}^n P(X_i = \omega_i)$$

marjinallerin çarpımıdır; Tanım 14.1 gereği $X_1, \dots, X_n$ bağımsızdır. Toplamları $X_1(\omega) + \dots + X_n(\omega) = \omega_1 + \dots + \omega_n$, ω uzaydaki başarı sayısıdır; dağılımı $B(n, p)$, yani X 'in dağılımıdır. ■

Şimdi somut bir hesap yapalım. Aşağıdaki örnek, olasılık fonksiyonunun tablosunu, olasılıkların toplamı hesabını ve dağılım fonksiyonunun birikimli toplamı oluşturulmasını bir arada gösteriyor.

Örnek 20.1 (Zar Beş Kez Atılıyor). Hilesiz bir zar beş kez atılıyor; X , 6 gelen atışların sayısı olsun.

(a) X 'in olasılık fonksiyonunu tablo hâlinde yazınız.

(b) Tam iki kez 6 gelme, en az bir kez 6 gelme ve en az iki kez 6 gelme olasılıklarını bulunuz.

(c) Dağılım fonksiyonu F_X 'i hesaplayıp açık biçimde yazınız; $P(1 \leq X \leq 3)$ olasılığını F_X yardımıyla bulunuz.

Çözüm.

Her atış bir Bernoulli deneyidir: başarı "6 gelmesi", başarı olasılığı $p = \frac{1}{6}$, dolayısıyla $q = \frac{5}{6}$. Atışlar bağımsızdır; Teorem 20.1 gereği $X \sim (B, \frac{1}{6})$ 'dir.

(a) $6^5 = 7776$ olduğundan

$$f_X(x) = \binom{5}{x} \left(\frac{1}{6}\right)^x \left(\frac{5}{6}\right)^{5-x} = \frac{\binom{5}{x} 5^{5-x}}{7776}, \quad x = 0, 1, \dots, 5.$$

$\binom{5}{x}$ katsayıları 1, 5, 10, 10, 5, 1 ve 5^{5-x} kuvvetleri 3125, 625, 125, 25, 5, 1'dir. Çarpımlar tabloyu verir:

x	0	1	2	3	4	5
$f_X(x)$	$\frac{3125}{7776}$	$\frac{3125}{7776}$	$\frac{1250}{7776}$	$\frac{250}{7776}$	$\frac{25}{7776}$	$\frac{1}{7776}$
yaklaşık	0,4019	0,4019	0,1608	0,0322	0,0032	0,0001

Sağlama: $3125 + 3125 + 1250 + 250 + 25 + 1 = 7776$, yani olasılıkların toplamı 1'dir. Tablodaki ilk iki değer eşit olması rastlantı değildir; nedenini Örnek 20.3 açıklayacak.

(b) Tam iki kez 6 gelme olasılığı doğrudan tablodan okunur:

$$P(X = 2) = \frac{1250}{7776} = \frac{625}{3888} \approx 0,1608.$$

“En az bir kez” olayının tümleyeni “hiç gelmemesi”dir:

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{3125}{7776} = \frac{4651}{7776} \approx 0,5981.$$

Benzer biçimde

$$P(X \geq 2) = 1 - P(X = 0) - P(X = 1) = 1 - \frac{6250}{7776} = \frac{1526}{7776} = \frac{763}{3888} \approx 0,1962.$$

Beş atışta en az bir 6 görme olasılığının $\frac{1}{2}$ 'den ancak biraz büyük olması çoğu kişiye şaşırtıcı gelir. Yaygın yanlış, beklenen sayıyı olasılıkla karıştırmaktır: beş atışta ortalama $\frac{5}{6} \approx 0,83$ tane 6 gelmesi, “en az bir 6 gelme” olasılığının 0,83 olduğu anlamına gelmez; o olasılık 0,5981’dir.

(c) Kesikli bir değişkenin dağılım fonksiyonu $F_X(x) = P(X \leq x) = \sum_{k \leq x} f_X(k)$ birikimli toplamıdır. Tablodaki payları soldan sağa toplayalım:

x	0	1	2	3	4	5
$F_X(x)$	$\frac{3125}{7776}$	$\frac{6250}{7776}$	$\frac{7500}{7776}$	$\frac{7750}{7776}$	$\frac{7775}{7776}$	1

F_X , tam sayı noktalarında sıçrayan ve aralarda sabit kalan bir basamak fonksiyonudur:

$$F_X(x) = \begin{cases} 0, & x < 0 \\ \frac{3125}{7776}, & 0 \leq x < 1 \\ \frac{6250}{7776}, & 1 \leq x < 2 \\ \frac{7500}{7776}, & 2 \leq x < 3 \\ \frac{7750}{7776}, & 3 \leq x < 4 \\ \frac{7775}{7776}, & 4 \leq x < 5 \\ 1, & x \geq 5. \end{cases}$$

X yalnızca tam sayı değerler aldığından $(1 \leq X \leq 3) = (0 < X \leq 3)$ ’tür ve Teorem 9.2 (a)’daki $P(a < X \leq b) = F_X(b) - F_X(a)$ formülü uygulanır:

$$P(1 \leq X \leq 3) = F_X(3) - F_X(0) = \frac{7750 - 3125}{7776} = \frac{4625}{7776} \approx 0,5948.$$

$$\text{Sağlama: } f_X(1) + f_X(2) + f_X(3) = \frac{3125+1250+250}{7776} = \frac{4625}{7776}.$$

■

20.3 Beklenen Değer ve Varyans

Binom dağılımının beklenen değerini ve varyansını iki bağımsız yolla hesaplayacağız. Birinci yol, tanım-daki toplamı doğrudan hesaplar ve binom katsayılarının bir özdeşliğine dayanır. İkinci yol, Önerme 20.2’deki toplam yazılışını kullanır ve neredeyse hesap gerektirmez. İki yolun aynı sonucu vermesi, hem formüllerin hem de araçların sağlamasıdır.

Teorem 20.2 (Binom Dağılımının Beklenen Değeri ve Varyansı). $X \sim B(n, p)$ ise

$$E(X) = np, \quad \text{Var}(X) = np(1-p) = npq, \quad \sigma_X = \sqrt{npq}$$

olur.

İspat.

Birinci yol: doğrudan toplam. Önce binom katsayıları için iki özdeşlik: $1 \leq x \leq n$ için

$$x \binom{n}{x} = x \cdot \frac{n!}{x!(n-x)!} = \frac{n \cdot (n-1)!}{(x-1)!(n-x)!} = n \binom{n-1}{x-1},$$

ve $2 \leq x \leq n$ için, aynı adımı bir kez daha uygulayarak,

$$x(x-1)\binom{n}{x} = n(n-1)\binom{n-2}{x-2}.$$

Beklenen değer [Tanım 16.1](#) gereği $\sum_x x f_X(x)$ 'tir; $x = 0$ terimi sıfırdır:

$$E(X) = \sum_{x=1}^n x \binom{n}{x} p^x q^{n-x} = \sum_{x=1}^n n \binom{n-1}{x-1} p^x q^{n-x} = np \sum_{x=1}^n \binom{n-1}{x-1} p^{x-1} q^{n-x}.$$

Son toplamda $y = x - 1$ koyalım; y , 0'dan $n - 1$ 'e gider ve $n - x = (n - 1) - y$ 'dir:

$$E(X) = np \sum_{y=0}^{n-1} \binom{n-1}{y} p^y q^{(n-1)-y} = np (p + q)^{n-1} = np.$$

Toplam, [Teorem 5.4](#) gereği $(p + q)^{n-1} = 1$ 'e eşittir; başka bir deyişle, toplanan sayılar $B(n - 1, p)$ dağılımının olasılıklarıdır.

Varyans için $E(X^2)$ 'yi doğrudan hesaplamak yerine önce $E(X(X - 1))$ 'i buluruz; ikinci özdeşlik tam bunun için hazırlandı. $x = 0$ ve $x = 1$ terimleri sıfırdır:

$$E(X(X - 1)) = \sum_{x=2}^n x(x-1) \binom{n}{x} p^x q^{n-x} = n(n-1)p^2 \sum_{x=2}^n \binom{n-2}{x-2} p^{x-2} q^{(n-2)-(x-2)} = n(n-1)p^2,$$

çünkü $y = x - 2$ dönüşümüyle son toplam $(p + q)^{n-2} = 1$ olur. $X^2 = X(X - 1) + X$ olduğundan [Teorem 16.1](#) ile $E(X^2) = E(X(X - 1)) + E(X) = n(n-1)p^2 + np$ ve [Önerme 16.3](#) 'nden

$$\text{Var}(X) = E(X^2) - (E(X))^2 = n(n-1)p^2 + np - n^2p^2 = np - np^2 = np(1 - p).$$

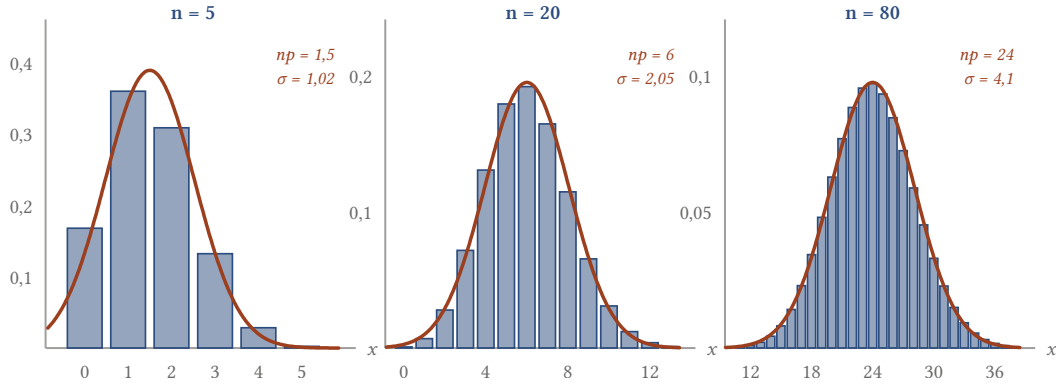
İkinci yol: Bernoulli toplamı. [Önerme 20.2](#) gereği X 'in dağılımı, bağımsız $B(1, p)$ değişkenlerinin toplamı $X_1 + \dots + X_n$ 'nin dağılımıyla aynıdır; beklenen değer ve varyans yalnızca dağılıma bağlı olduğundan hesabı bu toplam üzerinde yapabiliriz. Rastgele değişkenlerin toplamının beklenen değeri, beklenen değerlerin toplamıdır ([Önerme 18.2](#); n terim için tümevarımla, bunun için bağımsızlık gerekmez); [Önerme 20.1](#) ile

$$E(X) = E(X_1) + \dots + E(X_n) = \underbrace{p + \dots + p}_{n \text{ terim}} = np.$$

X_i 'ler bağımsız olduğundan kovaryansları sıfırdır ([Teorem 18.2](#)) ve [Teorem 18.4](#) gereği (n terim için tümevarımla) toplamın varyansı varyansların toplamıdır:

$$\text{Var}(X) = \text{Var}(X_1) + \dots + \text{Var}(X_n) = np(1 - p).$$

Standart sapma, varyansın kareköküdür. Bu iki sayı dağılımın biçimini de belirler: n büyüdükçe olasılık fonksiyonunun çubukları, ortalaması np ve varyansı $np(1 - p)$ olan normal yoğunluk eğrisine oturur.



Şekil 20.1: $p = 0,3$ için binom çubukları ve üstlerine çizilen $N(np, np(1 - p))$ normal yoğunluğu. $n = 5$ 'te çubuklar sağa çarpıktır ve çan eğrisine uymaz; n büyüdükçe çarpıklık silinir ve $n = 80$ 'de çubuk tepeleri neredeyse eğrinin üstündedir. Bu, binom dağılımının normal yaklaşımıdır; nedenini merkezi limit teoremiyle göreceğiz.

İkinci yolun varyans hesabında bağımsızlık gerçekten gereklidir: yalnızca toplamsallığa dayanan beklenen değer aksine, bağımlı Bernoulli değişkenlerinin toplamının varyansı, toplamın varyans formülünde beliren $2 \text{Cov}(X_i, X_j)$ terimleri yüzünden npq 'dan farklı olabilir; böyle bir toplamın dağılımı da genellikle binom olmaz.

Formülleri Örnek 20.1'na uygulayalım: $X \sim (B, \frac{1}{6})$ için

$$E(X) = 5 \cdot \frac{1}{6} = \frac{5}{6}, \quad \text{Var}(X) = 5 \cdot \frac{1}{6} \cdot \frac{5}{6} = \frac{25}{36}, \quad \sigma_X = \frac{5}{6}.$$

Beş atışta ortalama $\frac{5}{6} \approx 0,83$ kez 6 gelir; standart sapma da tam olarak aynı sayıdır: $\sigma_X = \sqrt{25/36} = \frac{5}{6} = E(X)$. Tablodaki olasılıkların 0 ve 1 etrafında yığılıp 3'ten sonra hızla sönmesi bununla uyumludur.

20.4 Moment Üreten Fonksiyon ve Toplamlar

Moment üreten fonksiyon, bir dağılımın bütün momentlerini tek bir fonksiyonda toplar ve bağımsız toplamların dağılımını bulmanın en kısa yoludur. Binom dağılımı için bu fonksiyon, binom teoremi sayesinde kapalı biçimde hesaplanır. Aynı hesap, küçük değişikliklerle, karakteristik fonksiyonu (Tanım 17.1) ve çarpımsal moment üreten fonksiyonu (Tanım 17.3) da verir; ikincisi, çarpımsal momentleri hesapsız üretir.

Teorem 20.3 (Binom Dağılımının Üretici Fonksiyonları). $X \sim B(n, p)$ olsun. O zaman:

- (a) Her $t \in \mathbb{R}$ için $M_X(t) = (pe^t + 1 - p)^n = (q + pe^t)^n$.
- (b) $M'_X(0) = np$ ve $M''_X(0) = n(n-1)p^2 + np$; dolayısıyla $E(X) = np$, $E(X^2) = n(n-1)p^2 + np$ ve $\text{Var}(X) = np(1-p)$.
- (c) Karakteristik fonksiyonu ve çarpımsal moment üreten fonksiyonu her $t \in \mathbb{R}$ için tanımlıdır ve

$$\varphi_X(t) = (q + pe^{it})^n, \quad N_X(t) = (q + pt)^n.$$

- (d) $k = 1, 2, \dots, n$ için k . çarpımsal moment

$$E[X(X-1)\cdots(X-k+1)] = n(n-1)\cdots(n-k+1)p^k$$

olur; $k > n$ için bu moment sıfırdır.

İspat.

- (a) X kesikli olduğundan Tanım 17.2 gereği

$$M_X(t) = E(e^{tX}) = \sum_{x=0}^n e^{tx} \binom{n}{x} p^x q^{n-x} = \sum_{x=0}^n \binom{n}{x} (pe^t)^x q^{n-x}.$$

Bu, $a = pe^t$ ve $b = q$ için $(a + b)^n$ 'nin binom açılımıdır (Teorem 5.4); öyleyse $M_X(t) = (pe^t + q)^n$ 'dir. Toplam sonlu olduğundan her $t \in \mathbb{R}$ için tanımlıdır.

Aynı sonuca Önerme 20.2 üzerinden de ulaşılır: $X_1, \dots, X_n$ bağımsız $B(1, p)$ değişkenleri ise (Tanım 14.1) $e^{tX_1}, \dots, e^{tX_n}$ de bağımsızdır ve bağımsız değişkenlerin çarpımının beklenen değeri beklenen değerlerin çarpımıdır (Teorem 18.1; n çarpan için, $e^{t(X_1 + \dots + X_{n-1})}$ ile e^{tX_n} bağımsız olduğundan tümevarımla). Öyleyse

$$M_{X_1 + \dots + X_n}(t) = E(e^{tX_1} \dots e^{tX_n}) = \prod_{i=1}^n E(e^{tX_i}) = \prod_{i=1}^n (q + pe^t) = (q + pe^t)^n;$$

burada her çarpan Önerme 20.1 (c)'den geldi.

(b) Zincir kuralıyla türev alalım. $q + pe^t$ 'nin türevi pe^t olduğundan

$$M'_X(t) = n(q + pe^t)^{n-1} pe^t,$$

$$M''_X(t) = n(n-1)(q + pe^t)^{n-2} (pe^t)^2 + n(q + pe^t)^{n-1} pe^t.$$

$t = 0$ 'da $q + pe^0 = q + p = 1$ olduğundan

$$M'_X(0) = n \cdot 1 \cdot p = np, \quad M''_X(0) = n(n-1)p^2 + np.$$

Teorem 17.4 gereği $E(X) = M'_X(0) = np$ ve $E(X^2) = M''_X(0) = n(n-1)p^2 + np$ 'dir. Önerme 16.3 ile

$$\text{Var}(X) = n(n-1)p^2 + np - (np)^2 = np - np^2 = np(1-p).$$

Bu, Teorem 20.2'nin üçüncü bir ispatıdır.

(c) M_X bütün $\mathbb{R}$ üzerinde var olduğundan Önerme 17.1 gereği $\varphi_X(t) = M_X(it) = (q + pe^{it})^n$ 'dir. Doğrudan hesap da aynı şeyi verir: (a)'daki toplamda e^t yerine e^{it} yazılır; binom açılımı karmaşık sayılar için de geçerlidir. Çarpımsal moment üreten fonksiyon için Tanım 17.3 gereği

$$N_X(t) = E(t^X) = \sum_{x=0}^n t^x \binom{n}{x} p^x q^{n-x} = \sum_{x=0}^n \binom{n}{x} (pt)^x q^{n-x} = (q + pt)^n;$$

bu, t 'nin n . dereceden bir polinomudur ve her gerçel t için tanımlıdır. Beklendiği gibi $M_X(t) = N_X(e^t)$ ve $\varphi_X(t) = N_X(e^{it})$ 'dir.

(d) N_X sonlu bir toplam olduğundan terim terim türevlenir. $\frac{d^k}{dt^k} t^x = x(x-1)\dots(x-k+1)t^{x-k}$ olduğundan

$$N_X^{(k)}(1) = \sum_{x=0}^n x(x-1)\dots(x-k+1) \binom{n}{x} p^x q^{n-x} = E[X(X-1)\dots(X-k+1)];$$

son eşitlik, Tanım 16.1'in $g(x) = x(x-1)\dots(x-k+1)$ fonksiyonuna uygulanmasıdır ($x < k$ için çarpım sıfırdır, dolayısıyla o terimler düşer). Öte yandan $(q + pt)^n$ kapalı biçimini zincir kuralıyla k kez türevlersek $k \leq n$ için

$$N_X^{(k)}(t) = n(n-1)\dots(n-k+1) p^k (q + pt)^{n-k},$$

$k > n$ için ise $N_X^{(k)}(t) = 0$ bulunur. $t = 1$ 'de $q + p = 1$ olduğundan $N_X^{(k)}(1) = n(n-1)\dots(n-k+1) p^k$ 'dir. İki hesabı karşılaştırmak (d)'yi verir.

$k = 1$ ve $k = 2$ için $E(X) = np$ ve $E[X(X-1)] = n(n-1)p^2$ çıkar; bunlar Teorem 20.2'nin birinci yolundaki ara sonuçlardır ve oradaki gibi $\text{Var}(X) = n(n-1)p^2 + np - n^2p^2 = npq$ verir. Böylece varyans için dördüncü bir yol elde edilmiş olur.

■

Moment üreten fonksiyonun asıl gücü, [Teorem 17.5](#) ile birleştğinde ortaya çıkar: moment üreten fonksiyonu $(q + pe^t)^n$ olan **her** rastgele değişken $B(n, p)$ dağılımlıdır. Bu, bir toplamın binom olup olmadığını anlamının en hızlı yoludur.

Teorem 20.4 (Bağımsız Binom Değişkenlerinin Toplamı). $X \sim B(n, p)$ ve $Y \sim B(m, p)$ bağımsız rastgele değişkenler olsun (**aynı** p ile). O zaman

$$X + Y \sim B(n + m, p)$$

olur. Daha genel olarak, $X_1, \dots, X_k$ bağımsız ve $X_i \sim B(n_i, p)$ ise $X_1 + \dots + X_k \sim B(n_1 + \dots + n_k, p)$ 'dir.

İspat.

X ve Y bağımsız olduğundan ([Tanım 14.1](#)) e^{tX} ve e^{tY} de bağımsızdır; [Teorem 18.1](#) gereği bağımsız değişkenlerin çarpımının beklenen değeri, beklenen değerlerin çarpımıdır. Öyleyse her $t \in \mathbb{R}$ için

$$M_{X+Y}(t) = E(e^{tX} e^{tY}) = E(e^{tX}) E(e^{tY}) = M_X(t) M_Y(t).$$

[Teorem 20.3](#) ile

$$M_{X+Y}(t) = (q + pe^t)^n (q + pe^t)^m = (q + pe^t)^{n+m}, \quad t \in \mathbb{R}.$$

Sağ taraf, $B(n + m, p)$ dağılımının moment üreten fonksiyonudur ve her t için tanımlıdır. Moment üreten fonksiyon dağılımı tek biçimde belirlediğinden ([Teorem 17.5](#)) $X + Y \sim B(n + m, p)$ 'dir. Genel ifade, bu sonucun $k - 1$ kez ardışık uygulanmasıyla çıkar: $X_1 + X_2 \sim B(n_1 + n_2, p)$; bu toplam X_3 'ten bağımsız olduğundan $X_1 + X_2 + X_3 \sim B(n_1 + n_2 + n_3, p)$; ve böyle devam eder.

■

Teoremin sezgisel anlamı açıktır: n denemede başarıları m denemede başarıları eklemek, $n + m$ denemede başarıları saymaktır; yeter ki bütün denemeler bağımsız ve aynı p 'li olsun. [Önerme 20.2](#) da bu teoremin $n_1 = \dots = n_k = 1$ özel hâlidir.

İ Konvolüsyon ve Vandermonde özdeşliği

Konvolüsyon formülü ([Önerme 15.3](#)), teoremi bir sayma özdeşliğine bağlar. $X \sim B(n, p)$ ve $Y \sim B(m, p)$ bağımsız ise $k = 0, 1, \dots, n + m$ için

$$P(X + Y = k) = \sum_j P(X = j) P(Y = k - j) = \sum_j \binom{n}{j} \binom{m}{k-j} p^j q^{n-j} p^{k-j} q^{m-k+j} = p^k q^{n+m-k} \sum_j \binom{n}{j} \binom{m}{k-j},$$

burada toplam $\max(0, k - m) \leq j \leq \min(n, k)$ üzerinden alınır. Öte yandan [Teorem 20.4](#) gereği bu olasılık $\binom{n+m}{k} p^k q^{n+m-k}$ 'ya eşittir. İkisini karşılaştırınca

$$\sum_j \binom{n}{j} \binom{m}{k-j} = \binom{n+m}{k}$$

Vandermonde özdeşliği çıkar: olasılık kuramı, bir sayma özdeşliğine ispat vermiş olur. Özdeşliğin doğrudan yorumu da aynıdır: $n + m$ nesneden k tanesini seçmek, ilk n 'den j tane ve son m 'den $k - j$ tane seçmenin bütün j değerleri üzerinden toplamıdır.

⚠ Başarı olasılıkları farklıysa toplam binom değildir

Teorem 20.4'nde “aynı p ” koşulu vazgeçilmezdir. $X \sim (B, \frac{1}{2})$ ve $Y \sim (B, \frac{1}{3})$ bağımsız olsun. O zaman

$$M_{X+Y}(t) = \left(\frac{1}{2} + \frac{1}{2}e^t\right)\left(\frac{2}{3} + \frac{1}{3}e^t\right) = \frac{1}{3} + \frac{1}{2}e^t + \frac{1}{6}e^{2t}.$$

$X + Y$ değerlerini $\{0, 1, 2\}$ 'de aldığından binom olsaydı ancak $B(2, p)$ olabilirdi; o dağılımın moment üreten fonksiyonu $(q + pe^t)^2 = q^2 + 2pqe^t + p^2e^{2t}$ 'dir. Katsayıları eşitlersek $p^2 = \frac{1}{6}$ ve $q^2 = \frac{1}{3}$ gerekir; ama o zaman $p + q = \frac{1}{\sqrt{6}} + \frac{1}{\sqrt{3}} \approx 0,986 \neq 1$ olur. Öyleyse $X + Y$ binom dağılımlı değildir. Farklı p 'li bağımsız Bernoulli değişkenlerinin toplamına **Poisson binom dağılımı** denir; onu bu kitapta ele almayacağız.

20.5 Uygulama: Kalite Kontrol

Binom dağılımının en yaygın uygulaması, bir yığından alınan örnekteki kusurlu birim sayısıdır. Aşağıdaki örnek hem olasılık hesaplarını hem de beklenen değer, varyans ve Chebyshev eşitsizliğinin somut bir dağılımda nasıl çalıştığını gösteriyor.

Örnek 20.2 (Kalite Kontrol). Bir üretim bandından çıkan parçaların %5'i kusurludur ve parçaların kusurlu olup olmaması birbirinden bağımsızdır. Banttı rastgele 20 parça alınıp inceleniyor; X , bunlar arasındaki kusurlu parça sayısı olsun.

- Örnekte en çok bir kusurlu parça bulunma ve en az iki kusurlu parça bulunma olasılıklarını bulunuz.
- $E(X)$, $\text{Var}(X)$ ve σ_X 'i hesaplayınız.
- $P(X \geq 4)$ olasılığını tam olarak hesaplayınız ve Chebyshev eşitsizliğinin verdiği üst sınırla karşılaştırınız.
- En az bir kusurlu parça bulma olasılığının 0,9'u aşması için kaç parça incelenmelidir?

Çözüm.

Her parça bir Bernoulli deneyidir: başarı “kusurlu çıkması”, $p = 0,05$, $q = 0,95$. Parçalar bağımsız olduğundan **Teorem 20.1** gereği $X \sim B(20; 0,05)$ ve

$$f_X(x) = \binom{20}{x} (0,05)^x (0,95)^{20-x}, \quad x = 0, 1, \dots, 20.$$

- (a) Küçük x değerleri için olasılıkları hesaplayalım. $\binom{20}{0} = 1$, $\binom{20}{1} = 20$, $\binom{20}{2} = 190$, $\binom{20}{3} = 1140$:

$$\begin{aligned} f_X(0) &= (0,95)^{20} \approx 0,3585, \\ f_X(1) &= 20 \cdot 0,05 \cdot (0,95)^{19} = (0,95)^{19} \approx 0,3774, \\ f_X(2) &= 190 \cdot (0,05)^2 \cdot (0,95)^{18} = 0,475 \cdot (0,95)^{18} \approx 0,1887, \\ f_X(3) &= 1140 \cdot (0,05)^3 \cdot (0,95)^{17} = 0,1425 \cdot (0,95)^{17} \approx 0,0596. \end{aligned}$$

Dolayısıyla

$$P(X \leq 1) = f_X(0) + f_X(1) \approx 0,3585 + 0,3774 = 0,7359, \quad P(X \geq 2) = 1 - P(X \leq 1) \approx 0,2642.$$

Yirmi parçalık bir örnekte iki ya da daha çok kusurlu parça bulunma olasılığı yaklaşık dörtte birdir.

(b) **Teorem 20.2** ile

$$E(X) = 20 \cdot 0,05 = 1, \quad \text{Var}(X) = 20 \cdot 0,05 \cdot 0,95 = 0,95, \quad \sigma_X = \sqrt{0,95} \approx 0,9747.$$

Ortalama olarak yirmi parçada bir kusurlu bulunur.

(c) Tam değer, tümleyen üzerinden:

$$P(X \geq 4) = 1 - (f_X(0) + f_X(1) + f_X(2) + f_X(3)) \approx 1 - 0,9841 = 0,0159.$$

Chebyshev eşitsizliği (**Teorem 16.4**), her $\varepsilon > 0$ için $P(|X - E(X)| \geq \varepsilon) \leq \text{Var}(X)/\varepsilon^2$ der. $E(X) = 1$ olduğundan $(X \geq 4) = (X - 1 \geq 3)$ 'tür; $X \geq 0$ olduğu için $X - 1 \leq -3$ olayı boştur ve dolayısıyla $(X \geq 4) = (|X - 1| \geq 3)$ 'tür. $\varepsilon = 3$ ile

$$P(X \geq 4) = P(|X - 1| \geq 3) \leq \frac{0,95}{9} \approx 0,1056.$$

Sınır doğrudur ama gevşektir: gerçek olasılık 0,0159, sınırın yaklaşık yedide biridir. Chebyshev eşitsizliği dağılım hakkında yalnızca beklenen değer ile varyansı kullandığından bu kabalık beklenir; dağılımı tam olarak bildiğimiz yerde olasılığı doğrudan hesaplamak her zaman daha iyidir.

(d) n parça incelenir; kusurlu sayısı $X_n \sim B(n; 0,05)$ olsun. En az bir kusurlu bulma olasılığı

$$P(X_n \geq 1) = 1 - P(X_n = 0) = 1 - (0,95)^n$$

olduğundan koşul $1 - (0,95)^n \geq 0,9$, yani $(0,95)^n \leq 0,1$ 'dir. Her iki tarafın logaritmasını alınca $n \ln 0,95 \leq \ln 0,1$ olur; $\ln 0,95 < 0$ olduğundan bu sayıya bölerken eşitsizlik yön değiştirir:

$$n \ln 0,95 \leq \ln 0,1 \Leftrightarrow n \geq \frac{\ln 0,1}{\ln 0,95} \approx \frac{-2,3026}{-0,0513} \approx 44,89.$$

Öyleyse en az 45 parça incelenmelidir. Sağlama: $(0,95)^{44} \approx 0,1047 > 0,1$, ama $(0,95)^{45} \approx 0,0994 < 0,1$.

■

20.6 En Olası Değer

Örnek 20.1'ndeki tabloda olasılıklar en büyük değerinden sonra düzenli olarak azalıyordu; $f_X(0) = f_X(1)$ eşitliği de dikkat çekiciydi. Bu davranış geneldir: binom olasılıkları bir tepeye kadar artar, sonra azalır. Tepenin yerini, yani olasılık fonksiyonunun en büyük değerini aldığı noktayı — **en olası değeri** ya da **modu** — kesin olarak belirleyebiliriz. Anahtar, ardışık olasılıkların oranıdır: oran 1'den büyükken dizi yükselir, 1'den küçükken alçalır.

Önerme 20.3 (Binom Dağılımının En Olası Değeri). $X \sim B(n, p)$ ve $0 < p < 1$ olsun. $x = 1, 2, \dots, n$ için

$$\frac{f_X(x)}{f_X(x-1)} = \frac{n-x+1}{x} \cdot \frac{p}{1-p}$$

ve dolayısıyla

$$f_X(x) > f_X(x-1) \Leftrightarrow x < (n+1)p, \quad f_X(x) = f_X(x-1) \Leftrightarrow x = (n+1)p, \quad f_X(x) < f_X(x-1) \Leftrightarrow x > (n+1)p.$$

Sonuç olarak:

(a) $(n+1)p$ bir tam sayı değilse f_X en büyük değerini yalnızca $x^* = \lfloor (n+1)p \rfloor$ noktasında alır; burada $\lfloor \cdot \rfloor$ tam kısım fonksiyonudur.

(b) $(n+1)p = m$ bir tam sayı ise f_X en büyük değerini tam olarak iki noktada, $m-1$ ve m 'de alır ve $f_X(m-1) = f_X(m)$ 'dir.

Her iki durumda da en olası değer $np - q \leq x^* \leq np + p$ aralığındadır; yani beklenen değer np 'den en çok 1 uzaktadır.

İspat.

Oran. $1 \leq x \leq n$ için

$$\frac{\binom{n}{x}}{\binom{n}{x-1}} = \frac{n!}{x!(n-x)!} \cdot \frac{(x-1)!(n-x+1)!}{n!} = \frac{n-x+1}{x}$$

olduğundan

$$\frac{f_X(x)}{f_X(x-1)} = \frac{\binom{n}{x} p^x q^{n-x}}{\binom{n}{x-1} p^{x-1} q^{n-x+1}} = \frac{n-x+1}{x} \cdot \frac{p}{q}.$$

$0 < p < 1$ olduğundan bütün $f_X(x)$ değerleri pozitifdir ve oran tanımlıdır.

İşaret. Oranın 1 ile karşılaştırılması, $xq > 0$ ile çarpılarak bir doğrusal eşitsizliğe dönüşür:

$$\frac{f_X(x)}{f_X(x-1)} > 1 \Leftrightarrow (n-x+1)p > xq = x(1-p) \Leftrightarrow np - xp + p > x - xp \Leftrightarrow (n+1)p > x.$$

Aynı adımlar “>” yerine “=” ve “<” ile de geçerlidir; bu, ifadedeki üç denkliği verir. Demek ki $f_X(0), f_X(1), \dots, f_X(n)$ dizisi, $x < (n+1)p$ olduğu sürece kesin artar, $x > (n+1)p$ olduğu andan itibaren kesin azalır.

(a) $(n+1)p$ tam sayı olmasın ve $x^* = \lfloor (n+1)p \rfloor$ olsun. $0 < (n+1)p < n+1$ olduğundan $0 \leq x^* \leq n$ 'dir. $1 \leq x \leq x^*$ olan her x için $x \leq x^* < (n+1)p$, dolayısıyla $f_X(x) > f_X(x-1)$:

$$f_X(0) < f_X(1) < \dots < f_X(x^*).$$

$x^* + 1 \leq x \leq n$ olan her x için $x \geq x^* + 1 > (n+1)p$, dolayısıyla $f_X(x) < f_X(x-1)$:

$$f_X(x^*) > f_X(x^* + 1) > \dots > f_X(n).$$

Öyleyse en büyük değer yalnızca x^* 'da alınır.

(b) $(n+1)p = m$ tam sayı olsun; $0 < m < n+1$ olduğundan $1 \leq m \leq n$ 'dir. $x < m$ için dizi kesin artar, $x = m$ için $f_X(m) = f_X(m-1)$ ve $x > m$ için kesin azalır:

$$f_X(0) < \dots < f_X(m-1) = f_X(m) > f_X(m+1) > \dots > f_X(n).$$

En büyük değer tam olarak $m-1$ ve m noktalarında alınır.

Sınırlar. (a) durumunda tam kısmın tanımı gereği $(n+1)p - 1 < x^* \leq (n+1)p$, yani $np - q < x^* \leq np + p$ 'dir. (b) durumunda iki mod $m-1 = (n+1)p - 1 = np - q$ ve $m = np + p$ 'dir. Her iki durumda da $np - q \leq x^* \leq np + p$ sağlanır; aralığın uzunluğu $p + q = 1$ 'dir.

■

Sonucun pratik özeti: **en olası değer, $(n+1)p$ 'nin tam kısmıdır**; $(n+1)p$ tam sayıysa onun bir eksiği de aynı olasılığa sahiptir. Beklenen değer np genellikle tam sayı değildir ve zaten “alınan bir değer” olmak zorunda değildir; en olası değer ise her zaman D_X 'in bir elemanıdır ve beklenen değerın hemen yanında durur.

Örnek 20.3 (En Olası Değer Hesapları). (a) $n = 10$ ve $p = 0, 2$; $p = 0, 5$; $p = 0, 8$ için binom dağılımının en olası değerini bulunuz ve olasılık fonksiyonlarını karşılaştırınız.

(b) Örnek 20.1'ndeki $(\mathcal{B}, \frac{1}{6})$ dağılımında $f_X(0) = f_X(1)$ eşitliğini Önerme 20.3 ile açıklayınız.

(c) Hilesiz bir para dokuz kez atıldığında tura sayısının en olası değerini bulunuz.

Çözüm.

(a) Üç durumda da $n+1 = 11$ 'dir:

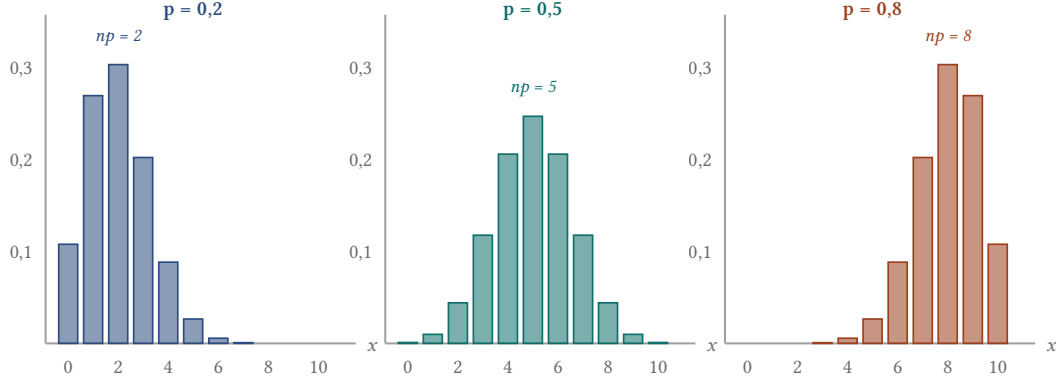
- $p = 0, 2$ için $(n+1)p = 2, 2$; tam sayı değildir, en olası değer $[2, 2] = 2$ 'dir. Beklenen değer $np = 2$ 'dir.
- $p = 0, 5$ için $(n+1)p = 5, 5$; en olası değer $[5, 5] = 5$ 'tir. Beklenen değer 5'tir.
- $p = 0, 8$ için $(n+1)p = 8, 8$; en olası değer $[8, 8] = 8$ 'dir. Beklenen değer 8'dir.

Olasılık fonksiyonlarını dört ondalıkla hesaplayalım ($f(x) = \binom{10}{x} p^x (1-p)^{10-x}$):

x	0	1	2	3	4	5	6	7	8	9	10
$p = 0, 2$	0,1074	0,2684	0,3020	0,2013	0,0881	0,0264	0,0055	0,0008	0,0001	0,0000	0,0000
$p = 0, 5$	0,0010	0,0098	0,0439	0,1172	0,2051	0,2461	0,2051	0,1172	0,0439	0,0098	0,0010
$p = 0, 8$	0,0000	0,0000	0,0001	0,0008	0,0055	0,0264	0,0881	0,2013	0,3020	0,2684	0,1074

Tablo üç şeyi gösterir. Birincisi, her satır tek tepelidir ve tepe, Önerme 20.3'un öngördüğü yerdedir. İkincisi, $p = 0, 5$ satırı $x = 5$ etrafında simetriktir; çünkü $\binom{10}{x} = \binom{10}{10-x}$ ve $p = q$ olduğundan $f(x) = f(10-x)$ 'tir. Üçüncüsü, $p = 0, 8$ satırı $p = 0, 2$ satırının ters çevrilmişidir: $B(10; 0, 8)$ 'in x 'teki olası-

lığı, $B(10; 0, 2)$ 'nin $10 - x$ 'teki olasılığına eşittir. Bunun genel nedeni, $X \sim B(n, p)$ iken $n - X \sim B(n, 1 - p)$ olmasıdır (Alıştırma 20.1, (e) şıkkı): başarı ile başarısızlığın adları değiştirildiğinde başarı sayısı $n - X$ olur.



Şekil 20.2: $n = 10$ için binom olasılık fonksiyonu. Kütle np çevresinde toplanır: $p = 0,5$ 'te simetrik, küçük p 'de sağa, büyük p 'de sola çarpıktır. Çubukların toplam yüksekliği her panelde 1'dir.

(b) $n = 5, p = \frac{1}{6}$ için $(n + 1)p = 6 \cdot \frac{1}{6} = 1$ bir tam sayıdır. Önerme 20.3 (b) gereği en büyük olasılık tam olarak $m - 1 = 0$ ve $m = 1$ noktalarında alınır ve $f_X(0) = f_X(1)$ 'dir. Tabloda gördüğümüz $f_X(0) = f_X(1) = \frac{3125}{7776}$ eşitliği budur. Oranla da doğrulanır: $\frac{f_X(1)}{f_X(0)} = \frac{5-1+1}{1} \cdot \frac{1/6}{5/6} = 5 \cdot \frac{1}{5} = 1$.

(c) Tura sayısı $X \sim (B, \frac{1}{2})$ ve $(n + 1)p = 10 \cdot \frac{1}{2} = 5$ tam sayıdır. En olası değerler 4 ve 5'tir:

$$f_X(4) = \binom{9}{4} \frac{1}{2^9} = \frac{126}{512} = \frac{63}{256}, \quad f_X(5) = \binom{9}{5} \frac{1}{2^9} = \frac{126}{512} = \frac{63}{256}.$$

Beklenen değer $np = 4,5$ alınabilen bir değer değildir; en olası değerler onun iki yanındaki tam sayılardır. Tek sayıda hilesiz para atışında durum hep böyledir: n tek ise $(n + 1)/2$ tam sayıdır ve $\frac{n-1}{2}$ ile $\frac{n+1}{2}$ eşit olasılıklı iki moddur.

■

20.7 Alıştırmalar

Alıştırma 20.1 (Bernoulli ve Binom Dağılımları Üzerine Alıştırmalar). (a) $X \sim (B, \frac{1}{3})$ olsun. Olasılık fonksiyonunu ve dağılım fonksiyonunu tablo hâlinde yazınız; $P(1 \leq X \leq 3)$, $E(X)$ ve $\text{Var}(X)$ 'i bulunuz.

(b) Hilesiz bir para sekiz kez atılıyor. Tam dört kez tura gelme olasılığını ve turaların yazılardan çok olma olasılığını bulunuz.

(c) $X \sim B(1, p)$ için $\text{Var}(X) \leq \frac{1}{4}$ olduğunu ve eşitliğin yalnızca $p = \frac{1}{2}$ 'de sağlandığını gösteriniz. Buradan $Y \sim B(n, p)$ için $\text{Var}(Y) \leq \frac{n}{4}$ olduğunu çıkarınız.

(d) Bir X rastgele değişkeninin moment üreten fonksiyonu $M_X(t) = (\frac{1}{4} + \frac{3}{4}e^t)^6$ olsun. X 'in dağılımını belirleyiniz; $E(X)$, $\text{Var}(X)$ ve $P(X \geq 5)$ 'i bulunuz.

(e) $X \sim B(n, p)$ ise $n - X \sim B(n, 1 - p)$ olduğunu hem olasılık fonksiyonuyla hem de moment üreten fonksiyonla gösteriniz.

(f) Hilesiz bir zar kaç kez atılmalıdır ki en az bir kez 6 gelme olasılığı 0,9'u aşsın?

(g) X ve Y bağımsız ve her ikisi de $(B, \frac{1}{2})$ dağılımlı olsun. $P(X + Y = 3)$ 'ü önce Teorem 20.4 ile, sonra konvolüsyon formülüyle hesaplayınız.

(h) $(B1, \frac{1}{4})$ ve $B(15; 0, 3)$ dağılımlarının en olası değerlerini bulunuz.

Çözüm.

(a) $f_X(x) = \binom{4}{x} (\frac{1}{3})^x (\frac{2}{3})^{4-x} = \frac{\binom{4}{x} 2^{4-x}}{81}$. Katsayılar 1, 4, 6, 4, 1; kuvvetler 16, 8, 4, 2, 1:

x	0	1	2	3	4
$f_X(x)$	$\frac{16}{81}$	$\frac{32}{81}$	$\frac{24}{81}$	$\frac{8}{81}$	$\frac{1}{81}$
$F_X(x)$	$\frac{16}{81}$	$\frac{48}{81}$	$\frac{72}{81}$	$\frac{80}{81}$	1

Sağlama: $16 + 32 + 24 + 8 + 1 = 81$. $P(1 \leq X \leq 3) = F_X(3) - F_X(0) = \frac{80-16}{81} = \frac{64}{81}$. **Teorem 20.2** ile $E(X) = 4 \cdot \frac{1}{3} = \frac{4}{3}$ ve $\text{Var}(X) = 4 \cdot \frac{1}{3} \cdot \frac{2}{3} = \frac{8}{9}$. En olası değer $\lfloor 5 \cdot \frac{1}{3} \rfloor = 1$ 'dir; tabloyla uyuyor.

(b) Tura sayısı $X \sim (\mathcal{B}, \frac{1}{2})$, $f_X(x) = \binom{8}{x}/256$.

$$P(X = 4) = \frac{\binom{8}{4}}{256} = \frac{70}{256} = \frac{35}{128} \approx 0,2734.$$

Turalar yazılardan çoksa $X > 8 - X$, yani $X \geq 5$ 'tir:

$$P(X \geq 5) = \frac{\binom{8}{5} + \binom{8}{6} + \binom{8}{7} + \binom{8}{8}}{256} = \frac{56 + 28 + 8 + 1}{256} = \frac{93}{256} \approx 0,3633.$$

Simetri gereği $P(X \leq 3)$ de $\frac{93}{256}$ 'dir ve $\frac{93}{256} + \frac{70}{256} + \frac{93}{256} = 1$ çıkar.

(c) **Önerme 20.1** ile $\text{Var}(X) = p(1-p) = p - p^2$. Kareye tamamlayarak

$$p - p^2 = \frac{1}{4} - \left(p^2 - p + \frac{1}{4}\right) = \frac{1}{4} - \left(p - \frac{1}{2}\right)^2 \leq \frac{1}{4},$$

ve eşitlik ancak $\left(p - \frac{1}{2}\right)^2 = 0$, yani $p = \frac{1}{2}$ iken sağlanır. $Y \sim B(n, p)$ için **Teorem 20.2** gereği $\text{Var}(Y) = np(1-p) \leq n \cdot \frac{1}{4} = \frac{n}{4}$ 'tür.

(d) $q = \frac{1}{4}$, $p = \frac{3}{4}$ ve $n = 6$ ile $M_X(t) = (q + pe^t)^6, (\mathcal{B}, \frac{3}{4})$ dağılımının moment üreten fonksiyonudur (**Teorem 20.3**). Moment üreten fonksiyon dağılımı tek biçimde belirlediğinden (**Teorem 17.5**) $X \sim (\mathcal{B}, \frac{3}{4})$ 'tür. Buradan $E(X) = 6 \cdot \frac{3}{4} = \frac{9}{2}$ ve $\text{Var}(X) = 6 \cdot \frac{3}{4} \cdot \frac{1}{4} = \frac{9}{8}$. Ayrıca

$$P(X \geq 5) = \binom{6}{5} \left(\frac{3}{4}\right)^5 \frac{1}{4} + \left(\frac{3}{4}\right)^6 = \frac{6 \cdot 243 + 729}{4096} = \frac{1458 + 729}{4096} = \frac{2187}{4096} \approx 0,5339.$$

(e) $Y = n - X$ olsun; Y de $\{0, 1, \dots, n\}$ 'de değer alır. y bu kümede iken

$$P(Y = y) = P(X = n - y) = \binom{n}{n-y} p^{n-y} q^y = \binom{n}{y} q^y p^{n-y} = \binom{n}{y} (1-p)^y (1 - (1-p))^{n-y},$$

çünkü $\binom{n}{n-y} = \binom{n}{y}$ 'dir. Bu, $B(n, 1-p)$ 'nin olasılık fonksiyonudur. Moment üreten fonksiyonla: $e^{tY} = e^{nt} e^{-tX}$ olduğundan

$$M_Y(t) = e^{nt} M_X(-t) = e^{nt} (q + pe^{-t})^n = (e^t q + p)^n = (p + (1-p)e^t)^n,$$

ve bu, $B(n, 1-p)$ 'nin moment üreten fonksiyonudur; **Teorem 17.5** sonucu verir.

(f) n atışta 6 gelme sayısı $X_n \sim (\mathcal{B}, \frac{1}{6})$ ve $P(X_n \geq 1) = 1 - \left(\frac{5}{6}\right)^n$ 'dir. Koşul $\left(\frac{5}{6}\right)^n \leq 0,1$, yani

$$n \geq \frac{\ln 0,1}{\ln(5/6)} \approx \frac{-2,3026}{-0,1823} \approx 12,63.$$

En az 13 atış gerekir. Sağlama: $\left(\frac{5}{6}\right)^{12} \approx 0,1122$, dolayısıyla 12 atışta olasılık $\approx 0,888 < 0,9$; $\left(\frac{5}{6}\right)^{13} \approx 0,0935$, dolayısıyla 13 atışta olasılık $\approx 0,907 > 0,9$.

(g) **Teorem 20.4** gereği $X + Y \sim (\mathcal{B}, \frac{1}{2})$ ve

$$P(X + Y = 3) = \binom{6}{3} \frac{1}{2^6} = \frac{20}{64} = \frac{5}{16}.$$

Konvolüsyonla (**Önerme 15.3**), $f_X(k) = f_Y(k) = \binom{3}{k}/8$ olduğundan

$$P(X + Y = 3) = \sum_{k=0}^3 f_X(k) f_Y(3-k) = \frac{1}{64} \sum_{k=0}^3 \binom{3}{k} \binom{3}{3-k} = \frac{1 \cdot 1 + 3 \cdot 3 + 3 \cdot 3 + 1 \cdot 1}{64} = \frac{20}{64} = \frac{5}{16}.$$

Toplamın $\binom{6}{3} = 20$ çıkması Vandermonde özdeşliğinin $n = m = k = 3$ hâlidir.

(h) $(B1, \frac{1}{4})$ için $(n+1)p = 12 \cdot \frac{1}{4} = 3$ tam sayıdır; [Önerme 20.3](#) (b) gereği en olası değerler 2 ve 3'tür. Doğrulama: $f(2) = \binom{11}{2} \frac{3^9}{4^{11}} = \frac{55 \cdot 19683}{4^{11}}$ ve $f(3) = \binom{11}{3} \frac{3^8}{4^{11}} = \frac{165 \cdot 6561}{4^{11}}$; paylar $55 \cdot 19683 = 1\,082\,565 = 165 \cdot 6561$ ile eşittir ve ortak değer yaklaşık 0,2581'dir. $B(15; 0, 3)$ için $(n+1)p = 16 \cdot 0,3 = 4,8$ tam sayı değildir; tek en olası değer $[4, 8] = 4$ 'tür ($f(4) \approx 0,2186$; komşuları $f(3) \approx 0,1700$ ve $f(5) \approx 0,2061$).

■

Binom dağılımında n büyürken p küçülüyor ve np çarpımı sabit kalıyorsa olasılıklar tek bir parametreye bağlı yeni bir dağılıma yaklaşır; nadir olayların sayısını betimleyen bu dağılımı bir sonraki bölümde inceliyoruz: [Poisson Dağılımı](#).

21. Poisson Dağılımı

Bernoulli ve Binom Dağılımları bölümünde, n bağımsız denemede gözlenen başarı sayısının binom dağılımına sahip olduğunu gördük. Binom dağılımının iki parametresi vardır: deneme sayısı n ve başarı olasılığı p . Uygulamada sık karşılaşılan bir durum, n 'nin çok büyük, p 'nin çok küçük, ama çarpımları np 'nin ılımlı bir sayı olmasıdır. Bir kitabın bir sayfasındaki baskı hataları, bir çağrı merkezine bir dakikada gelen çağrılar, bir radyoaktif maddenin bir saniyede yaydığı parçacıklar, bir şehirde bir günde olan trafik kazaları bu türdendir. “Deneme” sayısı (harfler, saniyeler, atomlar, sürücüler) çok büyüktür; her denemede “başarı” (hata, çağrı, bozunma, kaza) olasılığı çok küçüktür; ama ortalama başarı sayısı 1, 3 ya da 10 gibi makul bir sayıdır.

Böyle durumlarda binom olasılıklarını doğrudan hesaplamak zahmetlidir ve aslında gereksizdir: $n \rightarrow \infty$, $p \rightarrow 0$ ve $np \rightarrow \lambda$ iken binom olasılıkları, yalnızca λ 'ya bağlı olan bir limite yaklaşır. Bu limit dağılım **Poisson dağılımı**dır; adını, bu limiti 1837’de inceleyen Siméon Denis Poisson’dan alır. Poisson dağılımı, bir zaman aralığında ya da bir bölgede gözlenen “nadir olay” sayılarının doğal modeli olduğundan olasılık teorisinin en çok kullanılan kesikli dağılımlarından biridir.

Bu bölümde şunları öğreneceğiz: Poisson dağılımının tanımı ve olasılık fonksiyonunun toplamının bir olduğunun üstel seriyle gösterilmesi; ardışık olasılıkların oranı ve dağılımın en olası değeri; beklenen değerlerin de varyansın da λ olduğu; moment üreten, karakteristik ve çarpımsal moment üreten fonksiyonlar; bağımsız Poisson değişkenlerinin toplamının yine Poisson olduğu; ve binom dağılımına Poisson yaklaşımının her çarpanın limitini ayrı ayrı bulan tam ispatı. Her sonucu sayısal örneklerle sınayacağız.

21.1 Tanım ve Olasılık Fonksiyonu

Poisson dağılımını önce doğrudan, bir olasılık fonksiyonu formülüyle tanımlayacağız; binom dağılımıyla bağımlı bölümün sonunda kuracağız. Formülün nereden geldiğini şimdiden sezmek için şu kadarını akılda tutmak yeterlidir: binom olasılığındaki $(1-p)^{n-x}$ çarpanı, $p = \lambda/n$ ve n büyük iken $e^{-\lambda}$ ’ya, $\binom{n}{x}p^x$ çarpanı ise $\lambda^x/x!$ ’e yaklaşır.

Tanım 21.1 (Poisson Dağılımı). $\lambda > 0$ bir sabit olsun. Değer kümesi $D_X = \{0, 1, 2, \dots\}$ olan bir X kesikli rastgele değişkeninin olasılık fonksiyonu

$$f_X(x) = \begin{cases} \frac{e^{-\lambda} \lambda^x}{x!}, & x = 0, 1, 2, \dots \\ 0, & \text{diğer durumlarda} \end{cases}$$

ise X , λ parametrelili Poisson dağılımına sahiptir denir ve $X \sim \text{Poisson}(\lambda)$ yazılır.

Tanım üzerine birkaç gözlem. Değer kümesi sonsuz sayılabilir bir kümedir; binom dağılımından farklı olarak X 'in alabileceği değerlerin bir üst sınırı yoktur. Dağılımın tek parametresi λ 'dır ve λ 'nın tam sayı olması gerekmez. $0! = 1$ ve $\lambda^0 = 1$ olduğundan $f_X(0) = e^{-\lambda}$ ’dır; bu, “hiç olay olmaması” olasılığıdır ve Poisson hesaplarının çoğunda başlangıç noktasıdır.

Tanımda f_X ’e “olasılık fonksiyonu” dedik; ama bir fonksiyona bu adı verebilmek için değerlerinin negatif olmaması ve toplamının 1 olması gerekir (**Önerme 10.1**). Bunu doğrulamak zorundayız; hesabın tamamı üstel fonksiyonun kuvvet serisine dayanır.

Önerme 21.1 (Poisson Olasılık Fonksiyonunun Geçerliliği). $\lambda > 0$ olsun. $x = 0, 1, 2, \dots$ için $f(x) = \frac{e^{-\lambda} \lambda^x}{x!}$ ile tanımlanan fonksiyon şu iki koşulu sağlar:

(i) her $x = 0, 1, 2, \dots$ için $f(x) > 0$;

(ii) $\sum_{x=0}^{\infty} f(x) = 1$.

Dolayısıyla f , değer kümesi $\{0, 1, 2, \dots\}$ olan bir kesikli rastgele değişkenin olasılık fonksiyonudur.

İspat.

(i) $e^{-\lambda} > 0$, $\lambda^x > 0$ ve $x! > 0$ olduğundan üçünün çarpımı ve bölümü de pozitiftir.

(ii) Üstel fonksiyonun kuvvet serisi her $u \in \mathbb{R}$ için yakınsaktır ve

$$e^u = \sum_{x=0}^{\infty} \frac{u^x}{x!}$$

eşitliği geçerlidir. $e^{-\lambda}$ çarpanı x 'e bağlı olmayan bir sabit olduğundan toplamın dışına alınabilir; $u = \lambda$ alınırsa

$$\sum_{x=0}^{\infty} f(x) = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^{-\lambda} \cdot e^{\lambda} = 1$$

bulunur. Böylece f , **Önerme 10.1**'ndeki (1) ve (2) koşullarını sağlar. Böyle bir f 'nin gerçekten bir rastgele değişkenin olasılık fonksiyonu olduğunu görmek için $\Omega = \{0, 1, 2, \dots\}$, $\mathcal{U} = \mathcal{P}(\Omega)$ alalım ve her $\omega \in \Omega$ noktasına $p_{\omega} = f(\omega)$ ağırlığını verelim; **Önerme 4.3** gereği $P(A) = \sum_{\omega \in A} f(\omega)$ bir olasılık ölçüsüdür. $X(\omega) = \omega$ ile tanımlanan rastgele değişken için $P(X = x) = P(\{x\}) = f(x)$ olur; yani X 'in olasılık fonksiyonu (**Tanım 10.2**) tam olarak f 'dir. ■

$X \sim \text{Poisson}(\lambda)$ değişkeninin dağılım fonksiyonu (**Tanım 9.1**), $x < 0$ için $F_X(x) = 0$ ve $x \geq 0$ için

$$F_X(x) = P(X \leq x) = e^{-\lambda} \sum_{k=0}^{\lfloor x \rfloor} \frac{\lambda^k}{k!}$$

olan bir basamak fonksiyonudur; burada $\lfloor x \rfloor$, x 'i aşmayan en büyük tam sayıdır. Bu toplamın temel fonksiyonlarla yazılabilen kapalı bir formülü yoktur; olasılıklar terim terim toplanarak bulunur. En çok kullanılan özel durum, “en az bir olay” olasılığıdır:

$$P(X \geq 1) = 1 - P(X = 0) = 1 - e^{-\lambda}.$$

21.1.1 Ardışık Oran ve Mod

Poisson olasılıklarını art arda hesaplarken her seferinde faktöriyel ve kuvvet almak gerekmez: bir terimden bir sonrakine tek bir çarpma ile geçilir. Aynı gözlem dağılımın biçimini de belirler: olasılıklar önce artar, sonra azalır ve en yüksek değerini λ 'nın yakınında alır.

Önerme 21.2 (Ardışık Oran ve Mod). $X \sim \text{Poisson}(\lambda)$ ve $f = f_X$ olsun.

- (i) Her $x = 0, 1, 2, \dots$ için $\frac{f(x+1)}{f(x)} = \frac{\lambda}{x+1}$.
- (ii) $x + 1 < \lambda$ iken $f(x+1) > f(x)$; $x + 1 = \lambda$ iken $f(x+1) = f(x)$; $x + 1 > \lambda$ iken $f(x+1) < f(x)$.
- (iii) $m = \lfloor \lambda \rfloor$ olsun. λ tam sayı değilse f en büyük değerini yalnızca $x = m$ noktasında alır. λ tam sayı ise f en büyük değerini $x = \lambda - 1$ ve $x = \lambda$ noktalarında alır ve $f(\lambda - 1) = f(\lambda)$ 'dir. Yani X 'in **modu** (en olası değeri) $\lfloor \lambda \rfloor$ 'dir; λ tam sayı olduğunda $\lambda - 1$ de moddur.

İspat.

(i) $(x+1)! = (x+1) \cdot x!$ olduğundan

$$\frac{f(x+1)}{f(x)} = \frac{e^{-\lambda} \lambda^{x+1} / (x+1)!}{e^{-\lambda} \lambda^x / x!} = \frac{\lambda^{x+1}}{\lambda^x} \cdot \frac{x!}{(x+1)!} = \frac{\lambda}{x+1}.$$

(ii) $f(x) > 0$ olduğundan $f(x+1) > f(x)$ eşitsizliği, oranın 1'den büyük olmasına, yani $\frac{\lambda}{x+1} > 1$, yani $x+1 < \lambda$ olmasına denktir. Aynı nedenle $f(x+1) = f(x)$ eşitliği $\frac{\lambda}{x+1} = 1$, yani $x+1 = \lambda$ olmasına; $f(x+1) < f(x)$ eşitsizliği ise $\frac{\lambda}{x+1} < 1$, yani $x+1 > \lambda$ olmasına denktir. ($x+1 > 0$ olduğundan $\frac{\lambda}{x+1}$ ile 1'i karşılaştırmak, λ ile $x+1$ 'i karşılaştırmakla aynı şeydir.)

(iii) $m = \lfloor \lambda \rfloor$ için $m \leq \lambda < m+1$ olduğunu hatırlayalım. Önce λ 'nın tam sayı olmadığını varsayalım; o zaman $m < \lambda$ 'dır. $x = 0, 1, \dots, m-1$ için $x+1 \leq m < \lambda$ olduğundan (ii) gereği

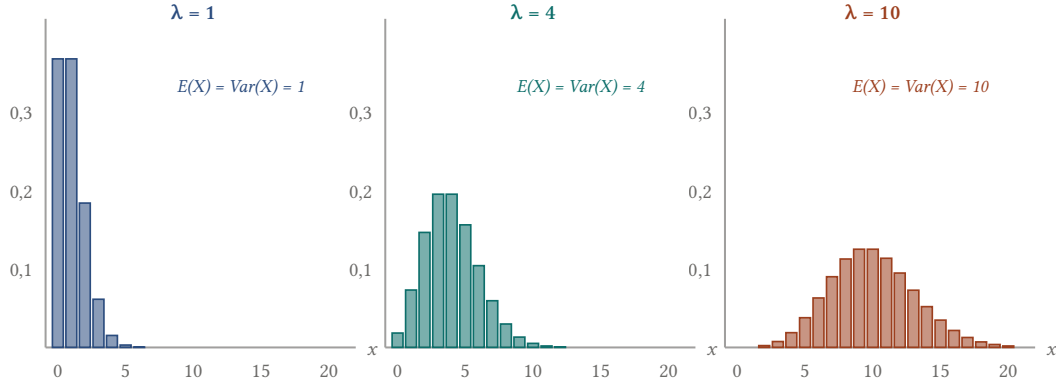
$$f(0) < f(1) < \dots < f(m)$$

olur. $x \geq m$ için $x+1 \geq m+1 > \lambda$ olduğundan

$$f(m) > f(m+1) > f(m+2) > \dots$$

olur. Demek ki f en büyük değerini yalnızca m 'de alır.

Şimdi λ tam sayı olsun; o zaman $m = \lambda$ 'dır. $x = 0, 1, \dots, m-2$ için $x+1 \leq m-1 < \lambda$ olduğundan $f(0) < f(1) < \dots < f(m-1)$; $x = m-1$ için $x+1 = \lambda$ olduğundan $f(m) = f(m-1)$; $x \geq m$ için $x+1 > \lambda$ olduğundan $f(m) > f(m+1) > \dots$ bulunur. Demek ki en büyük değer tam olarak $m-1 = \lambda-1$ ve $m = \lambda$ noktalarında alınır.



Şekil 21.1: Poisson olasılık fonksiyonu. λ küçükken kütle sıfıra yakın ve sağa çarpık; λ büyüdükçe dağılım λ çevresinde simetrikleşir ve normal eğriye benzemeye başlar. Ortalama da varyans da λ 'dır.

■

Şekil, üç ayrı λ için dağılımın biçimini gösterir. $\lambda = 1, 4$ ve 10 tam sayı olduklarından her birinde en yüksek iki çubuk eşit uzunluktadır: sırasıyla $x = 0$ ile 1 , $x = 3$ ile 4 ve $x = 9$ ile 10 . $\lambda < 1$ olduğunda mod 0 'dır ve olasılıklar baştan itibaren azalır. λ büyüdükçe dağılım sağa kayar, yayılır ve giderek simetrik bir çan biçimine yaklaşır; bu biçimin adı olan normal dağılımı ileride tanıyacağız. Yayılmanın nedeni ise bir sonraki teoremden ortaya çıkacak: varyans da λ 'dır.

Tanımı ve ardışık oranı ilk örnekte kullanalım. Örnek, parametrenin “ortalama olay sayısı” olarak yorumlanmasına dayanır; bu yorumun doğruluğunu Teorem 21.1’nde ispatlayacağız.

Örnek 21.1 (Sayfa Başına Baskı Hatası). Bir kitabın rastgele seçilen bir sayfasındaki baskı hatası sayısı X , sayfa başına ortalama $\frac{1}{2}$ hata olacak biçimde Poisson dağılımına sahip olsun; yani $X \sim \text{Poisson}(\frac{1}{2})$.

- Sayfanın hatasız olma olasılığını bulunuz.
- Sayfada en az iki hata bulunma olasılığını bulunuz.
- Bir sayfadaki en olası hata sayısını bulunuz ve $x = 0, 1, 2, 3, 4$ için olasılıkları tablo hâlinde veriniz.
- Kitap 400 sayfa ise ve sayfalardaki hata sayıları birbirinden bağımsızsa, hatasız sayfa sayısının beklenen değerini bulunuz.

Çözüm.

(a) $\lambda = \frac{1}{2}$ için

$$P(X = 0) = e^{-1/2} \approx 0,6065.$$

Sayfaların yaklaşık yüzde 61’i hatasızdır.

(b) Tümleyen olaydan gidelim:

$$P(X \geq 2) = 1 - P(X = 0) - P(X = 1) = 1 - e^{-1/2} - e^{-1/2} \cdot \frac{1}{2} = 1 - \frac{3}{2}e^{-1/2} \approx 1 - 0,9098 = 0,0902.$$

(c) $\lambda = \frac{1}{2}$ tam sayı olmadığından Önerme 21.2 gereği mod $\lfloor \frac{1}{2} \rfloor = 0$ 'dır: en olası durum sayfanın hatasız olmasıdır. Olasılıkları ardışık oranla üretelim: $f(x+1) = f(x) \cdot \frac{\lambda}{x+1}$ ve $\lambda = \frac{1}{2}$ olduğundan her terim bir öncekinin $\frac{1}{2(x+1)}$ katıdır.

x	0	1	2	3	4
$f_X(x)$	$e^{-1/2}$	$\frac{1}{2}e^{-1/2}$	$\frac{1}{8}e^{-1/2}$	$\frac{1}{48}e^{-1/2}$	$\frac{1}{384}e^{-1/2}$
yaklaşık	0,6065	0,3033	0,0758	0,0126	0,0016

İlk beş olasılığın toplamı 0,9998'dir; kalan 0,0002, beş ve daha çok hatalı sayfaların payıdır.

(d) Her sayfa, ötekilerden bağımsız olarak, $p = e^{-1/2}$ olasılıkla hatasızdır. Dolayısıyla 400 sayfanın her biri bir Bernoulli denemesidir (Tanım 20.2) ve hatasız sayfa sayısı $Y \sim B(400, e^{-1/2})$ 'dir (Tanım 20.3). Teorem 20.2 gereği

$$E(Y) = 400 e^{-1/2} \approx 242,6.$$

Kitapta ortalama 243 dolayında hatasız sayfa beklenir.

■

21.2 Beklenen Değer ve Varyans

Şimdiye dek λ 'ya "ortalama olay sayısı" dedik; bunu ispatlamanın zamanı geldi. Hesabı doğrudan beklenen değer tanımıyla yapacağız. Varyans için ise $E(X^2)$ 'yi doğrudan hesaplamak yerine **çarpımsal momentleri** kullanacağız: **Beklenen Değer, Momentler ve Varyans** bölümünde tanımlanan (Tanım 16.2) k . çarpımsal moment $E[X(X-1)\cdots(X-k+1)]$, Poisson dağılımında olağanüstü sade çıkar, çünkü $\frac{x(x-1)\cdots(x-k+1)}{x!} = \frac{1}{(x-k)!}$ sadeleşmesi faktöriyeli ortadan kaldırır.

Teorem 21.1 (Poisson Dağılımının Momentleri). $X \sim \text{Poisson}(\lambda)$ olsun.

(i) $E(X) = \lambda$.

(ii) Her $k = 1, 2, \dots$ için k . çarpımsal moment $E[X(X-1)\cdots(X-k+1)] = \lambda^k$ 'dir. Özel olarak $E[X(X-1)] = \lambda^2$.

(iii) $\text{Var}(X) = \lambda$ ve $\sigma_X = \sqrt{\lambda}$.

İspat.

(i) Tanım 16.1 gereği $E(X) = \sum_{x=0}^{\infty} x f_X(x)$ 'tir. $x = 0$ terimi sıfırdır; $x \geq 1$ için $\frac{x}{x!} = \frac{1}{(x-1)!}$ olduğundan

$$E(X) = \sum_{x=1}^{\infty} x \cdot \frac{e^{-\lambda} \lambda^x}{x!} = \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} = \lambda e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^y}{y!} = \lambda e^{-\lambda} e^{\lambda} = \lambda.$$

Üçüncü eşitlikte $y = x - 1$ yazdık; son toplam üstel seridir. Serinin terimleri pozitif ve toplamı sonlu olduğundan seri mutlak yakınsaktır; yani beklenen değer vardır.

(ii) $x < k$ için $x(x-1)\cdots(x-k+1)$ çarpımının çarpanlarından biri sıfırdır; bu terimler toplama katkısı vermez. $x \geq k$ için

$$\frac{x(x-1)\cdots(x-k+1)}{x!} = \frac{x!}{(x-k)!x!} = \frac{1}{(x-k)!}$$

olduğundan

$$E[X(X-1)\cdots(X-k+1)] = \sum_{x=k}^{\infty} \frac{e^{-\lambda} \lambda^x}{(x-k)!} = \lambda^k e^{-\lambda} \sum_{x=k}^{\infty} \frac{\lambda^{x-k}}{(x-k)!} = \lambda^k e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^y}{y!} = \lambda^k e^{-\lambda} e^{\lambda} = \lambda^k$$

bulunur; burada $y = x - k$ yazdık. $k = 2$ için $E[X(X-1)] = \lambda^2$ 'dir.

(iii) $X^2 = X(X-1) + X$ olduğundan beklenen değer doğrusallığıyla

$$E(X^2) = E[X(X-1)] + E(X) = \lambda^2 + \lambda.$$

Önerme 16.3 gereği

$$\text{Var}(X) = E(X^2) - [E(X)]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

ve $\sigma_X = \sqrt{\text{Var}(X)} = \sqrt{\lambda}$ olur.

İ Ortalama ile varyans eşittir

Poisson dağılımının ayırt edici özelliği, beklenen değer ile varyansın aynı sayı olmasıdır: $E(X) = \text{Var}(X) = \lambda$. Binom dağılımında ise $\text{Var}(X) = np(1-p) = E(X)(1-p) < E(X)$ 'tir (**Teorem 20.2**); $p \rightarrow 0$ iken bu iki büyüklüğün oranı $1-p \rightarrow 1$ 'e gider. Bu, binom dağılımının küçük p için Poisson dağılımına yaklaşmasıyla tutarlıdır.

Aynı gözlem uygulamada bir sağlama aracıdır: gözlenen sayımların örnek ortalaması ile örnek varyansı birbirine yakınsa Poisson modeli makuldür; varyans ortalamayı belirgin biçimde aşıyorsa Poisson modeli yetersizdir.

Standart sapmanın $\sqrt{\lambda}$ olması, “yayılmanın ortalamaya göre daralması” demektir: $\lambda = 100$ için standart sapma yalnızca 10’dur, yani gözlemler ortalamanın yüzde 10’u kadar bir bantta yoğunlaşır. Bu, önceki şekilde λ büyüdükçe çubukların mutlak anlamda yayılmasına karşın ortalamaya oranla toplanmasını açıklar.

21.3 Moment Üreten ve Karakteristik Fonksiyon

Bundan sonraki sonuçlardan biri — bağımsız Poisson değişkenlerinin toplamı — en kısa yoldan üretici fonksiyonlarla ispatlanır; binom dağılımına yaklaşım için de üretici fonksiyonlarla ikinci bir okuma vereceğiz. Bu yüzden Poisson dağılımının moment üreten fonksiyonunu (**Tanım 17.2**), karakteristik fonksiyonunu (**Tanım 17.1**) ve çarpımsal moment üreten fonksiyonunu (**Tanım 17.3**) hesaplıyoruz. Üçü de aynı üstel seri hesabına dayanır.

Teorem 21.2 (Poisson Dağılımının Üretici Fonksiyonları). $X \sim \text{Poisson}(\lambda)$ olsun.

(i) Moment üreten fonksiyon her $t \in \mathbb{R}$ için tanımlıdır ve

$$M_X(t) = e^{\lambda(e^t-1)}, \quad t \in \mathbb{R}.$$

(ii) Karakteristik fonksiyon

$$\varphi_X(t) = e^{\lambda(e^{it}-1)}, \quad t \in \mathbb{R}.$$

(iii) Çarpımsal moment üreten fonksiyon

$$N_X(t) = E(t^X) = e^{\lambda(t-1)}, \quad t \in \mathbb{R}.$$

İspat.

Bu hesapları **Örnek 17.5** ve **Örnek 17.7**'de yapmıştık; burada üçünü toplu olarak kayda geçiriyoruz.

(i) **Tanım 17.2** gereği $M_X(t) = E(e^{tX})$ 'dir. Kesikli değişkende bu beklenen değer bir seridir:

$$M_X(t) = \sum_{x=0}^{\infty} e^{tx} \cdot \frac{e^{-\lambda} \lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!}.$$

Burada $e^{tx} \lambda^x = (\lambda e^t)^x$ yazdık. Son toplam, $u = \lambda e^t$ için üstel seridir ve her $t \in \mathbb{R}$ için yakınsar. Dolayısıyla

$$M_X(t) = e^{-\lambda} e^{\lambda e^t} = e^{\lambda e^t - \lambda} = e^{\lambda(e^t - 1)}.$$

Seri her t için yakınsadığından M_X bütün $\mathbb{R}$ 'de, özellikle sıfırın bir komşuluğunda tanımlıdır; bu yüzden **Teorem 17.4** ve **Teorem 17.5** Poisson dağılımına uygulanabilir.

(ii) **Tanım 17.1** gereği $\varphi_X(t) = E(e^{itX})$ 'dir. Aynı hesap e^t yerine e^{it} ile yapılır:

$$\varphi_X(t) = \sum_{x=0}^{\infty} e^{itx} \cdot \frac{e^{-\lambda} \lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^{it})^x}{x!} = e^{-\lambda} e^{\lambda e^{it}} = e^{\lambda(e^{it} - 1)}.$$

Üstel seri karmaşık sayılar için de geçerlidir ve $|\lambda e^{it}|^x / x! = \lambda^x / x!$ terimlerinin toplamı sonlu olduğundan seri mutlak yakınsaktır.

(iii) **Tanım 17.3** gereği $N_X(t) = E(t^X)$ 'dir. Aynı seri hesabıyla

$$N_X(t) = \sum_{x=0}^{\infty} t^x \cdot \frac{e^{-\lambda} \lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda t)^x}{x!} = e^{-\lambda} e^{\lambda t} = e^{\lambda(t-1)}$$

bulunur.

Sağlama. **Teorem 17.4** gereği $E(X) = M'_X(0)$ ve $E(X^2) = M''_X(0)$ 'dır. Zincir kuralıyla

$$M'_X(t) = \lambda e^t e^{\lambda(e^t-1)} = \lambda e^t M_X(t), \quad M'_X(0) = \lambda \cdot 1 \cdot 1 = \lambda.$$

Bir kez daha türev alırsak

$$M''_X(t) = \lambda e^t M_X(t) + \lambda e^t M'_X(t) = \lambda e^t M_X(t) + \lambda^2 e^{2t} M_X(t), \quad M''_X(0) = \lambda + \lambda^2.$$

Böylece $\text{Var}(X) = M''_X(0) - [M'_X(0)]^2 = \lambda + \lambda^2 - \lambda^2 = \lambda$; **Teorem 21.1** ile uyuşur. Çarpımsal momentler için $N_X(t) = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda t)^x}{x!}$ bütün $\mathbb{R}$ 'de yakınsak bir kuvvet serisi olduğundan terim terim türetilir; yani **Teorem 17.6**'ın türev ile beklenen değerin yer değiştirme koşulu sağlanır. Zincir kuralıyla $N_X^{(k)}(t) = \lambda^k e^{\lambda(t-1)}$ ve $N_X^{(k)}(1) = \lambda^k$; **Teorem 17.6** gereği bu sayı k . çarpımsal momenttir ve **Teorem 21.1**'nin (ii) şikkıyla uyuşur. ■

Tanım 17.3'in ardından belirtildiği gibi $M_X(t) = N_X(e^t)$ 'dir; X 'in değerleri negatif olmayan tam sayılar olduğundan aynı kuvvet serisi karmaşık argüman için de mutlak yakınsar ve $\varphi_X(t) = N_X(e^{it})$ yazılabilir. Bu, yalnızca Poisson dağılımına özgü değildir; t^X 'in $t = e^s$ ve $t = e^{is}$ değerlerinde e^{sX} ve e^{isX} olmasından gelir.

21.4 Bağımsız Poisson Değişkenlerinin Toplamı

Bir dakikada gelen çağrı sayısı Poisson ise iki dakikada gelen çağrı sayısı ne olur? Birinci ve ikinci dakikadaki sayılar bağımsızsa yanıt, iki bağımsız Poisson değişkeninin toplamıdır. Aşağıdaki teorem bu toplamın yine Poisson olduğunu, parametrelerin ise toplandığını söyler. Bu özellik, Poisson dağılımının zaman ve bölge modeli olarak tutarlı kullanılabilmesinin temelidir.

Teorem 21.3 (Bağımsız Poisson Değişkenlerinin Toplamı). $X \sim \text{Poisson}(\lambda_1)$ ve $Y \sim \text{Poisson}(\lambda_2)$ bağımsız rastgele değişkenler olsun. O zaman

$$X + Y \sim \text{Poisson}(\lambda_1 + \lambda_2).$$

Daha genel olarak, $X_1, \dots, X_n$ bağımsız ve her i için $X_i \sim \text{Poisson}(\lambda_i)$ ise

$$X_1 + \dots + X_n \sim \text{Poisson}(\lambda_1 + \dots + \lambda_n).$$

İspat.

Birinci yol: moment üreten fonksiyon. X ile Y bağımsız olduğundan (**Tanım 14.1**) **Teorem 18.1**'i $g_1(u) = g_2(u) = e^{tu}$ ile uygulayabiliriz; her $t \in \mathbb{R}$ için

$$M_{X+Y}(t) = E(e^{t(X+Y)}) = E(e^{tX} e^{tY}) = E(e^{tX}) E(e^{tY}) = M_X(t) M_Y(t).$$

Teorem 21.2 ile

$$M_{X+Y}(t) = e^{\lambda_1(e^t-1)} \cdot e^{\lambda_2(e^t-1)} = e^{(\lambda_1+\lambda_2)(e^t-1)}, \quad t \in \mathbb{R}.$$

Sağ taraf, Poisson $(\lambda_1 + \lambda_2)$ dağılımının moment üreten fonksiyonudur ve iki fonksiyon bütün $\mathbb{R}$ 'de çakışır. **Teorem 17.5** gereği $X + Y$ 'nin dağılımı Poisson $(\lambda_1 + \lambda_2)$ 'dir.

İkinci yol: konvolüsyon. Aynı sonuca olasılık fonksiyonlarını doğrudan toplayarak da ulaşılır. $X + Y$ 'nin değer kümesi $\{0, 1, 2, \dots\}$ 'dir ve **Önerme 15.3** gereği $z = 0, 1, 2, \dots$ için

$$f_{X+Y}(z) = \sum_{k=0}^z f_X(k) f_Y(z-k) = \sum_{k=0}^z \frac{e^{-\lambda_1} \lambda_1^k}{k!} \cdot \frac{e^{-\lambda_2} \lambda_2^{z-k}}{(z-k)!}.$$

$e^{-(\lambda_1+\lambda_2)}$ çarpanını dışarı alıp toplamı $\frac{z!}{z!}$ ile genişletelim:

$$f_{X+Y}(z) = \frac{e^{-(\lambda_1+\lambda_2)}}{z!} \sum_{k=0}^z \frac{z!}{k!(z-k)!} \lambda_1^k \lambda_2^{z-k} = \frac{e^{-(\lambda_1+\lambda_2)}}{z!} \sum_{k=0}^z \binom{z}{k} \lambda_1^k \lambda_2^{z-k}.$$

Son toplam, [Teorem 5.4](#) gereği $(\lambda_1 + \lambda_2)^z$ 'dir. Böylece

$$f_{X+Y}(z) = \frac{e^{-(\lambda_1+\lambda_2)} (\lambda_1 + \lambda_2)^z}{z!}, \quad z = 0, 1, 2, \dots$$

olur; bu, Poisson $(\lambda_1 + \lambda_2)$ dağılımının olasılık fonksiyonudur.

Genel durum. n üzerinden tümevarım yapalım. $n = 1$ için söylenecek bir şey yoktur. $S_{n-1} = X_1 + \dots + X_{n-1}$ toplamının Poisson $(\lambda_1 + \dots + \lambda_{n-1})$ dağılımına sahip olduğunu varsayalım. S_{n-1} ile X_n 'in bağımsız olduğunu görelim: bağımsızlık gereği ortak olasılık fonksiyonu marjinallerin çarpımına ayrıldığından, her s ve x için

$$P(S_{n-1} = s, X_n = x) = \sum_{x_1 + \dots + x_{n-1} = s} P(X_1 = x_1) \dots P(X_{n-1} = x_{n-1}) P(X_n = x) = P(S_{n-1} = s) P(X_n = x)$$

olur. İki değişken için ispatlanan sonuç $S_n = S_{n-1} + X_n$ 'ye uygulanırsa $S_n \sim \text{Poisson}(\lambda_1 + \dots + \lambda_n)$ bulunur.

■

⚠ Fark ve katlar Poisson değildir

Teorem yalnızca **bağımsız** değişkenlerin **toplamı** için geçerlidir. $X - Y$ farkı negatif değerler alabildiğinden Poisson olamaz. $2X$ gibi bir kat da Poisson değildir: yalnızca çift değerler alır ve $\text{Var}(2X) = 4\lambda \neq 2\lambda = E(2X)$ 'dir ([Teorem 16.2](#)); oysa Poisson dağılımında ortalama ile varyans eşittir. Bağımsızlık koşulu da atlanamaz: $Y = X$ alınırsa $X + Y = 2X$ elde edilir.

Toplam teoremini, zaman ölçeğinin değiştiği bir çağrı merkezi örneğinde kullanalım.

Örnek 21.2 (Çağrı Merkezi). Bir çağrı merkezine gelen çağrılar için şu model kabul edilsin: uzunluğu t dakika olan herhangi bir zaman aralığında gelen çağrı sayısı Poisson $(3t)$ dağılımına sahiptir ve ayrıık zaman aralıklarında gelen çağrı sayıları bağımsızdır. Yani dakikada ortalama 3 çağrı gelmektedir.

- Belirli bir dakikada hiç çağrı gelmeme olasılığını bulunuz.
- Belirli bir dakikada en az 5 çağrı gelme olasılığını bulunuz.
- Belirli bir 2 dakikalık sürede en çok 3 çağrı gelme olasılığını bulunuz ve modelin [Teorem 21.3](#) ile tutarlı olduğunu gösteriniz.
- Belirli bir 20 saniyelik sürede en az bir çağrı gelme olasılığını bulunuz.
- Merkeze, birincisinden bağımsız ve dakikada ortalama 2 çağrı taşıyan ikinci bir hat eklenmiştir. Bir dakikada iki hatta toplam tam 5 çağrı gelme olasılığını bulunuz.

Çözüm.

Bir dakikadaki çağrı sayısına X diyelim: $X \sim \text{Poisson}(3)$.

(a) $P(X = 0) = e^{-3} \approx 0,0498$. Yaklaşık her 20 dakikadan biri çağrısız geçer.

(b) Tümüleyeninden hesaplayalım:

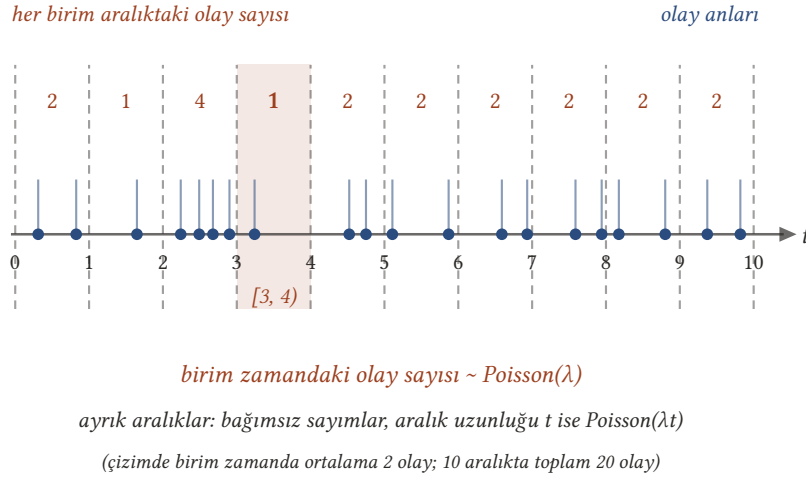
$$P(X \geq 5) = 1 - \sum_{x=0}^4 \frac{e^{-3} 3^x}{x!} = 1 - e^{-3} \left(1 + 3 + \frac{9}{2} + \frac{27}{6} + \frac{81}{24} \right) = 1 - e^{-3} \left(1 + 3 + \frac{9}{2} + \frac{9}{2} + \frac{27}{8} \right) = 1 - \frac{131}{8} e^{-3}.$$

Sayısal olarak $\frac{131}{8} e^{-3} \approx 0,8153$, dolayısıyla $P(X \geq 5) \approx 0,1847$.

(c) Modele göre 2 dakikadaki çağrı sayısı $Y \sim \text{Poisson}(6)$ 'dır. O hâlde

$$P(Y \leq 3) = e^{-6} \left(1 + 6 + \frac{36}{2} + \frac{216}{6} \right) = e^{-6} (1 + 6 + 18 + 36) = 61e^{-6} \approx 0,1512.$$

Tutarlılık: 2 dakikalık süre, ayrık iki 1 dakikalık aralığın birleşimidir. Bu aralıklardaki çağrı sayıları X_1 ve X_2 bağımsız ve her biri Poisson (3) olduğundan Teorem 21.3 gereği $X_1 + X_2 \sim \text{Poisson}(3 + 3) = \text{Poisson}(6)$ 'dır. Bu, modelin doğrudan verdiği dağılımla aynıdır; yani " t dakikada Poisson ($3t$)" kuralı toplam teoremiyle çelişmez; tersine, ardışık aralıklardaki çağrı sayılarının toplanabilirliğiyle uyum içindedir.



Şekil 21.2: Zaman ekseninde bir Poisson süreci: mavi noktalar olay anlarıdır, kesikli çizgiler eksen birim aralıklara böler ve turuncu sayılar her aralıktaki olay sayısını verir. Her birim aralıktaki sayım aynı Poisson(λ) dağılımına sahiptir; ayrık aralıklardaki sayımlar bağımsızdır ve uzunluğu t olan bir aralıktaki sayım Poisson(λt) dağılımlıdır. Çağrı merkezi örneğindeki " t dakikada Poisson($3t$)" modeli tam olarak bu tablodur.

(d) 20 saniye $\frac{1}{3}$ dakikadır; bu süredeki çağrı sayısı $Z \sim \text{Poisson}(3 \cdot \frac{1}{3}) = \text{Poisson}(1)$ 'dir. Buradan

$$P(Z \geq 1) = 1 - P(Z = 0) = 1 - e^{-1} \approx 0,6321.$$

(e) Birinci hattın bir dakikadaki çağrı sayısı $X \sim \text{Poisson}(3)$, ikincininki $W \sim \text{Poisson}(2)$ ve ikisi bağımsızdır. Teorem 21.3 gereği $X + W \sim \text{Poisson}(5)$ ve

$$P(X + W = 5) = \frac{e^{-5} 5^5}{5!} = \frac{3125}{120} e^{-5} = \frac{625}{24} e^{-5} \approx 0,1755.$$

$\lambda = 5$ tam sayı olduğundan Önerme 21.2 gereği 4 ve 5 birlikte moddur: $P(X + W = 4) = P(X + W = 5)$; gerçekten $\frac{e^{-5} 5^4}{4!} = \frac{625}{24} e^{-5}$ 'tir.

■

21.5 Binom Dağılımına Poisson Yaklaşımı

Bölümün başındaki vaadi yerine getirme zamanı: n büyük, p küçük ve np ılımlı iken binom olasılıklarının Poisson olasılıklarına yaklaştığını ispatlayacağız. İspatın kalbi, $(1 - \frac{\lambda}{n})^n \rightarrow e^{-\lambda}$ limitidir; ama np 'nin tam olarak λ 'ya eşit olmasını değil, yalnızca λ 'ya yakınsamasını isteyeceğimizden limitin biraz daha genel bir biçimine ihtiyacımız var. Önce onu ispatlayalım.

Lemma 21.1 (Üstel Limit). (a_n) gerçel sayı dizisi bir $a \in \mathbb{R}$ sayısına yakınsasın. O zaman

$$\lim_{n \rightarrow \infty} \left(1 + \frac{a_n}{n} \right)^n = e^a.$$

Özel olarak, her $\lambda \in \mathbb{R}$ için $(1 - \frac{\lambda}{n})^n \rightarrow e^{-\lambda}$ 'dir.

İspat.

Adım 1: bir logaritma eşitsizliği. Her $u > -1$ için

$$\frac{u}{1+u} \leq \ln(1+u) \leq u$$

olduğunu gösterelim. $g(u) = u - \ln(1+u)$ fonksiyonu $(-1, \infty)$ üzerinde türevlenebilirdir, $g(0) = 0$ ve

$$g'(u) = 1 - \frac{1}{1+u} = \frac{u}{1+u}$$

olur. g' , $(-1, 0)$ üzerinde negatif ve $(0, \infty)$ üzerinde pozitifdir; dolayısıyla g önce azalır, sonra artar ve en küçük değerini $u = 0$ 'da alır: her $u > -1$ için $g(u) \geq g(0) = 0$, yani $\ln(1+u) \leq u$. Sağ eşitsizlik budur. Sol eşitsizlik için bu sonucu $v = -\frac{u}{1+u}$ sayısına uygulayalım. $1+v = \frac{1}{1+u} > 0$ olduğundan $v > -1$ 'dir ve

$$\ln(1+v) = \ln \frac{1}{1+u} = -\ln(1+u) \leq v = -\frac{u}{1+u}$$

bulunur; her iki tarafı -1 ile çarpınca $\ln(1+u) \geq \frac{u}{1+u}$ çıkar.

Adım 2: sıkıştırma. Yakınsak dizi sınırlıdır: her n için $|a_n| \leq K$ olacak biçimde bir $K > 0$ vardır. $n > K$ için $u_n = \frac{a_n}{n}$ sayısı $|u_n| < 1$ 'i sağlar; özellikle $u_n > -1$ ve $1+u_n > 0$ 'dır. Adım 1'deki eşitsizliği $u = u_n$ için yazıp n ile çarpalım:

$$\frac{nu_n}{1+u_n} \leq n \ln(1+u_n) \leq nu_n.$$

Sağdaki dizi $nu_n = a_n \rightarrow a$ 'dır. Soldaki dizi için $|a_n/n| \leq K/n \rightarrow 0$ olduğundan

$$\frac{nu_n}{1+u_n} = \frac{a_n}{1+a_n/n} \rightarrow \frac{a}{1+0} = a.$$

Sıkıştırma teoremi gereği $n \ln(1+u_n) \rightarrow a$.

Adım 3: sonuç. $1+u_n > 0$ olduğundan $(1+u_n)^n = \exp(n \ln(1+u_n))$ yazılabilir. Üstel fonksiyon sürekli olduğundan

$$\left(1 + \frac{a_n}{n}\right)^n = \exp(n \ln(1+u_n)) \rightarrow \exp(a) = e^a.$$

Özel durum, sabit $a_n = -\lambda$ dizisiyle elde edilir.

■

Şimdi ana teoreme geçebiliriz. İfadeyi, p 'nin n 'ye bağlı olarak değiştiği bir dizi olarak kuruyoruz; “ n büyük, p küçük” sezgisinin kesin karşılığı budur.

Teorem 21.4 (Binom Dağılımına Poisson Yaklaşımı). $\lambda > 0$ olsun. Her n için $p_n \in (0, 1)$ verilmiş ve $n \rightarrow \infty$ iken $np_n \rightarrow \lambda$ olsun; en basit örnek, $n > \lambda$ için $p_n = \frac{\lambda}{n}$ alınmasıdır. $X_n \sim B(n, p_n)$ ise her sabit $x \in \{0, 1, 2, \dots\}$ için

$$\lim_{n \rightarrow \infty} P(X_n = x) = \lim_{n \rightarrow \infty} \binom{n}{x} p_n^x (1-p_n)^{n-x} = \frac{e^{-\lambda} \lambda^x}{x!}.$$

Yani n büyük ve p küçük iken $B(n, p)$ olasılıkları, $\lambda = np$ alınarak Poisson (np) olasılıklarıyla yaklaşık olarak hesaplanabilir.

İspat.

$\lambda_n = np_n$ diyelim; varsayım gereği $\lambda_n \rightarrow \lambda$ ve $p_n = \frac{\lambda_n}{n}$ 'dir. x 'i sabitleyelim ve $n \geq x$ alalım. Tanım 20.3 ve Teorem 5.2 ile

$$\binom{n}{x} = \frac{n!}{x!(n-x)!} = \frac{n(n-1)\cdots(n-x+1)}{x!}$$

olduğundan

$$P(X_n = x) = \frac{n(n-1)\cdots(n-x+1)}{x!} \left(\frac{\lambda_n}{n}\right)^x \left(1 - \frac{\lambda_n}{n}\right)^{n-x}.$$

$\left(1 - \frac{\lambda_n}{n}\right)^{n-x} = \left(1 - \frac{\lambda_n}{n}\right)^n \left(1 - \frac{\lambda_n}{n}\right)^{-x}$ yazıp $(1 - \lambda_n/n = 1 - p_n > 0)$ olduğundan negatif kuvvet tanımlıdır) çarpanları yeniden gruplayalım:

$$P(X_n = x) = \underbrace{\frac{\lambda_n^x}{x!}}_{A_n} \cdot \underbrace{\frac{n(n-1)\cdots(n-x+1)}{n^x}}_{B_n} \cdot \underbrace{\left(1 - \frac{\lambda_n}{n}\right)^n}_{C_n} \cdot \underbrace{\left(1 - \frac{\lambda_n}{n}\right)^{-x}}_{D_n}.$$

Dört çarpanın limitini ayrı ayrı bulalım.

A_n çarpanı. $u \mapsto u^x$ sürekli olduğundan $\lambda_n \rightarrow \lambda$ iken $\lambda_n^x \rightarrow \lambda^x$; dolayısıyla $A_n \rightarrow \frac{\lambda^x}{x!}$.

B_n çarpanı. Pay x tane çarpandan oluşur; her birini n 'ye bölelim:

$$B_n = \frac{n}{n} \cdot \frac{n-1}{n} \cdots \frac{n-x+1}{n} = \prod_{k=0}^{x-1} \left(1 - \frac{k}{n}\right).$$

$k = 0, 1, \dots, x-1$ sabit sayılar olduğundan her çarpan 1'e yakınsar; sonlu sayıda (x tane) yakınsak dizinin çarpımının limiti limitlerin çarpımıdır. Dolayısıyla $B_n \rightarrow 1$. ($x = 0$ için çarpım boştur ve $B_n = 1$ 'dir.)

C_n çarpanı. $C_n = \left(1 + \frac{-\lambda_n}{n}\right)^n$ ve $-\lambda_n \rightarrow -\lambda$ olduğundan [Lemma 21.1](#) ($a_n = -\lambda_n$, $a = -\lambda$) gereği $C_n \rightarrow e^{-\lambda}$.

D_n çarpanı. (λ_n) yakınsak olduğundan sınırlıdır; dolayısıyla $\frac{\lambda_n}{n} \rightarrow 0$. $u \mapsto (1-u)^{-x}$ fonksiyonu $u = 0$ 'da sürekli olduğundan $D_n \rightarrow (1-0)^{-x} = 1$.

Yakınsak dizilerin çarpımı limitlerin çarpımına yakınsadığından

$$\lim_{n \rightarrow \infty} P(X_n = x) = \frac{\lambda^x}{x!} \cdot 1 \cdot e^{-\lambda} \cdot 1 = \frac{e^{-\lambda} \lambda^x}{x!}.$$

■

i Aynı limit moment üreten fonksiyonlarla

Teoremin sonucunu üretici fonksiyonlar dilinde de görebiliriz. [Teorem 20.3](#) gereği $X_n \sim B(n, p_n)$ için $M_{X_n}(t) = (1 - p_n + p_n e^t)^n$ 'dir. $p_n = \lambda_n/n$ yazılırsa

$$M_{X_n}(t) = \left(1 + \frac{\lambda_n(e^t - 1)}{n}\right)^n.$$

Sabit t için $a_n = \lambda_n(e^t - 1) \rightarrow \lambda(e^t - 1)$ olduğundan [Lemma 21.1](#) gereği

$$M_{X_n}(t) \rightarrow e^{\lambda(e^t - 1)},$$

ki bu [Teorem 21.2](#) gereği Poisson (λ) dağılımının moment üreten fonksiyonudur. Moment üreten fonksiyonların sıfırın bir komşuluğundaki her t için yakınsamasının dağılımların yakınsamasını gerektirdiği bilinen bir sonuçtur; burada ona gerek duymadan olasılıkların yakınsadığını doğrudan ispatladık, ama iki hesabın aynı yere çıkması güzel bir sağlamadır.

❓ Yaklaşım ne zaman kullanılır

Teorem bir limit ifadesidir; uygulamada sonlu n ve p için ne kadar iyi olduğunu bilmek gerekir. Yaygın kaba kural, $n \geq 50$ ve $p \leq 0,05$ olduğunda $B(n, p)$ yerine Poisson (np) kullanmanın yeterli olduğudur. Daha kesin bir ifade de vardır: bütün x değerleri üzerinden toplanan mutlak hata

$$\sum_{x=0}^{\infty} \left| P(X_n = x) - \frac{e^{-np}(np)^x}{x!} \right| \leq 2np^2$$

eşitsizliğini sağlar (Le Cam eşitsizliği; ispatı bu kitabın kapsamı dışındadır). $np = \lambda$ sabit tutulursa sağ taraf $2\lambda p$ olur; yani hata, p ile orantılı olarak sıfıra gider. Örneğin $n = 100$, $p = 0,02$ için sınır $0,08$ 'dir; gerçek toplam hata ise yaklaşık $0,009$ 'dur.

Yaklaşımın niteliğini sayılarla görelim. Aşağıdaki örnekte $np = 2$ sabit tutulup n büyütülmektedir; tablo, her satırda binom olasılıklarının Poisson (2) değerine nasıl yaklaştığını gösterir.

Örnek 21.3 (Nadir Olay Yaklaşımının Sayısal Sınanması). Bir üretim hattında her parçanın, ötekilerden bağımsız olarak, $p = 0,02$ olasılıkla kusurlu olduğu bilinmektedir. 100 parçalık bir kutudaki kusurlu parça sayısı X olsun.

(a) X 'in dağılımını yazınız ve $P(X \leq 1)$ 'i tam olarak hesaplayınız.

(b) Aynı olasılığı Poisson yaklaşımıyla hesaplayınız.

(c) $np = 2$ sabit kalacak biçimde $n = 20$, $p = 0,1$; $n = 100$, $p = 0,02$ ve $n = 1000$, $p = 0,002$ için $P(X = x)$ binom olasılıklarını Poisson (2) olasılıklarıyla tablo hâlinde karşılaştırınız.

Çözüm.

(a) Parçalar bağımsız Bernoulli denemeleridir; $X \sim B(100; 0,02)$. Tam hesap:

$$P(X \leq 1) = (0,98)^{100} + 100 \cdot (0,02) \cdot (0,98)^{99} = (0,98)^{100} + 2(0,98)^{99} \approx 0,1326 + 0,2707 = 0,4033.$$

(b) $\lambda = np = 100 \cdot 0,02 = 2$ ile

$$P(X \leq 1) \approx e^{-2}(1 + 2) = 3e^{-2} \approx 0,4060.$$

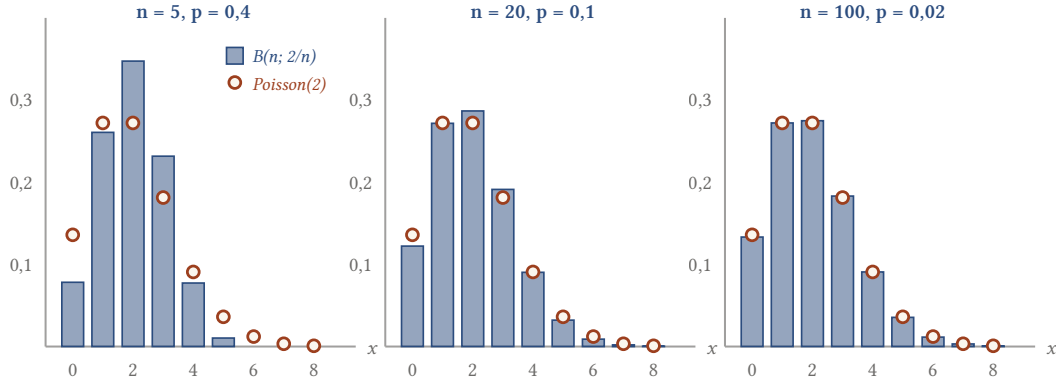
İki değer arasındaki fark $0,003$ 'ten küçüktür; ikinci hesap ise 100 'üncü kuvvet almadan yapılmıştır.

(c) Binom olasılıkları $\binom{n}{x}p^x(1-p)^{n-x}$ formülüyle, Poisson olasılıkları $e^{-2}2^x/x!$ ile hesaplanır. Dört basamağa yuvarlanmış değerler şöyledir:

x	$n = 20, p = 0,1$	$n = 100, p = 0,02$	$n = 1000, p = 0,002$	Poisson (2)
0	0,1216	0,1326	0,1351	0,1353
1	0,2702	0,2707	0,2707	0,2707
2	0,2852	0,2734	0,2709	0,2707
3	0,1901	0,1823	0,1806	0,1804
4	0,0898	0,0902	0,0902	0,0902
5	0,0319	0,0353	0,0360	0,0361
6	0,0089	0,0114	0,0120	0,0120
7	0,0020	0,0031	0,0034	0,0034

Tablodan üç gözlem çıkar. Birincisi, $x = 0$ satırı tam olarak $(1 - \frac{2}{n})^n$ dizisidir ve $e^{-2} \approx 0,1353$ 'e yakınsar; **Lemma 21.1**'in sayısal görüntüsü budur. İkincisi, $n = 20$ için sapma dördüncü basamakta değil ikinci basamaktadır ($x = 2$ 'de $0,2852$ yerine $0,2707$); $n = 1000$ için ise her satırdaki fark $0,0003$ 'ü aşmaz. Üçüncüsü, Poisson (2) için $\lambda = 2$ tam sayı olduğundan **Önerme 21.2** gereği $x = 1$ ve $x = 2$ eşit olasılıklı iki moddur; tabloda $0,2707 = 0,2707$ bunu doğrular. Binom sütunlarında ise $x = 2$ tek moddur: $(\frac{2}{n}, \frac{2}{n})$ için $(n+1)p = 2 + \frac{2}{n}$, $n \geq 3$ iken tam sayı değildir ve **Önerme 20.3** gereği tek mod

$\lfloor 2 + \frac{2}{n} \rfloor = 2$ 'dir. n büyüdükçe $(n+1)p$ değeri 2'ye yaklaşır; limitte $x = 1$ ile $x = 2$ 'nin eşit olasılıklı olması bu yaklaşmanın sonucudur.



Şekil 21.3: $np = \lambda = 2$ sabit tutulup n büyütüldüğünde $B(n, p)$ çubukları Poisson(2) olasılıklarına (içi boş noktalar) yaklaşır. $n = 5$ 'te $x = 2$ çubuğu Poisson değerinin epey üstünde, $x = 0$ çubuğu ise altındadır; $n = 100$ 'de fark artık gözle seçilemez. Limitte $x = 1$ ile $x = 2$ eşit olasılıklı olur.

🔍 Poisson dağılımı nerede ortaya çıkar

Poisson yaklaşımı, dağılımın neden bu kadar yaygın olduğunu açıklar. Bir zaman aralığında ya da bir bölgede gözlenen olay sayısını düşünelim ve aralığı çok sayıda küçük parçaya bölelim. Eğer (1) ayrı parçalardaki olay sayıları bağımsızsa, (2) küçük bir parçada bir olay olma olasılığı parçanın uzunluğuyla orantılıysa ve (3) küçük bir parçada iki ya da daha çok olay olma olasılığı ihmal edilebilir düzeydeyse, toplam olay sayısı yaklaşık olarak $B(n, p)$ 'dir; n büyük, p küçük ve $np = \lambda$ sabit olduğundan Teorem 21.4 gereği bu sayı Poisson (λ)'ya yakındır. Bu üç varsayımı sağlayan olay akışına **Poisson süreci** denir; Örnek 21.2'ndeki " t dakikada Poisson ($3t$)" modeli tam olarak budur. Klasik örnekler: bir radyoaktif kaynağın belirli sürede yaydığı parçacık sayısı, bir kavşaktan belirli sürede geçen araç sayısı, bir dokuma kumaşının metrekaresindeki kusur sayısı, bir kekin dilimindeki üzüm sayısı ve bu bölümde gördüğümüz baskı hataları ile çağrılar. Ortak nokta her zaman aynıdır: çok sayıda fırsat, her birinde küçük olasılık, ılımlı bir ortalama.

21.6 Alıştırmalar

Alıştırma 21.1 (Poisson Dağılımı Üzerine Alıştırmalar). (a) $X \sim \text{Poisson}(3)$ için $P(X = 2)$ ve $P(X \geq 2)$ olasılıklarını hesaplayınız.

(b) $X \sim \text{Poisson}(\lambda)$ ve $P(X = 1) = P(X = 2)$ ise λ 'yı, $P(X = 0)$ 'ı ve $P(X = 3)$ 'ü bulunuz.

(c) $X \sim \text{Poisson}(\lambda)$ için $E(X^3) = \lambda^3 + 3\lambda^2 + \lambda$ olduğunu gösteriniz.

(d) $X \sim \text{Poisson}(\lambda)$ için X 'in çift değer alma olasılığının $\frac{1+e^{-2\lambda}}{2}$ olduğunu gösteriniz ve $\lambda = 2$ için değerini bulunuz.

(e) $X \sim \text{Poisson}(\lambda_1)$ ile $Y \sim \text{Poisson}(\lambda_2)$ bağımsız olsun. $n \geq 0$ tam sayısı ve $k = 0, 1, \dots, n$ için

$$P(X = k \mid X + Y = n) = \binom{n}{k} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^k \left(\frac{\lambda_2}{\lambda_1 + \lambda_2} \right)^{n-k}$$

olduğunu gösteriniz; yani $X + Y = n$ verildiğinde X 'in koşullu dağılımı $\left(B, \frac{\lambda_1}{\lambda_1 + \lambda_2} \right)$ 'dir.

(f) Bir aşının her uygulamada, ötekilerden bağımsız olarak, 0,005 olasılıkla hafif bir yan etkiye yol açtığı bilinmektedir. 300 uygulamada en az iki yan etki görülme olasılığını tam olarak ve Poisson yaklaşımıyla hesaplayınız; aradaki farkı Le Cam sınırıyla karşılaştırınız.

(g) $X \sim \text{Poisson}(100)$ olsun. Chebyshev eşitsizliğiyle $P(80 < X < 120)$ için bir alt sınır bulunuz.

Çözüm.

(a) $\lambda = 3$ için

$$P(X = 2) = \frac{e^{-3}3^2}{2!} = \frac{9}{2}e^{-3} \approx 0,2240, \quad P(X \geq 2) = 1 - e^{-3}(1 + 3) = 1 - 4e^{-3} \approx 0,8009.$$

(b) $P(X = 1) = \lambda e^{-\lambda}$ ve $P(X = 2) = \frac{\lambda^2 e^{-\lambda}}{2!}$ dir. Eşitlikten, $\lambda > 0$ ve $e^{-\lambda} > 0$ olduğu için $\lambda e^{-\lambda}$ ile sadeleştirerek $1 = \frac{\lambda}{2}$, yani $\lambda = 2$ bulunur. (Önerme 21.2'in (ii) şikkı da aynı şeyi söyler: $f(2) = f(1)$ ancak $1 + 1 = \lambda$ iken olur.) Buradan

$$P(X = 0) = e^{-2} \approx 0,1353, \quad P(X = 3) = \frac{e^{-2}2^3}{3!} = \frac{4}{3}e^{-2} \approx 0,1804.$$

(c) x^3 'ü çarpımsal momentlere ayıralım. $x(x-1)(x-2) = x^3 - 3x^2 + 2x$ ve $3x(x-1) = 3x^2 - 3x$ olduğundan

$$x(x-1)(x-2) + 3x(x-1) + x = x^3 - 3x^2 + 2x + 3x^2 - 3x + x = x^3.$$

Beklenen değerin doğrusallığı ve Teorem 21.1'nin (ii) şikkı ($k = 3, 2, 1$ için çarpımsal momentler $\lambda^3, \lambda^2, \lambda$) ile

$$E(X^3) = E[X(X-1)(X-2)] + 3E[X(X-1)] + E(X) = \lambda^3 + 3\lambda^2 + \lambda.$$

Sağlama: $\lambda = 2$ için formül $8 + 12 + 2 = 22$ verir; $\sum_x x^3 e^{-2} 2^x / x!$ serisinin sayısal değeri de 22,000'dir.

(d) x çift iken $1 + (-1)^x = 2$, tek iken 0 olduğundan çift x üzerinden toplam, bütün x üzerinden $\frac{1+(-1)^x}{2}$ ağırlıklı toplamdır:

$$P(X \text{ çift}) = \sum_{x=0}^{\infty} \frac{1 + (-1)^x}{2} \cdot \frac{e^{-\lambda} \lambda^x}{x!} = \frac{e^{-\lambda}}{2} \left(\sum_{x=0}^{\infty} \frac{\lambda^x}{x!} + \sum_{x=0}^{\infty} \frac{(-\lambda)^x}{x!} \right) = \frac{e^{-\lambda}}{2} (e^{\lambda} + e^{-\lambda}) = \frac{1 + e^{-2\lambda}}{2}.$$

Her iki seri de mutlak yakınsak olduğundan toplamı terim terim ayırmak geçerlidir. Sonuç her λ için $\frac{1}{2}$ 'den büyüktür: çift değerler her zaman biraz daha olasıdır. $\lambda = 2$ için $\frac{1+e^{-4}}{2} \approx 0,5092$.

(e) $(X = k) \cap (X + Y = n) = (X = k) \cap (Y = n - k)$ olduğundan bağımsızlıkla

$$P(X = k, X + Y = n) = P(X = k)P(Y = n - k) = \frac{e^{-\lambda_1} \lambda_1^k}{k!} \cdot \frac{e^{-\lambda_2} \lambda_2^{n-k}}{(n-k)!}.$$

Teorem 21.3 gereği $X + Y \sim \text{Poisson}(\lambda_1 + \lambda_2)$, yani $P(X + Y = n) = \frac{e^{-(\lambda_1 + \lambda_2)} (\lambda_1 + \lambda_2)^n}{n!} > 0$ 'dır. Koşullu olasılık tanımıyla

$$P(X = k | X + Y = n) = \frac{e^{-(\lambda_1 + \lambda_2)} \lambda_1^k \lambda_2^{n-k} / (k! (n-k)!)}{e^{-(\lambda_1 + \lambda_2)} (\lambda_1 + \lambda_2)^n / n!} = \frac{n!}{k! (n-k)!} \cdot \frac{\lambda_1^k \lambda_2^{n-k}}{(\lambda_1 + \lambda_2)^n}.$$

$(\lambda_1 + \lambda_2)^n = (\lambda_1 + \lambda_2)^k (\lambda_1 + \lambda_2)^{n-k}$ yazılırsa istenen biçim elde edilir:

$$P(X = k | X + Y = n) = \binom{n}{k} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^k \left(\frac{\lambda_2}{\lambda_1 + \lambda_2} \right)^{n-k}, \quad k = 0, 1, \dots, n.$$

Bu, $\left(\frac{\lambda_1}{\lambda_1 + \lambda_2} \right)$ dağılımının olasılık fonksiyonudur. Yorum: iki bağımsız Poisson kaynağından toplam n olay gözlemlendiğinde, bu olayların her biri sanki birbirinden bağımsız olarak $\frac{\lambda_1}{\lambda_1 + \lambda_2}$ olasılıkla birinci kaynağa aittir.

(f) Yan etki sayısı $X \sim B(300; 0,005)$ 'tir. Tümleyen olaydan gidelim; tam hesap:

$$P(X \geq 2) = 1 - (0,995)^{300} - 300 \cdot (0,005) \cdot (0,995)^{299} = 1 - (0,995)^{300} - \frac{3}{2} (0,995)^{299} \approx 1 - 0,22229 - 0,33511 = 0,44260.$$

Poisson yaklaşımı $\lambda = 300 \cdot 0,005 = \frac{3}{2}$ ile

$$P(X \geq 2) \approx 1 - e^{-3/2} \left(1 + \frac{3}{2}\right) = 1 - \frac{5}{2}e^{-3/2} \approx 1 - 0,55783 = 0,44217.$$

Fark 0,0005'ten küçüktür; $n = 300$, $p = 0,005$ kaba kuralın rahatça içindedir. Le Cam sınırı, bütün x değerleri üzerinden toplanan mutlak hata için $2np^2 = 2 \cdot 300 \cdot (0,005)^2 = 0,015$ verir; tek bir olayda gözlenen bu farkı fazlasıyla aşar.

(g) **Teorem 21.1** gereği $E(X) = 100$, $\text{Var}(X) = 100$ ve $\sigma_X = 10$ 'dur. **Teorem 16.4** ile

$$P(|X - 100| \geq 20) \leq \frac{\text{Var}(X)}{20^2} = \frac{100}{400} = \frac{1}{4}.$$

$80 < X < 120$ olayı $|X - 100| < 20$ olayının kendisidir; dolayısıyla

$$P(80 < X < 120) = 1 - P(|X - 100| \geq 20) \geq 1 - \frac{1}{4} = \frac{3}{4}.$$

Poisson olasılıkları terim terim toplanarak bulunan kesin değer yaklaşık 0,9491'dir; Chebyshev sınırı kaba, ama hiçbir hesap gerektirmeden dağılımın ortalama etrafında toplandığını garanti eder.

■

Bu bölümde başarı sayısını sınırsız bırakan ilk dağılımı, Poisson dağılımını inceledik. Sonraki bölümde deneme sayısını sabit tutmak yerine “ilk başarıya kadar” ya da “ r 'inci başarıya kadar” bekleyen dağılımlara ve iadesiz çekilişin dağılımına geçiyoruz: [Geometrik](#), [Negatif Binom](#) ve [Hipergeometrik Dağılımlar](#).

22. Geometrik, Negatif Binom ve Hipergeometrik Dağılımlar

Poisson Dağılımı bölümünde binom dağılımının, deneme sayısı büyürken ve başarı olasılığı küçülürken Poisson dağılımına yaklaştığını gördük. Bernoulli, binom ve Poisson dağılımlarının ortak yanı, üçünün de **başarı sayısını** saymasıdır: deneme sayısı (ya da gözlem aralığı) baştan sabittir, kaç başarı geleceği rastgeledir. Bu bölümde soruyu tersine çevireceğiz: başarı sayısını sabitleyip **kaç deneme gerektiğini** sayacağız. İlk başarıya kadar yapılan deneme sayısı **geometrik dağılıma**, r -inci başarıya kadar yapılan deneme sayısı **negatif binom dağılımına** uyar.

Bölümün üçüncü dağılımı başka bir yönden gelir. Binom dağılımı denemelerin bağımsız olduğunu varsayar; bir torbadan **iadeli** çekiliş bu varsayımı sağlar. Oysa uygulamada çekilişler çoğu zaman **iadesizdir**: bir partiden denetlenmek üzere ayrılan ürünler, bir desteden dağıtılan kartlar, bir sınıftan seçilen öğrenciler geri konmaz. İadesiz çekilişte başarı sayısının dağılımı **hipergeometrik dağılımdır**. Onun binom dağılımından nasıl ayrıldığını, varyansındaki “sonlu kitle düzeltmesi” çarpanının nereden geldiğini ve kitle büyüdükçe iki dağılımın nasıl birleştiğini göreceğiz.

Bölüme en yalın kesikli dağılım olan kesikli düzgün dağılımla başlayacağız; sonunda ise şimdiye kadar tanıdığımız bütün kesikli dağılımları tek bir tabloda toplayıp aralarındaki ilişkileri bir önermede özetleyeceğiz. Ana araçlarımız yine moment üreten fonksiyon ([Tanım 17.2](#)), onun momentleri üreten ([Teorem 17.4](#)) ve dağılımı tek türlü belirleyen ([Teorem 17.5](#)) özellikleri ile toplamın varyansı formülü ([Teorem 18.4](#)) olacaktır.

22.1 Kesikli Düzgün Dağılım

Bir Laplace uzayında ([Tanım 4.1](#)) bütün sonuçlar eşit olasılıklıdır. Sonuçları $1, 2, \dots, n$ sayılarıyla etiketlersek “gelen sayı” rastgele değişkeni her değeri aynı olasılıkla alır. Düzgün bir zarın gösterdiği sayı, iyice karıştırılmış n numaralı karttan çekilen birinin numarası, bir piyango çekilişinde çıkan numara hep bu türdendir.

Tanım 22.1 (Kesikli Düzgün Dağılım). $n \in \mathbb{N}$ olsun. Değer kümesi $D_X = \{1, 2, \dots, n\}$ ve olasılık fonksiyonu

$$f_X(x) = \frac{1}{n}, \quad x = 1, 2, \dots, n$$

olan X rastgele değişkenine $\{1, 2, \dots, n\}$ kümesi üzerinde **kesikli düzgün dağılıma** sahiptir denir.

Olasılık fonksiyonunun toplamı $n \cdot \frac{1}{n} = 1$ olduğundan f_X gerçekten bir olasılık fonksiyonudur ([Tanım 10.2](#)). Dağılım fonksiyonu ([Tanım 9.1](#)), $\lfloor x \rfloor$ ile x 'in tam kısmını göstermek üzere,

$$F_X(x) = \begin{cases} 0, & x < 1 \\ \frac{\lfloor x \rfloor}{n}, & 1 \leq x < n \\ 1, & x \geq n \end{cases}$$

biçiminde $\frac{1}{n}$ yüksekliğinde n basamaktan oluşur. Momentleri, tamsayıların ve karelerinin toplam formüllerinden hemen çıkar.

Önerme 22.1 (Kesikli Düzgün Dağılımın Momentleri). X , $\{1, 2, \dots, n\}$ üzerinde kesikli düzgün dağılıma sahip olsun. O zaman

$$E(X) = \frac{n+1}{2}, \quad \text{Var}(X) = \frac{n^2-1}{12}.$$

İspat.

Beklenen değerin tanımı (Tanım 16.1) ve $1 + 2 + \dots + n = \frac{n(n+1)}{2}$ özdeşliğiyle

$$E(X) = \sum_{x=1}^n x \cdot \frac{1}{n} = \frac{1}{n} \cdot \frac{n(n+1)}{2} = \frac{n+1}{2}.$$

İkinci moment için $1^2 + 2^2 + \dots + n^2 = \frac{n(n+1)(2n+1)}{6}$ özdeşliğini kullanalım:

$$E(X^2) = \sum_{x=1}^n x^2 \cdot \frac{1}{n} = \frac{1}{n} \cdot \frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6}.$$

Varyans hesap formülü (Önerme 16.3) gereği

$$\text{Var}(X) = E(X^2) - (E(X))^2 = \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} = \frac{(n+1)}{12} [2(2n+1) - 3(n+1)] = \frac{(n+1)(n-1)}{12} = \frac{n^2-1}{12}.$$

■

Değer kümesi $\{a, a+1, \dots, a+n-1\}$ olan düzgün dağılım için ayrı bir formül gerekmez: böyle bir Y için $X = Y - a + 1$ değişkeni $\{1, \dots, n\}$ üzerinde düzgündür ve Teorem 16.2 gereği $E(Y) = a + \frac{n-1}{2}$, $\text{Var}(Y) = \text{Var}(X) = \frac{n^2-1}{12}$ olur. Örneğin düzgün bir zar için ($n = 6$) $E(X) = \frac{7}{2}$ ve $\text{Var}(X) = \frac{35}{12}$ 'dir.

22.2 Geometrik Dağılım

Bernoulli deneme dizisine geri dönelim: her denemede başarı olasılığı aynı p sayısıdır ($0 < p < 1$) ve denemeler birbirinden bağımsızdır. Bu bölüm boyunca $q = 1 - p$ ile **başarısızlık olasılığı**nı göstereceğiz. Binom dağılımında (Tanım 20.3) deneme sayısı n baştan sabittir. Şimdi deneme sayısını sabitlemiyoruz; denemelere **ilk başarı gelene kadar** devam ediyor ve kaç deneme yapıldığını sayıyoruz. Bir zarı ilk 6 gelene kadar atmak, bir hedefe ilk isabete kadar ateş etmek, bir parçayı ilk kusurlu çıkana kadar denetlemek hep bu modeldir.

X , ilk başarının geldiği denemenin sırası olsun. $X = x$ olayı, ilk $x - 1$ denemenin başarısız ve x -inci denemenin başarılı olması demektir. Denemeler bağımsız olduğundan (Tanım 7.1) bu olayın olasılığı

$$P(X = x) = \underbrace{q \cdot q \cdots q}_{x-1 \text{ kez}} \cdot p = q^{x-1}p$$

olur. Bu, X 'in olasılık fonksiyonudur.

Tanım 22.2 (Geometrik Dağılım). $0 < p < 1$ ve $q = 1 - p$ olsun. Değer kümesi $D_X = \{1, 2, 3, \dots\}$ ve olasılık fonksiyonu

$$f_X(x) = q^{x-1}p, \quad x = 1, 2, 3, \dots$$

olan X rastgele değişkenine **p parametrelili geometrik dağılım**a sahiptir denir ve kısaca $X \sim \text{Geo}(p)$ yazılır. X , başarı olasılığı p olan bağımsız Bernoulli denemelerinde **ilk başarıya kadar yapılan deneme sayısıdır**.

Geometrik dağılımın literatürde iki parametrisasyonu vardır. İkincisi, ilk başarıdan **önceki başarısızlık sayısını** sayar: $Y = X - 1$ için

$$f_Y(y) = q^y p, \quad y = 0, 1, 2, \dots$$

Bu kitapta **deneme sayısı** X standart alınır; “geometrik dağılım” dendiğinde $D_X = \{1, 2, \dots\}$ ve $f_X(x) = q^{x-1}p$ kastedilir. İki parametrisasyon arasında $Y = X - 1$ dönüşümüyle geçilir.

Adını, olasılıkların $p, qp, q^2p, \dots$ biçiminde ortak çarpanı q olan bir **geometrik dizi** oluşturmamasından alır. Bu yapı, dağılımla ilgili her hesabı geometrik seri toplamına indirger. Aşağıdaki teorem, geometrik dağılımla çalışırken gereken her şeyi bir arada verir: olasılıkların toplamının 1 olduğunu, kuyruk olasılığının kapalı biçimini, moment üreten fonksiyonu ve ilk iki momenti.

Teorem 22.1 (Geometrik Dağılımın Kuyruk Olasılığı, Moment Üreten Fonksiyonu ve Momentleri). $X \sim \text{Geo}(p)$ ve $q = 1 - p$ olsun.

- (a) $\sum_{x=1}^{\infty} f_X(x) = 1$ 'dir ve her $k \in \{0, 1, 2, \dots\}$ için $P(X > k) = q^k$ 'dir.
 (b) Moment üreten fonksiyon $t < -\ln q$ için tanımlıdır ve

$$M_X(t) = \frac{pe^t}{1 - qe^t}.$$

- (c) $E(X) = \frac{1}{p}$ ve $\text{Var}(X) = \frac{q}{p^2} = \frac{1-p}{p^2}$.

İspat.

(a) $0 < q < 1$ olduğundan geometrik seri yakınsar:

$$\sum_{x=1}^{\infty} q^{x-1}p = p \sum_{j=0}^{\infty} q^j = p \cdot \frac{1}{1-q} = \frac{p}{p} = 1.$$

Kuyruk olasılığı için

$$P(X > k) = \sum_{x=k+1}^{\infty} q^{x-1}p = pq^k \sum_{j=0}^{\infty} q^j = \frac{pq^k}{1-q} = q^k.$$

Aynı sonuca hesapsız da varılır: $(X > k)$ olayı “ilk k denemenin hepsi başarısız” olayıdır ve bağımsızlık gereği olasılığı q^k 'dir. Buradan dağılım fonksiyonu da çıkar: $x \geq 1$ için $F_X(x) = 1 - q^{[x]}$.

(b) Moment üreten fonksiyonun tanımına (Tanım 17.2) göre

$$M_X(t) = E(e^{tX}) = \sum_{x=1}^{\infty} e^{tx} q^{x-1}p = pe^t \sum_{x=1}^{\infty} (qe^t)^{x-1} = pe^t \sum_{j=0}^{\infty} (qe^t)^j.$$

Bu geometrik seri ancak $qe^t < 1$, yani $t < -\ln q$ iken yakınsar ve toplamı $\frac{1}{1-qe^t}$ 'dir. Öyleyse

$$M_X(t) = \frac{pe^t}{1 - qe^t}, \quad t < -\ln q.$$

$q < 1$ olduğundan $-\ln q > 0$ 'dır; yani M_X , sıfırın bir komşuluğunda tanımlıdır ve Teorem 17.4 uygulanabilir.

(c) $M_X(t) = pe^t (1 - qe^t)^{-1}$ yazıp çarpım ve zincir kurallarıyla türev alalım:

$$M'_X(t) = \frac{pe^t}{1 - qe^t} + \frac{pe^t \cdot qe^t}{(1 - qe^t)^2} = \frac{pe^t [(1 - qe^t) + qe^t]}{(1 - qe^t)^2} = \frac{pe^t}{(1 - qe^t)^2}.$$

Bir kez daha türev alırsak

$$M''_X(t) = \frac{pe^t}{(1 - qe^t)^2} + \frac{2pe^t \cdot qe^t}{(1 - qe^t)^3} = \frac{pe^t [(1 - qe^t) + 2qe^t]}{(1 - qe^t)^3} = \frac{pe^t (1 + qe^t)}{(1 - qe^t)^3}.$$

$t = 0$ koyup $1 - q = p$ kullanalım. Teorem 17.4 gereği

$$E(X) = M'_X(0) = \frac{p}{p^2} = \frac{1}{p}, \quad E(X^2) = M''_X(0) = \frac{p(1+q)}{p^3} = \frac{1+q}{p^2}.$$

Son olarak Önerme 16.3 ile

$$\text{Var}(X) = E(X^2) - (E(X))^2 = \frac{1+q}{p^2} - \frac{1}{p^2} = \frac{q}{p^2}.$$

■

$E(X) = \frac{1}{p}$ sonucu sezgiyle uyuşur: başarı olasılığı $\frac{1}{6}$ ise ilk başarı için ortalama 6 deneme beklenir. Varyans ise p küçüldükçe hızla büyür: $p = \frac{1}{6}$ için $\text{Var}(X) = 30$ ve standart sapma $\sigma_X \approx 5,48$, yani beklenen değere neredeyse eşittir. Geometrik dağılım sağa doğru uzun kuyrukludur.

💡 Beklenen değeri seriyi türevleyerek de bulabilirsiniz

Moment üreten fonksiyona başvurmadan, $|q| < 1$ için kuvvet serisinin terim terim türevlenebilmesini kullanarak

$$E(X) = \sum_{x=1}^{\infty} x q^{x-1} p = p \sum_{x=1}^{\infty} \frac{d}{dq} q^x = p \frac{d}{dq} \left(\frac{q}{1-q} \right) = p \cdot \frac{1}{(1-q)^2} = \frac{1}{p}$$

bulunur. İkinci türevle de ikinci faktöriyel moment çıkar: $x(x-1)q^{x-2} = \frac{d^2}{dq^2} q^x$ olduğundan ($x=1$ terimi sıfırdır)

$$E(X(X-1)) = \sum_{x=1}^{\infty} x(x-1) q^{x-1} p = pq \sum_{x=1}^{\infty} \frac{d^2}{dq^2} q^x = pq \frac{d^2}{dq^2} \left(\frac{q}{1-q} \right) = pq \cdot \frac{2}{(1-q)^3} = \frac{2q}{p^2}$$

ve $E(X^2) = E(X(X-1)) + E(X)$ ile $\text{Var}(X) = \frac{2q}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{2q+p-1}{p^2} = \frac{q}{p^2}$ elde edilir.

Örnek 22.1 (Zarı İlk Altı Gelene Kadar Atmak). Düzgün bir zar, ilk kez 6 gelene kadar atılıyor. X yapılan atış sayısı olsun.

- (a) X 'in dağılımını belirleyiniz ve $P(X=3)$ 'ü hesaplayınız.
- (b) Dört atışta hiç 6 gelmeme olasılığını, yani $P(X > 4)$ 'ü bulunuz.
- (c) İlk 6'nın en geç altıncı atışta gelme olasılığını bulunuz.
- (d) $E(X)$ ve $\text{Var}(X)$ 'i hesaplayınız.
- (e) İlk 6'nın gelmesi için en az yüzde 95 güvenceyle kaç atış yeterlidir?
- (f) $P(X \geq 18)$ olasılığını Chebyshev eşitsizliğiyle yukarıdan sınırlayıp gerçek değeriyle karşılaştırınız.

Çözüm.

(a) Her atışta 6 gelme olasılığı $p = \frac{1}{6}$ ve atışlar bağımsızdır; öyleyse $X \sim G(\frac{1}{6})$, $q = \frac{5}{6}$ ve

$$f_X(x) = \left(\frac{5}{6}\right)^{x-1} \frac{1}{6}, \quad x = 1, 2, 3, \dots$$

Buradan $P(X=3) = \left(\frac{5}{6}\right)^2 \frac{1}{6} = \frac{25}{216} \approx 0,1157$.

(b) Teorem 22.1 (a) gereği $P(X > 4) = q^4 = \left(\frac{5}{6}\right)^4 = \frac{625}{1296} \approx 0,4823$. Dört atışın dördünde de 6 gelmemesi, neredeyse yazı tura kadar olasıdır.

(c) Tümleyen olay üzerinden hesaplayalım:

$$P(X \leq 6) = 1 - P(X > 6) = 1 - \left(\frac{5}{6}\right)^6 = 1 - \frac{15625}{46656} = \frac{31031}{46656} \approx 0,6651.$$

Bu aynı zamanda “altı atışta en az bir 6 gelir” olayının olasılığıdır.

(d) Teorem 22.1 (c) ile

$$E(X) = \frac{1}{p} = 6, \quad \text{Var}(X) = \frac{q}{p^2} = \frac{5/6}{1/36} = 30, \quad \sigma_X = \sqrt{30} \approx 5,48.$$

(e) $P(X \leq n) \geq 0,95$, yani $q^n \leq 0,05$ olacak en küçük n aranıyor. $\ln q < 0$ olduğundan eşitsizlik yön değiştirir:

$$n \geq \frac{\ln 0,05}{\ln(5/6)} \approx \frac{-2,9957}{-0,1823} \approx 16,43.$$

Demek ki $n = 17$ atış yeterlidir. Gerçekten $\left(\frac{5}{6}\right)^{16} \approx 0,0541 > 0,05$ ve $\left(\frac{5}{6}\right)^{17} \approx 0,0451 < 0,05$ 'tir; $P(X \leq 16) \approx 0,9459$, $P(X \leq 17) \approx 0,9549$.

(f) $X \geq 1$ olduğundan $X - 6 \leq -12$ olamaz; dolayısıyla $(X \geq 18) = (|X - 6| \geq 12)$ 'dir. Teorem 16.4 ile

$$P(X \geq 18) = P(|X - E(X)| \geq 12) \leq \frac{\text{Var}(X)}{12^2} = \frac{30}{144} = \frac{5}{24} \approx 0,208.$$

Gerçek değer $P(X \geq 18) = P(X > 17) = \left(\frac{5}{6}\right)^{17} \approx 0,045$ 'tir. Chebyshev sınırı doğrudur ama dağılımın biçimini kullanmadığı için gerçek değer yaklaşık dört buçuk katıdır.

■

22.2.1 Hafızasızlık Özelliği

Zar örneğinde şu soruyu soralım: ilk dört atışta 6 gelmediğini biliyoruz; bundan sonraki üç atışta da gelmeme olasılığı nedir? Sezgi, “zarın hafızası yoktur” der: dört başarısız atış, beşinci atışın 6 gelme şansını değiştirmez. Aşağıdaki teorem bu sezgiyi kesinleştirir ve daha fazlasını söyler: bu özellik, $\{1, 2, \dots\}$ üzerindeki ve $P(X = 1) < 1$ olan kesikli dağılımlar arasında **yalnızca** geometrik dağılıma aittir. Yani hafızasızlık, geometrik dağılımın tanımlayıcı özelliğidir.

Teorem 22.2 (Geometrik Dağılımın Hafızasızlık Özelliği). (a) $X \sim \text{Geo}(p)$ ise her $m, n \in \{0, 1, 2, \dots\}$ için

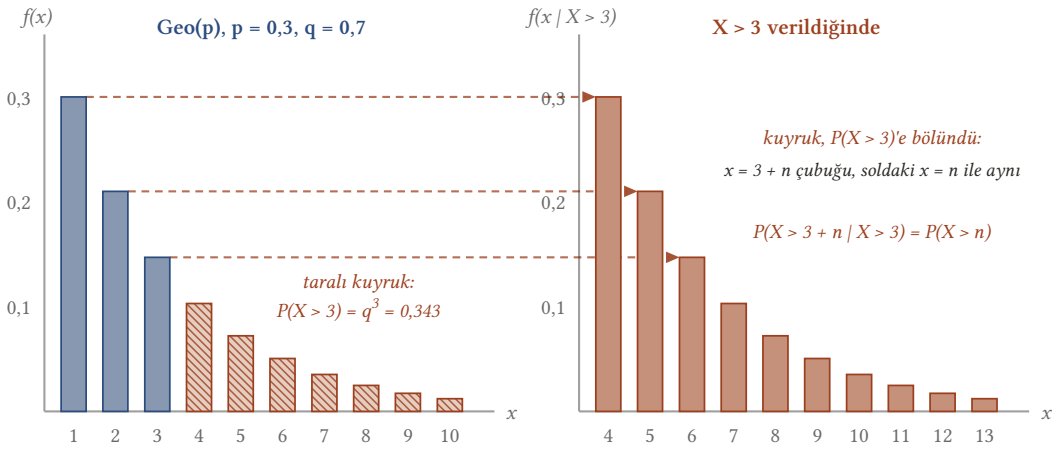
$$P(X > m + n \mid X > m) = P(X > n).$$

(b) Tersine, X değerlerini $\{1, 2, 3, \dots\}$ kümesinde alan ve $P(X = 1) < 1$ koşulunu sağlayan bir kesikli rastgele değişken olsun. Her $m, n \in \{0, 1, 2, \dots\}$ için $P(X > m + n \mid X > m) = P(X > n)$ sağlanıyorsa $X \sim \text{Geo}(p)$ 'dir; burada $p = P(X = 1)$ 'dir.

İspat.

(a) $m + n \geq m$ olduğundan $(X > m + n) \subseteq (X > m)$ 'dir; dolayısıyla $(X > m + n) \cap (X > m) = (X > m + n)$. Teorem 22.1 (a) gereği $P(X > m) = q^m > 0$ olduğundan koşullu olasılık tanımlıdır ve

$$P(X > m + n \mid X > m) = \frac{P(X > m + n)}{P(X > m)} = \frac{q^{m+n}}{q^m} = q^n = P(X > n).$$



Şekil 22.1: Geometrik dağılımın hafızasızlığı ($p = 0,3$). Solda $X > 3$ kuyruğu taralıdır ve toplam olasılığı $q^3 = 0,343$ 'tür. Sağda bu kuyruk $P(X > 3)$ 'e bölünerek koşullu dağılıma dönüştürülmüştür: $x = 4, 5, 6, \dots$ çubukları soldaki $x = 1, 2, 3, \dots$ çubuklarıyla tam aynı yüksekliktedir. Üç başarısızlıktan sonra kalan bekleme süresi, en baştaki bekleme süresiyle aynı dağılıma sahiptir; kesikli oklar bu eşlemeyi gösterir.

(b) $G(k) = P(X > k)$ yazalım; $X \geq 1$ olduğundan $G(0) = 1$ 'dir. Varsayım, $G(m) > 0$ olan her m için

$$\frac{G(m+n)}{G(m)} = G(n), \quad \text{yani} \quad G(m+n) = G(m)G(n)$$

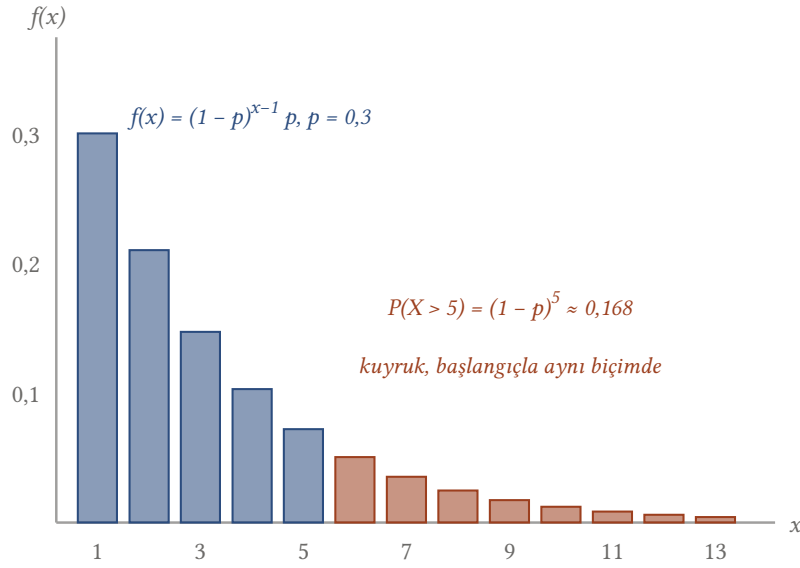
demektir. ($G(m) = 0$ ise $G(m+n) \leq G(m) = 0$ olduğundan bu eşitlik zaten sağlanır; öyleyse eşitlik her m, n için geçerlidir.) $q = G(1) = P(X > 1) = 1 - P(X = 1)$ diyelim. $P(X = 1) < 1$ olduğundan

$q > 0$ 'dır. Tümevarımla $G(n) = q^n$ olduğunu gösterelim: $n = 0$ ve $n = 1$ için doğrudur; $G(n) = q^n$ ise $G(n+1) = G(n)G(1) = q^n \cdot q = q^{n+1}$ 'dir.

Şimdi $q < 1$ olduğunu görelim. $q = 1$ olsaydı her n için $P(X > n) = 1$ olurdu. Ama X gerçel değerli olduğundan $(X > n)$ olayları azalarak $\bigcap_n (X > n) = \emptyset$ kümesine iner; olasılığın azalan dizilerdeki sürekliliği (Teorem 3.5) gereği $P(X > n) \rightarrow P(\emptyset) = 0$ olmalıdır. Bu çelişki $q < 1$ verir. Öyleyse $p = 1 - q = P(X = 1) \in (0, 1)$ 'dir ve her $x \geq 1$ için

$$P(X = x) = P(X > x - 1) - P(X > x) = q^{x-1} - q^x = q^{x-1}(1 - q) = q^{x-1}p,$$

yani $X \sim \text{Geo}(p)$.



Şekil 22.2: İlk başarıya kadar yapılan deneme sayısının (geometrik dağılım, $p = 0,3$) olasılık fonksiyonu. Çubuklar sabit oranla $(1-p)$ küçülür; bu yüzden 5. denemeden sonraki kuyruk (turuncu), baştan başlayan dağılımın küçültülmüş kopyasıdır. Hafızasızlık özelliği tam olarak budur: ilk beş deneme başarısızsa, kalan bekleme yine geometriktir.

■

Teoremin (a) kısmı $m = 4, n = 3$ ile zar sorusunu yanıtlar: $P(X > 7 \mid X > 4) = P(X > 3) = \left(\frac{5}{6}\right)^3 \approx 0,5787$. Dört başarısız atıştan sonra bile ilk 6'ya "yaklaşmış" olmayız; kalan bekleme süresinin dağılımı en baştaki dağılımın aynısıdır. (b) kısmı ise sürekli dağılımlar için de bir habercidir: bir sonraki bölümde, $(0, \infty)$ üzerindeki sürekli dağılımlar arasında hafızasız olan tek dağılımın üstel dağılım olduğunu göreceğiz.

⚠ Kumarbaz yanılgısı

Hafızasızlık, gündelik sezginin sık yanıldığı bir noktadır. Bir rulet çarkında art arda on kez kırmızı geldiğinde "artık siyahın gelme vakti" düşüncesi, geçmiş sonuçların gelecekteki bağımsız denemeyi etkilediğini varsayar; oysa denemeler bağımsızsa on birinci dönüşte siyahın olasılığı ilk dönüştekiyle aynıdır. Geometrik dağılım bu gerçeği tam olarak ifade eder: m başarısızlıktan sonra ilk başarıya kalan bekleme süresi, başlangıçtaki bekleme süresiyle aynı dağılıma sahiptir. "Uzun süredir gelmeyen" bir başarı, gecikmiş bir borç gibi birikirmez.

22.3 Negatif Binom Dağılımı

Geometrik dağılım ilk başarıyı bekler. Şimdi r -inci başarıyı bekleyelim: bir hedefe üçüncü isabet gelene kadar kaç atış yapılır, bir denetimde ikinci kusurlu parçaya kaç parçada rastlanır? $r \in \mathbb{N}$ sabit olmak üzere X , r -inci başarının geldiği denemenin sırası olsun. X 'in alabileceği en küçük değer r 'dir (ilk r denemenin hepsi başarılı olursa).

$X = x$ olayını sayarak olasılık fonksiyonunu türetilim. r -inci başarının tam x -inci denemede gelmesi iki koşulun birlikte sağlanması demektir: x -inci deneme başarılıdır ve ilk $x - 1$ denemede tam $r - 1$ başarı vardır. İlk $x - 1$ denemede $r - 1$ başarının hangi konumlara düşeceği $\binom{x-1}{r-1}$ biçimde seçilir (Teorem 5.2). Bu seçimlerin her biri, sabit bir sonuç dizisine karşılık gelir: dizide toplam r başarı ve $x - r$ başarısızlık vardır ve bağımsızlık gereği dizinin olasılığı $p^r q^{x-r}$ 'dir. Diziler ayırık olaylar olduğundan olasılıklar toplanır:

$$P(X = x) = \binom{x-1}{r-1} p^r q^{x-r}, \quad x = r, r+1, r+2, \dots$$

Tanım 22.3 (Negatif Binom Dağılımı). $r \in \mathbb{N}$, $0 < p < 1$ ve $q = 1 - p$ olsun. Değer kümesi $D_X = \{r, r+1, r+2, \dots\}$ ve olasılık fonksiyonu

$$f_X(x) = \binom{x-1}{r-1} p^r q^{x-r}, \quad x = r, r+1, r+2, \dots$$

olan X rastgele değişkenine **r ve p parametrelili negatif binom dağılımına** sahiptir denir ve $X \sim \text{NB}(r, p)$ yazılır. X , başarı olasılığı p olan bağımsız Bernoulli denemelerinde **r -inci başarıya kadar yapılan deneme sayısıdır**. $r = 1$ için $\binom{x-1}{0} = 1$ olduğundan $\text{NB}(1, p) = \text{Geo}(p)$ 'dir.

Geometrik dağılımda olduğu gibi burada da ikinci bir parametrisasyon vardır: r -inci başarıdan önceki başarısızlık sayısı $Y = X - r$ için

$$f_Y(y) = \binom{y+r-1}{r-1} p^r q^y, \quad y = 0, 1, 2, \dots$$

Bu kitapta deneme sayısı X standart alınır.

Dağılımın adı ikinci parametrisasyondan gelir: $\binom{y+r-1}{r-1} = \binom{y+r-1}{y}$ ve genelleştirilmiş binom katsayısı için $\binom{y+r-1}{y} = (-1)^y \binom{-r}{y}$ özdeşliği sayesinde $f_Y(y) = \binom{-r}{y} p^r (-q)^y$ yazılabilir; yani olasılıklar, üssü **negatif** olan $(1 - q)^{-r}$ binom açılımının terimleridir. Bu gözlemi kanıtlamak için aşağıdaki yardımcı sonuca ihtiyacımız var: geometrik serinin art arda türevleri, negatif binom olasılıklarının toplamını verir.

Lemma 22.1 (Geometrik Serinin Türevleri). $r \in \mathbb{N}$ ve $|s| < 1$ için

$$\sum_{k=0}^{\infty} \binom{k+r-1}{r-1} s^k = \frac{1}{(1-s)^r}.$$

İspat.

$r = 1$ için ifade $\sum_{k \geq 0} s^k = \frac{1}{1-s}$ geometrik serisidir. $r \geq 2$ olsun. $|s| < 1$ üzerinde kuvvet serisi terim terim türevlenebilir; geometrik seriyi $r - 1$ kez türevleyelim. Sol tarafta s^k teriminin $(r - 1)$ -inci türevi $k \geq r - 1$ için $k(k-1)\dots(k-r+2) s^{k-r+1}$, $k < r - 1$ için sıfırdır. $j = k - r + 1$ dersek

$$\frac{d^{r-1}}{ds^{r-1}} \sum_{k=0}^{\infty} s^k = \sum_{j=0}^{\infty} (j+r-1)(j+r-2)\dots(j+1) s^j = \sum_{j=0}^{\infty} \frac{(j+r-1)!}{j!} s^j.$$

Sağ tarafta $\frac{d}{ds}(1-s)^{-m} = m(1-s)^{-m-1}$ olduğundan $r - 1$ türev sonunda

$$\frac{d^{r-1}}{ds^{r-1}} \frac{1}{1-s} = 1 \cdot 2 \cdot \dots \cdot (r-1) \cdot (1-s)^{-r} = \frac{(r-1)!}{(1-s)^r}$$

bulunur. İki tarafı $(r-1)!$ 'e bölersek

$$\sum_{j=0}^{\infty} \frac{(j+r-1)!}{j!(r-1)!} s^j = \sum_{j=0}^{\infty} \binom{j+r-1}{r-1} s^j = \frac{1}{(1-s)^r}.$$

■

Lemma elimizdeyken negatif binom dağılımının temel nicelikleri geometrik dağılımdakiyle aynı yoldan çıkar. Momentleri doğrudan seriden hesaplamak yerine, moment üreten fonksiyonun geometrik dağılımının r -inci kuvveti olduğunu görüp geometrik dağılım için yaptığımız türev hesaplarını yeniden kullanacağız.

Teorem 22.3 (Negatif Binom Dağılımının Moment Üreten Fonksiyonu ve Momentleri). $X \sim \text{NB}(r, p)$ ve $q = 1 - p$ olsun.

(a) $\sum_{x=r}^{\infty} f_X(x) = 1$.

(b) Moment üreten fonksiyon $t < -\ln q$ için tanımlıdır ve

$$M_X(t) = \left(\frac{pe^t}{1-qe^t} \right)^r.$$

(c) $E(X) = \frac{r}{p}$ ve $\text{Var}(X) = \frac{rq}{p^2} = \frac{r(1-p)}{p^2}$.

İspat.

(a) $k = x - r$ dönüşümüyle ve Lemma 22.1’nde $s = q$ alarak

$$\sum_{x=r}^{\infty} \binom{x-1}{r-1} p^r q^{x-r} = p^r \sum_{k=0}^{\infty} \binom{k+r-1}{r-1} q^k = p^r \cdot \frac{1}{(1-q)^r} = \frac{p^r}{p^r} = 1.$$

(b) Aynı dönüşümle

$$M_X(t) = \sum_{x=r}^{\infty} e^{tx} \binom{x-1}{r-1} p^r q^{x-r} = p^r e^{rt} \sum_{k=0}^{\infty} \binom{k+r-1}{r-1} (qe^t)^k.$$

$qe^t < 1$, yani $t < -\ln q$ iken lemmayı $s = qe^t$ ile uygulayabiliriz:

$$M_X(t) = p^r e^{rt} \cdot \frac{1}{(1-qe^t)^r} = \left(\frac{pe^t}{1-qe^t} \right)^r.$$

(c) $g(t) = \frac{pe^t}{1-qe^t}$, geometrik dağılımın moment üreten fonksiyonu olsun; $M_X = g^r$ ’dir. Teorem 22.1’nin ispatında $g(0) = 1$, $g'(0) = \frac{1}{p}$ ve $g''(0) = \frac{1+q}{p^2}$ hesaplanmıştı. Zincir ve çarpım kurallarıyla

$$M'_X(t) = r g(t)^{r-1} g'(t), \quad M''_X(t) = r(r-1) g(t)^{r-2} (g'(t))^2 + r g(t)^{r-1} g''(t).$$

$t = 0$ koyarsak (Teorem 17.4)

$$E(X) = M'_X(0) = r \cdot \frac{1}{p} = \frac{r}{p}, \quad E(X^2) = M''_X(0) = \frac{r(r-1)}{p^2} + \frac{r(1+q)}{p^2}.$$

Buradan

$$\text{Var}(X) = \frac{r(r-1) + r(1+q)}{p^2} - \frac{r^2}{p^2} = \frac{r^2 - r + r + rq - r^2}{p^2} = \frac{rq}{p^2}.$$

■

Moment üreten fonksiyonun geometrik dağılımının r -inci kuvveti çıkması rastlantı değildir. r -inci başarıyı beklemek, ilk başarıyı beklemek, sonra ikinciye beklemek, ... diye r bağımsız beklemeye ayrılır. Aşağıdaki önerme bunu moment üreten fonksiyonun teklik özelliğiyle kesinleştirir; aynı zamanda $E(X) = \frac{r}{p}$ ve $\text{Var}(X) = \frac{rq}{p^2}$ formüllerinin neden geometrik dağılımların r katı olduğunu açıklar.

Önerme 22.2 (Bağımsız Geometrik Değişkenlerin Toplamı Negatif Binomdur). $T_1, T_2, \dots, T_r$ bağımsız rastgele değişkenler ve her biri Geo (p) dağılımına sahip olsun. O zaman

$$X = T_1 + T_2 + \dots + T_r \sim \text{NB} (r, p).$$

İspat.

$t < -\ln q$ olsun. $T_1, \dots, T_r$ bağımsız olduğundan (Tanım 14.1) $e^{tT_1}, \dots, e^{tT_r}$ de bağımsızdır ve bağımsız değişkenlerin çarpımının beklenen değeri beklenen değerlerin çarpımıdır (Teorem 18.1; r çarpan için, $e^{t(T_1+\dots+T_{r-1})}$ ile e^{tT_r} bağımsız olduğundan tümevarımla). Öyleyse

$$M_X(t) = E(e^{t(T_1+\dots+T_r)}) = E(e^{tT_1} e^{tT_2} \dots e^{tT_r}) = \prod_{i=1}^r E(e^{tT_i}) = \prod_{i=1}^r M_{T_i}(t) = \left(\frac{pe^t}{1-qe^t} \right)^r.$$

Son adımda Teorem 22.1 (b) kullanıldı. Bulunan fonksiyon, Teorem 22.3 (b) gereği NB (r, p) dağılımının moment üreten fonksiyonudur ve sıfırın bir komşuluğunda tanımlıdır. Moment üreten fonksiyonun teklik özelliği (Teorem 17.5) gereği $X \sim \text{NB} (r, p)$ 'dir.

Önermenin olasılıksal yorumu şudur: T_1 ilk başarıya kadar geçen deneme sayısı, T_2 ilk başarıdan ikinci başarıya kadar geçen deneme sayısı, ..., T_r ise $(r-1)$ -inci başarıdan r -inci başarıya kadar geçen deneme sayısıdır. Denemeler bağımsız ve aynı dağılımlı olduğundan bu bekleme süreleri de bağımsız ve her biri Geo (p) dağılımlıdır; toplamı r -inci başarının sırasındır. Bu ayrıştırma, momentleri ikinci bir yoldan verir: beklenen değer toplamsallığı (Önerme 18.2, tümevarımla r toplanana genişletilerek) ile $E(X) = r \cdot \frac{1}{p}$, bağımsız toplamın varyansı (Sonuç 18.2, tümevarımla r toplanana genişletilerek) ile $\text{Var}(X) = r \cdot \frac{q}{p^2}$.

Örnek 22.2 (Üçüncü Tura ve İkinci Kusurlu Parça). (a) Düzgün bir madeni para atılıyor. Üçüncü turanın tam beşinci atışta gelme olasılığını ve üçüncü turanın en geç beşinci atışta gelme olasılığını bulunuz. Üçüncü turaya kadar yapılan atış sayısının beklenen değerini ve varyansını hesaplayınız. (b) Bir makinenin ürettiği parçaların yüzde 10'u kusurludur ve parçalar birbirinden bağımsız olarak kusurlu çıkmaktadır. Parçalar tek tek denetleniyor ve ikinci kusurlu parça bulununca denetim durduruluyor. Tam 10 parça denetlenme olasılığını, denetlenen parça sayısının beklenen değerini ve standart sapmasını bulunuz.

Çözüm.

(a) X , üçüncü turanın geldiği atışın sırası olsun; $X \sim \text{NB}(\mathbf{B}, \frac{1}{2})$ ve

$$f_X(x) = \binom{x-1}{2} \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^{x-3} = \binom{x-1}{2} \frac{1}{2^x}, \quad x = 3, 4, 5, \dots$$

Buradan $P(X=5) = \binom{4}{2} \frac{1}{2^5} = \frac{6}{32} = \frac{3}{16}$. En geç beşinci atış için

$$P(X \leq 5) = f_X(3) + f_X(4) + f_X(5) = \binom{2}{2} \frac{1}{8} + \binom{3}{2} \frac{1}{16} + \binom{4}{2} \frac{1}{32} = \frac{1}{8} + \frac{3}{16} + \frac{3}{16} = \frac{1}{2}.$$

Sonucun $\frac{1}{2}$ çıkması bir sağlama sunar: “üçüncü tura en geç beşinci atışta gelir” olayı, “beş atışta en az üç tura gelir” olayıyla aynıdır ve düzgün parada beş atışta en az üç tura ile en çok iki tura bakışımı olduğundan bu olasılık $\frac{1}{2}$ ’dir. Momentler Teorem 22.3 (c) ile

$$E(X) = \frac{r}{p} = \frac{3}{1/2} = 6, \quad \text{Var}(X) = \frac{rq}{p^2} = \frac{3 \cdot \frac{1}{2}}{\frac{1}{4}} = 6.$$

(b) Her parça için “kusurlu” olayını başarı sayalım: $p = 0, 1$, $q = 0, 9$, $r = 2$. Denetlenen parça sayısı $X \sim \text{NB}(2; 0, 1)$ ve

$$P(X=10) = \binom{9}{1} (0, 1)^2 (0, 9)^8 = 9 \cdot 0, 01 \cdot 0, 43046721 \approx 0, 0387.$$

Beklenen değer ve varyans

$$E(X) = \frac{2}{0,1} = 20, \quad \text{Var}(X) = \frac{2 \cdot 0,9}{(0,1)^2} = 180, \quad \sigma_X = \sqrt{180} \approx 13,42.$$

Ortalama 20 parça denetlenir; ama standart sapma 13'ün üzerindedir, yani denetimin çok kısa ya da çok uzun sürmesi olağandır. Karşılaştırma için, ikinci kusurlunun en geç onuncu parçada bulunma olasılığı $P(X \leq 10) = \sum_{x=2}^{10} (x-1)(0,1)^2(0,9)^{x-2} \approx 0,2639$ 'dur.

■

22.4 Hipergeometrik Dağılım

Buraya kadarki bütün dağılımlar bağımsız denemelere dayanıyordu. Şimdi bağımsızlığın bozulduğu en temel duruma bakalım: **iadesiz çekiliş**. N nesneden oluşan bir topluluğun M tanesi “başarılı” (kusurlu parça, as kart, kırmızı bilye), $N - M$ tanesi “başarısız” olsun. Topluluktan rastgele ve iadesiz olarak n nesne çekiliyor ($1 \leq n \leq N$) ve X , çekilenler arasındaki başarılı nesne sayısını gösteriyor. Çekilişler iadeli olsaydı her çekilişte başarı olasılığı $\frac{M}{N}$ olur ve X binom dağılımına uyardı. İadesiz çekilişte ise ikinci çekilişteki başarı olasılığı ilkinin sonucuna bağlıdır; binom modeli geçerli değildir.

Olasılık fonksiyonunu doğrudan sayarak bulalım. Çekilen n nesnenin sırası önemli olmadığından sonuç, N nesneden seçilmiş n elemanlı bir alt kümedir; “rastgele çekiliş” bu $\binom{N}{n}$ alt kümenin eşit olasılıklı olması demektir (Tanım 4.1, Teorem 5.2). $X = x$ olan alt kümeler, M başarılıdan x tanesi ve $N - M$ başarısızdan $n - x$ tanesi seçilerek oluşur; sayıları çarpım kuralıyla $\binom{M}{x} \binom{N-M}{n-x}$ ’dir. Öyleyse

$$P(X = x) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}.$$

Bu ifadenin anlamlı olması için $0 \leq x \leq M$ ve $0 \leq n - x \leq N - M$ gerekir; yani x , $\max\{0, n - (N - M)\}$ ile $\min\{n, M\}$ arasında değişir.

Tanım 22.4 (Hipergeometrik Dağılım). $N \in \mathbb{N}$, $M \in \{0, 1, \dots, N\}$ ve $n \in \{1, 2, \dots, N\}$ olsun. Değer kümesi

$$D_X = \{x \in \mathbb{Z} : \max\{0, n - N + M\} \leq x \leq \min\{n, M\}\}$$

ve olasılık fonksiyonu

$$f_X(x) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}, \quad x \in D_X$$

olan X rastgele değişkenine N, M, n parametrelili **hipergeometrik dağılıma** sahiptir denir ve $X \sim \text{HG}(N, M, n)$ yazılır. X, M tanesi başarılı olan N nesneden iadesiz çekilen n nesne arasındaki **başarılı nesne sayısıdır**.

$b > a$ ya da $b < 0$ iken $\binom{a}{b} = 0$ uzlaşımı benimsenirse formül $x = 0, 1, \dots, n$ için yazılabilir; D_X dışındaki değerlerde pay kendiliğinden sıfırlanır.

Hipergeometrik dağılımın momentlerini binom dağılımındaki gibi moment üreten fonksiyondan çıkaramayız; $M_X(t)$ ’nin kullanışlı bir kapalı biçimi yoktur. Bunun yerine binom dağılımı için de işe yarayan ikinci yolu izleyeceğiz: X ’i **gösterge değişkenlerinin toplamı** olarak yazmak. Fark şudur: burada göstergeler bağımsız değildir, bu yüzden varyans hesabında kovaryans terimleri ortaya çıkar ve sonucu binom varyansından ayıran çarpan tam olarak bu terimlerden gelir.

Teorem 22.4 (Hipergeometrik Dağılımın Momentleri). $X \sim \text{HG}(N, M, n)$ olsun ve $p = \frac{M}{N}$ yazalım.

- (a) $\sum_{x \in D_X} f_X(x) = 1.$
- (b) $E(X) = n \frac{M}{N} = np.$
- (c) $N \geq 2$ için $\text{Var}(X) = n \frac{M}{N} \left(1 - \frac{M}{N}\right) \frac{N-n}{N-1} = np(1-p) \frac{N-n}{N-1}.$

İspat.

(a) Binom teoremi (Teorem 5.4) ile $(1+t)^M (1+t)^{N-M} = (1+t)^N$ özdeşliğinin iki yanında t^n ’nin katsayısını karşılaştıralım. Sol tarafta t^n , birinci çarpandan t^x ve ikincisinden t^{n-x} alınarak oluşur;

katsayısı $\sum_x \binom{M}{x} \binom{N-M}{n-x}$ 'dir (toplam, her iki binom katsayısının sıfırdan farklı olduğu x 'ler, yani $x \in D_X$ üzerindedir). Sağ tarafta katsayı $\binom{N}{n}$ 'dir. Öyleyse

$$\sum_{x \in D_X} \binom{M}{x} \binom{N-M}{n-x} = \binom{N}{n} \quad (\text{Vandermonde özdeşliği})$$

ve iki yanı $\binom{N}{n}$ 'ye bölmek $\sum_x f_X(x) = 1$ verir.

(b) Çekilişi sıralı düşünelim: nesneler tek tek çekilsin. Sıralı sonuçlar, N nesneden n tanesinin sıralı dizilişleridir; sayıları $N(N-1)\cdots(N-n+1)$ 'dir ve hepsi eşit olasılıklıdır. Bu model, sırasız modelle tutarlıdır: belli bir n elemanlı alt kümeye tam $n!$ sıralı diziliş karşılık gelir, dolayısıyla her alt kümenin olasılığı $\frac{n!}{N(N-1)\cdots(N-n+1)} = \frac{1}{\binom{N}{n}}$ 'dir. Öyleyse sıralı modelde de başarılı nesne sayısı HG (N, M, n) dağılımlıdır.

$k = 1, \dots, n$ için I_k , k -inci çekilen nesne başarılıysa 1, değilse 0 olan gösterge değişkeni olsun. Açıkça $X = I_1 + I_2 + \cdots + I_n$ 'dir. k -inci çekilişte başarılı gelen sıralı dizilişleri sayalım: k -ıncı konuma M başarılıdan biri gelir (M seçenek), geri kalan $n-1$ konum kalan $N-1$ nesneden sırayla doldurulur ($(N-1)(N-2)\cdots(N-n+1)$ seçenek). Dolayısıyla

$$P(I_k = 1) = \frac{M \cdot (N-1)(N-2)\cdots(N-n+1)}{N(N-1)\cdots(N-n+1)} = \frac{M}{N} = p.$$

Bu, her çekilişin, ne kadar sonra yapılsa yapılsın, ilk çekilişle aynı başarı olasılığına sahip olduğunu söyler. $E(I_k) = 1 \cdot p + 0 \cdot (1-p) = p$ olduğundan, beklenen değer toplamsallığıyla (Önerme 18.2, tümevarımla n toplanana genişletilerek)

$$E(X) = \sum_{k=1}^n E(I_k) = np = n \frac{M}{N}.$$

(c) I_k yalnızca 0 ve 1 değerlerini aldığından $I_k^2 = I_k$ ve $E(I_k^2) = p$; dolayısıyla $\text{Var}(I_k) = p - p^2 = p(1-p)$. Göstergeler bağımsız olmadığından kovaryansları gerekir. $j \neq k$ için $I_j I_k = 1$ olması, j -inci ve k -ıncı konumların ikisine de başarılı gelmesi demektir: bu konumlar $M(M-1)$ biçimde doldurulur, kalan $n-2$ konum kalan $N-2$ nesneden $(N-2)(N-3)\cdots(N-n+1)$ biçimde doldurulur. Öyleyse

$$E(I_j I_k) = P(I_j = 1, I_k = 1) = \frac{M(M-1)(N-2)\cdots(N-n+1)}{N(N-1)(N-2)\cdots(N-n+1)} = \frac{M(M-1)}{N(N-1)}$$

ve

$$\text{Cov}(I_j, I_k) = E(I_j I_k) - E(I_j)E(I_k) = \frac{M(M-1)}{N(N-1)} - \frac{M^2}{N^2} = \frac{M}{N^2} \left[\frac{N(M-1) - M(N-1)}{N-1} \right] = \frac{M(M-N)}{N^2(N-1)} = -\frac{p(1-p)}{N-1}.$$

Kovaryans negatiftir: bir çekilişte başarılı gelmesi, kalan başarılı sayısını azalttığından ötekilerin başarı şansını düşürür. Toplamın varyansı formülünün n değişkenli biçimiyle (Teorem 18.4), n tane varyans ve $\binom{n}{2} = \frac{n(n-1)}{2}$ tane sırasız çift olduğundan

$$\begin{aligned} \text{Var}(X) &= \sum_{k=1}^n \text{Var}(I_k) + 2 \sum_{1 \leq j < k \leq n} \text{Cov}(I_j, I_k) = np(1-p) + 2 \cdot \frac{n(n-1)}{2} \cdot \left(-\frac{p(1-p)}{N-1} \right) \\ &= np(1-p) \left[1 - \frac{n-1}{N-1} \right] = np(1-p) \frac{N-1-(n-1)}{N-1} = np(1-p) \frac{N-n}{N-1}. \end{aligned}$$

■

Sonuçları binom dağılımıyla (Teorem 20.2) karşılaştıralım. $p = \frac{M}{N}$ olmak üzere hipergeometrik beklenen değer np , binomunkiyle aynıdır: iadesiz çekiliş ortalama başarı sayısını değiştirmez. Varyans ise $np(1-p)$ 'nin $\frac{N-n}{N-1}$ katıdır. Bu çarpana **sonlu kitle düzeltmesi** denir; $n \geq 2$ iken 1'den küçüktür, yani iadesiz çekiliş yayılımı azaltır. İki uç durum anlamlıdır: $n = 1$ için çarpan 1'dir (tek çekilişte iadeli-iadesiz farkı yoktur); $n = N$ için çarpan 0'dır (bütün nesneler çekilince $X = M$ kesindir). N büyüdükçe çarpan 1'e yaklaşır; bunu aşağıda kesinleştireceğiz.

Örnek 22.3 (Torbadan İadesiz Çekiliş ve Parti Denetimi). (a) Bir torbada 4 sarı ve 2 kırmızı bilye vardır. Torbadan iadesiz olarak art arda 3 bilye çekiliyor. X , çekilen kırmızı bilye sayısı olsun. X 'in olasılık fonksiyonunu tablo hâlinde yazınız; $E(X)$ ve $\text{Var}(X)$ 'i hem tablodan hem de Teorem 22.4 ile hesaplayınız.

(b) 50 parçalık bir partide 5 kusurlu parça vardır. Partiden rastgele 10 parça iadesiz çekilip denetleniyor. Örnekte hiç kusurlu parça çıkmama olasılığını, en az bir kusurlu çıkma olasılığını, kusurlu sayısının beklenen değerini ve varyansını bulunuz. Sonuçları $B(10; 0, 1)$ dağılımının verdiği değerlerle karşılaştırınız.

(c) (b) şıkkındaki durumu kusurlu oranı $\frac{1}{10}$ 'da sabit tutarak $N = 500$, $N = 5000$ ve $N = 50000$ parçalık partiler için yineleyiniz ve gözlemlerinizi yorumlayınız.

Çözüm.

(a) Bu örnek Örnek 10.3'de elle sayarak çözülmüştü; şimdi onun bir hipergeometrik dağılım olduğunu görüyoruz. $N = 6$ nesne, $M = 2$ başarılı (kırmızı), $n = 3$ iadesiz çekiliş: $X \sim \text{HG}(6, 2, 3)$. $\max\{0, 3 - 4\} = 0 \leq x \leq \min\{3, 2\} = 2$ ve $\binom{6}{3} = 20$ olduğundan

$$f_X(x) = \frac{\binom{2}{x} \binom{4}{3-x}}{20}, \quad x = 0, 1, 2.$$

x	0	1	2
$f_X(x)$	$\frac{\binom{2}{0} \binom{4}{3}}{20} = \frac{4}{20} = \frac{1}{5}$	$\frac{\binom{2}{1} \binom{4}{2}}{20} = \frac{12}{20} = \frac{3}{5}$	$\frac{\binom{2}{2} \binom{4}{1}}{20} = \frac{4}{20} = \frac{1}{5}$

Toplam $\frac{1}{5} + \frac{3}{5} + \frac{1}{5} = 1$ 'dir. Tablodan

$$E(X) = 0 \cdot \frac{1}{5} + 1 \cdot \frac{3}{5} + 2 \cdot \frac{1}{5} = 1, \quad E(X^2) = 0 + \frac{3}{5} + \frac{4}{5} = \frac{7}{5}, \quad \text{Var}(X) = \frac{7}{5} - 1 = \frac{2}{5}.$$

Formüllerle: $p = \frac{M}{N} = \frac{1}{3}$ için $E(X) = 3 \cdot \frac{1}{3} = 1$ ve

$$\text{Var}(X) = 3 \cdot \frac{1}{3} \cdot \frac{2}{3} \cdot \frac{6-3}{6-1} = \frac{2}{3} \cdot \frac{3}{5} = \frac{2}{5}.$$

İki yol da aynı sonucu verir. Karşılaştırma için, çekilişler iadeli olsaydı $X \sim (B, \frac{1}{3})$ ve $\text{Var}(X) = 3 \cdot \frac{1}{3} \cdot \frac{2}{3} = \frac{2}{3}$ olurdu; sonlu kitle düzeltmesi $\frac{3}{5}$ varyansı $\frac{2}{5}$ 'e indirir.

(b) $X \sim \text{HG}(50, 5, 10)$. Hiç kusurlu çıkmaması:

$$P(X = 0) = \frac{\binom{5}{0} \binom{45}{10}}{\binom{50}{10}} = \frac{3190187286}{10272278170} = \frac{82251}{264845} \approx 0,3106.$$

Buradan $P(X \geq 1) = 1 - P(X = 0) \approx 0,6894$. Ayrıca $P(X = 1) = \frac{5 \binom{45}{9}}{\binom{50}{10}} \approx 0,4313$. Momentler, $p = \frac{5}{50} = \frac{1}{10}$ ile

$$E(X) = 10 \cdot \frac{1}{10} = 1, \quad \text{Var}(X) = 10 \cdot \frac{1}{10} \cdot \frac{9}{10} \cdot \frac{50-10}{50-1} = \frac{9}{10} \cdot \frac{40}{49} = \frac{36}{49} \approx 0,735.$$

Binom modeli $B(10; 0,1)$ ise $P(X = 0) = (0,9)^{10} \approx 0,3487$, $P(X = 1) = 10 \cdot 0,1 \cdot (0,9)^9 \approx 0,3874$ ve $\text{Var}(X) = 0,9$ verir. Beklenen değerler aynı, ama olasılıklar ve varyans belirgin biçimde farklıdır: örnek partinin beşte birini kapsadığından $(\frac{n}{N} = 0,2)$ iadesizlik önemlidir.

(c) Kusurlu oranı $\frac{1}{10}$ olan N parçalık partide $M = \frac{N}{10}$ ve $n = 10$ için hipergeometrik olasılıklar aşağıdaki gibidir.

N	50	500	5000	50000	binom ($N \rightarrow \infty$)
$P(X = 0)$	0,3106	0,3452	0,3483	0,3486	0,3487
$P(X = 1)$	0,4313	0,3913	0,3878	0,3875	0,3874
$\text{Var}(X)$	0,7347	0,8838	0,8984	0,8998	0,9

N büyüdükçe hipergeometrik değerler binom değerlerine yaklaşır. Nedeni açıktır: 50000 parçadan 10 tanesini çekmek partiyi neredeyse hiç değiştirmez, her çekilişte kusurlu oranı $\frac{1}{10}$ 'a çok yakın kalır ve çekilişler pratik olarak bağımsızdır. Bu gözlemi bir teorem olarak ispatlayacağız.

■

22.4.1 Binom Dağılımına Yaklaşım

Örneğin (c) şikkındaki tablo, kitle büyüdükçe iadesizliğin önemini yitirdiğini gösterdi. Teorem, bunun bir limit ifadesi olduğunu söyler: örnek büyüklüğü n sabit tutulup kitle sonsuza götürüldüğünde, başarılı oranı p 'ye yaklaştığı sürece, hipergeometrik olasılıklar binom olasılıklarına yakınsar.

Teorem 22.5 (Hipergeometrik Dağılımın Binom Dağılımına Yaklaşımı). $n \in \mathbb{N}$ ve $x \in \{0, 1, \dots, n\}$ sabit olsun. $(M_N)_{N \geq 1}$, $M_N \in \{0, 1, \dots, N\}$ olan ve

$$\lim_{N \rightarrow \infty} \frac{M_N}{N} = p \in (0, 1)$$

koşulunu sağlayan bir dizi olsun. $X_N \sim \text{HG}(N, M_N, n)$ ise

$$\lim_{N \rightarrow \infty} P(X_N = x) = \binom{n}{x} p^x (1-p)^{n-x}.$$

İspat.

$M = M_N$ yazalım. $\frac{M}{N} \rightarrow p > 0$ olduğundan $M \rightarrow \infty$, $\frac{N-M}{N} \rightarrow 1-p > 0$ olduğundan $N-M \rightarrow \infty$ 'dur; dolayısıyla yeterince büyük N için $M \geq x$ ve $N-M \geq n-x$ olur ve $x \in D_{X_N}$ 'dir. Binom katsayılarını açık çarpımlar olarak yazalım:

$$\binom{M}{x} = \frac{M(M-1)\cdots(M-x+1)}{x!}, \quad \binom{N-M}{n-x} = \frac{(N-M)(N-M-1)\cdots(N-M-n+x+1)}{(n-x)!}, \quad \binom{N}{n} = \frac{N(N-1)\cdots(N-n+1)}{n!}.$$

Böylece

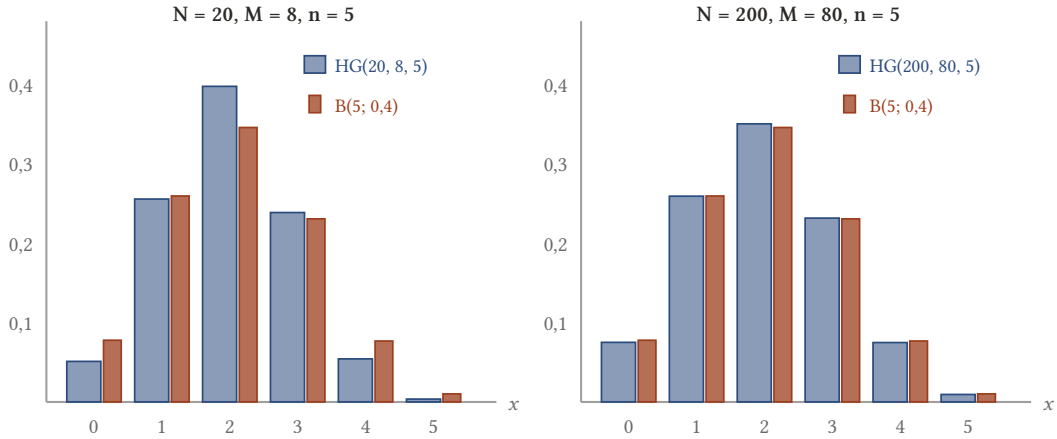
$$P(X_N = x) = \frac{n!}{x!(n-x)!} \cdot \frac{M(M-1)\cdots(M-x+1) \cdot (N-M)(N-M-1)\cdots(N-M-n+x+1)}{N(N-1)\cdots(N-n+1)}.$$

İlk çarpan $\binom{n}{x}$ 'tir. İkinci çarpanın payında $x + (n-x) = n$, paydasında n çarpan vardır; her birini N 'ye bölelim:

$$P(X_N = x) = \binom{n}{x} \cdot \frac{\prod_{i=0}^{x-1} \left(\frac{M}{N} - \frac{i}{N}\right) \cdot \prod_{j=0}^{n-x-1} \left(\frac{N-M}{N} - \frac{j}{N}\right)}{\prod_{k=0}^{n-1} \left(1 - \frac{k}{N}\right)}.$$

$N \rightarrow \infty$ iken çarpan sayıları sabit kaldığından limitler çarpım içine alınabilir: $\frac{M}{N} - \frac{i}{N} \rightarrow p$, $\frac{N-M}{N} - \frac{j}{N} \rightarrow 1-p$ ve $1 - \frac{k}{N} \rightarrow 1$. Öyleyse

$$\lim_{N \rightarrow \infty} P(X_N = x) = \binom{n}{x} \cdot \frac{p^x (1-p)^{n-x}}{1} = \binom{n}{x} p^x (1-p)^{n-x}.$$



Şekil 22.3: İadesiz çekilişte başarı sayısının hipergeometrik dağılımı (geniş mavi çubuk) ile aynı n ve $p = M/N$ için binom dağılımı (ince turuncu çubuk). Solda kitle küçükken ($n/N = 0,25$) hipergeometrik daha sivri: iadesiz çekiliş varyansı küçültür. Sağda kitle on kat büyüyünce ($n/N = 0,025$) çubuklar neredeyse çakışır; $N \rightarrow \infty$ iken hipergeometrik dağılım binom dağılımına yaklaşır.

■

Teorem, varyans formülüyle de uyumludur: $\frac{M}{N} \rightarrow p$ ve $\frac{N-n}{N-1} \rightarrow 1$ olduğundan $\text{Var}(X_N) \rightarrow np(1-p)$, yani binom varyansına gider. Örnekteki tablonun son satırı tam olarak bunu göstermektedir.

🔍 Uygulamada hangi model?

Bir kitleden örnek çekerken doğru model hipergeometriktir; binom, hesabı kolaylaştıran bir yaklaşımdır. Yaygın kural şudur: örnek kitlenin yüzde 5'inden küçükse ($n \leq 0,05N$) binom yaklaşımı yeterince iyidir. Bu durumda $N - n \geq 0,95N$ olduğundan sonlu kitle düzeltmesi $\frac{N-n}{N-1} > 0,95$ 'tir ve binom varyansı gerçek varyansı en çok yüzde 5 dolayında abartır. Örnekteki $N = 50$, $n = 10$ durumu bu kuralı sağlamaz ($\frac{n}{N} = 0,2$) ve gerçekten de iki model belirgin biçimde ayrışır; $N = 500$ için ($\frac{n}{N} = 0,02$) olasılıklardaki bağıl fark yüzde bir dolayına, $N = 5000$ için binde bir dolayına iner.

22.5 Kesikli Dağılımlar Arasındaki İlişkiler

Bu kısımda, şimdiye kadar tanıdığımız yedi kesikli dağılımı bir araya getirelim. Tabloda $q = 1 - p$ 'dir; hipergeometrik dağılım için $p = \frac{M}{N}$ kısaltması kullanılmıştır. Önceki bölümlerdeki gibi Bernoulli dağılımı $B(1, p)$, binom dağılımı $B(n, p)$ ve Poisson dağılımı $\text{Poisson}(\lambda)$ ile gösterilmiştir.

dağılım	D_X	$f_X(x)$	$E(X)$	$\text{Var}(X)$	$M_X(t)$
Bernoulli $B(1, p)$	$\{0, 1\}$	$p^x q^{1-x}$	p	pq	$q + pe^t$
$B(n, p)$	$\{0, \dots, n\}$	$\binom{n}{x} p^x q^{n-x}$	np	npq	$(q + pe^t)^n$
Poisson (λ)	$\{0, 1, \dots\}$	$e^{-\lambda} \frac{\lambda^x}{x!}$	λ	λ	$e^{\lambda(e^t-1)}$
kesikli düzgün	$\{1, \dots, n\}$	$\frac{1}{n}$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$	$\frac{e^t(1-e^{nt})}{n(1-e^t)}$
Geo (p)	$\{1, 2, \dots\}$	$q^{x-1} p$	$\frac{1}{p}$	$\frac{q}{p^2}$	$\frac{pe^t}{1-qe^t}$
NB (r, p)	$\{r, r+1, \dots\}$	$\binom{x-1}{r-1} p^r q^{x-r}$	$\frac{r}{p}$	$\frac{rq}{p^2}$	$\left(\frac{pe^t}{1-qe^t}\right)^r$

dağılım	D_X	$f_X(x)$	$E(X)$	$\text{Var}(X)$	$M_X(t)$
HG (N, M, n)	$\{\max\{0, n - \frac{(\frac{M}{x})(\frac{N-M}{N-x})}{N+M}\}, \dots, \min\{n, \frac{M}{N}\}\}$	$\frac{(\frac{M}{x})(\frac{N-M}{N-x})}{N+M}$	np	$npq \frac{N-n}{N-1}$	kullanışlı ka- palı biçimi yok

Tablo satırları birbirinden bağımsız değildir; dağılımlar, toplam alma ve limit işlemleriyle birbirine bağlıdır. Aşağıdaki önerme bu bağları toplu olarak verir. Çoğu, daha önce ispatladığımız sonuçların yeniden ifadesidir; yeni olan (e) ve (f) maddeleri, negatif binom ile binom dağılımının aynı deneme dizisi üzerindeki ilişkisini ve hipergeometrik dağılımın $n = 1$ durumunu açıklar.

Önerme 22.3 (Kesikli Dağılımlar Arasındaki İlişkiler). (a) Bernoulli dağılımı $B(1, p)$, $n = 1$ parametrelili binom dağılımıdır; n bağımsız $B(1, p)$ değişkeninin toplamı $B(n, p)$ dağılımıdır.

(b) $np_n \rightarrow \lambda > 0$ olacak biçimde $p_n \rightarrow 0$ ise $B(n, p_n)$ olasılıkları Poisson (λ) olasılıklarına yakınsar.

(c) Geo (p) dağılımı NB $(1, p)$ dağılımıdır; r bağımsız Geo (p) değişkeninin toplamı NB (r, p) dağılımıdır.

(d) $\frac{M}{N} \rightarrow p \in (0, 1)$ ve n sabit iken HG (N, M, n) olasılıkları $B(n, p)$ olasılıklarına yakınsar.

(e) Aynı Bernoulli deneme dizisinde X , r -inci başarıya kadar yapılan deneme sayısı ve Y_n , ilk n denemede başarı sayısı olsun ($n \geq r$). O zaman

$$P(X \leq n) = P(Y_n \geq r), \quad \text{yani} \quad \sum_{x=r}^n \binom{x-1}{r-1} p^r q^{x-r} = \sum_{y=r}^n \binom{n}{y} p^y q^{n-y}.$$

(f) HG $(N, M, 1)$ dağılımı, $p = \frac{M}{N}$ parametrelili $B(1, p)$ Bernoulli dağılımıdır.

İspat.

(a) Tanım 20.2 ve Tanım 20.3’nde $n = 1$ için olasılık fonksiyonları çıkarılır. Toplam iddiası Önerme 20.2’dir; moment üreten fonksiyonla da görülür: n bağımsız Bernoulli değişkeninin toplamının moment üreten fonksiyonu $(q + pe^t)^n$, yani binom dağılımının moment üreten fonksiyonudur (Teorem 20.3) ve teklik teoremi (Teorem 17.5) sonucu verir.

(b) Bu, Teorem 21.4’nin ifadesidir.

(c) İlk kısım Tanım 22.3’nde gözlenmişti; ikincisi Önerme 22.2’dir.

(d) Bu, Teorem 22.5’nin ifadesidir.

(e) İki olayın aynı olduğunu gösterelim. “ r -inci başarı en geç n -inci denemede gelir” demek, ilk n denemede en az r başarı bulunması demektir: r -inci başarı $x \leq n$ sırasında geldiyse ilk n denemede en az r başarı vardır; tersine ilk n denemede en az r başarı varsa r -inci başarı bu n denemenin birinde gelmiştir. Öyleyse $(X \leq n) = (Y_n \geq r)$ ve olasılıkları eşittir. $X \sim \text{NB}(r, p)$ ve $Y_n \sim B(n, p)$ olduğundan olasılıkları yazmak formülü verir.

(f) $0 < M < N$ olsun (aksi hâlde X sabittir). $n = 1$ için $D_X = \{0, 1\}$ ve $\binom{N}{1} = N$; dolayısıyla

$$f_X(1) = \frac{\binom{M}{1}\binom{N-M}{0}}{N} = \frac{M}{N}, \quad f_X(0) = \frac{\binom{M}{0}\binom{N-M}{1}}{N} = \frac{N-M}{N} = 1 - \frac{M}{N}.$$

Bu, $p = \frac{M}{N}$ parametrelili Bernoulli dağılımıdır.

■

(e) maddesi, negatif binom dağılımının dağılım fonksiyonunu binom tablosundan okumayı sağlar; Örnek 22.2 (a)’da $P(X \leq 5) = P(Y_5 \geq 3) = \frac{1}{2}$ eşitliğini bu yolla doğrulamıştık. Resmin bütünü şudur: Bernoulli denemeleri, bütün bu dağılımların ortak kaynağıdır. Deneme sayısını sabitleyip başarıları sayarsak binom; başarı sayısını sabitleyip denemeleri sayarsak geometrik ve negatif binom; nadir olaylar için binomu limite götürürsek Poisson; bağımsızlığı iadesiz çekilişle bozarsak hipergeometrik dağılım ortaya çıkar ve kitle büyüyünce bağımsızlık, dolayısıyla binom dağılımı geri gelir.

22.6 Alıştırmalar

Alıştırma 22.1 (Kesikli Düzgün, Geometrik, Negatif Binom ve Hipergeometrik Dağılımlar). (a) İyice karıştırılmış, 1’den 10’a kadar numaralanmış 10 karttan biri çekiliyor. Kart numarasının beklenen değerini ve varyansını bulunuz.

(b) Bir okçu her atışta hedefi bağımsız olarak 0, 2 olasılıkla vuruyor. İlk isabetin en geç üçüncü atışta gelme olasılığını ve ilk iki atışın iskaladığı bilindiğinde ilk isabetin beşinci atıştan sonra gelme olasılığını hesaplayınız. Atış sayısının beklenen değerini ve varyansını yazınız.

(c) $X \sim \text{Geo}(p)$ ve $Y = X - 1$ (ilk başarıdan önceki başarısızlık sayısı) olsun. Y 'nin moment üreten fonksiyonunu, beklenen değerini ve varyansını bulunuz.

(d) $X \sim \text{Geo}(p)$ için X 'in çift sayı olma olasılığının $\frac{q}{1+q}$ olduğunu gösteriniz ve zar örneği ($p = \frac{1}{6}$) için sayısal değerini bulunuz.

(e) Bir zar, ikinci kez 5 ya da 6 gelene kadar atılıyor. Tam dört atış yapılma olasılığını, atış sayısının beklenen değerini ve varyansını bulunuz.

(f) İyice karıştırılmış 52 kartlık bir desteden 5 kart çekiliyor. X , çekilen as sayısı olsun. $P(X = 0)$, $P(X = 1)$, $P(X \geq 2)$ olasılıklarını, $E(X)$ ve $\text{Var}(X)$ 'i hesaplayınız.

(g) $X \sim \text{HG}(N, M, n)$ ve $p = \frac{M}{N}$ olsun. $\text{Var}(X) \leq np(1-p)$ olduğunu, eşitliğin ancak $n = 1$ için sağlandığını ve $n = N$ için $\text{Var}(X) = 0$ olduğunu gösteriniz; son durumu olasılıksal olarak yorumlayınız.

Çözüm.

(a) Kart numarası X , $\{1, \dots, 10\}$ üzerinde kesikli düzgün dağılımlıdır. Önerme 22.1 ile

$$E(X) = \frac{10 + 1}{2} = \frac{11}{2}, \quad \text{Var}(X) = \frac{10^2 - 1}{12} = \frac{99}{12} = \frac{33}{4}.$$

(b) Atış sayısı $X \sim \text{Geo}(0, 2)$, $q = 0, 8$. Teorem 22.1 (a) ile

$$P(X \leq 3) = 1 - q^3 = 1 - 0,512 = 0,488.$$

Hafızasızlık (Teorem 22.2) ile $P(X > 5 \mid X > 2) = P(X > 3) = q^3 = 0,512$. Momentler: $E(X) = \frac{1}{0,2} = 5$ ve $\text{Var}(X) = \frac{0,8}{0,04} = 20$.

(c) $Y = X - 1$ olduğundan $M_Y(t) = E(e^{t(X-1)}) = e^{-t} M_X(t)$; Teorem 22.1 (b) ile

$$M_Y(t) = e^{-t} \cdot \frac{pe^t}{1 - qe^t} = \frac{p}{1 - qe^t}, \quad t < -\ln q.$$

Doğrudan da görülebilir: $M_Y(t) = \sum_{y \geq 0} e^{ty} q^y p = p \sum_{y \geq 0} (qe^t)^y = \frac{p}{1 - qe^t}$. Teorem 16.2 gereği

$$E(Y) = E(X) - 1 = \frac{1}{p} - 1 = \frac{q}{p}, \quad \text{Var}(Y) = \text{Var}(X) = \frac{q}{p^2}.$$

(d) X 'in çift olması, $X \in \{2, 4, 6, \dots\}$ demektir. Ayırık olayların olasılıkları toplanır:

$$P(X \text{ çift}) = \sum_{k=1}^{\infty} f_X(2k) = \sum_{k=1}^{\infty} q^{2k-1} p = pq \sum_{k=0}^{\infty} (q^2)^k = \frac{pq}{1 - q^2} = \frac{pq}{(1 - q)(1 + q)} = \frac{q}{1 + q}.$$

$p = \frac{1}{6}$, $q = \frac{5}{6}$ için $P(X \text{ çift}) = \frac{5/6}{11/6} = \frac{5}{11} \approx 0,4545$. Olasılığın $\frac{1}{2}$ 'den küçük olması doğaldır: tek sayılı atışlar her zaman “önce” gelir ve ilk atışın tek başına $\frac{1}{6}$ olasılığı vardır.

(e) Her atışta “5 ya da 6 gelmesi” başarısı $p = \frac{2}{6} = \frac{1}{3}$ olasılıklıdır, $q = \frac{2}{3}$. Atış sayısı $X \sim \text{NB}(\frac{1}{3})$ ve

$$P(X = 4) = \binom{3}{1} \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^2 = 3 \cdot \frac{1}{9} \cdot \frac{4}{9} = \frac{12}{81} = \frac{4}{27} \approx 0,1481.$$

Teorem 22.3 (c) ile $E(X) = \frac{2}{1/3} = 6$ ve $\text{Var}(X) = \frac{2 \cdot \frac{2}{3}}{1/9} = 12$.

(f) $X \sim \text{HG}(52, 4, 5)$; $\binom{52}{5} = 2598960$.

$$P(X = 0) = \frac{\binom{4}{0} \binom{48}{5}}{\binom{52}{5}} = \frac{1712304}{2598960} = \frac{35673}{54145} \approx 0,6588, \quad P(X = 1) = \frac{\binom{4}{1} \binom{48}{4}}{\binom{52}{5}} = \frac{4 \cdot 194580}{2598960} = \frac{778320}{2598960} \approx 0,2995.$$

Dolayısıyla $P(X \geq 2) = 1 - 0,6588 - 0,2995 \approx 0,0417$. Beş kartlık bir elde iki ya da daha çok as bulunma olasılığı yaklaşık yüzde 4'tür. Momentler, $p = \frac{4}{52} = \frac{1}{13}$ ile

$$E(X) = 5 \cdot \frac{1}{13} = \frac{5}{13} \approx 0,385, \quad \text{Var}(X) = 5 \cdot \frac{1}{13} \cdot \frac{12}{13} \cdot \frac{52 - 5}{52 - 1} = \frac{60}{169} \cdot \frac{47}{51} = \frac{940}{2873} \approx 0,327.$$

(g) **Teorem 22.4** (c) gereği $\text{Var}(X) = np(1-p) \frac{N-n}{N-1}$ 'dir. $1 \leq n \leq N$ olduğundan $N-n \leq N-1$, yani $\frac{N-n}{N-1} \leq 1$; $np(1-p) \geq 0$ ile çarpınca $\text{Var}(X) \leq np(1-p)$ çıkar. Eşitlik için ya $np(1-p) = 0$ (yani $M = 0$ ya da $M = N$; bu durumda X zaten sabittir ve her iki taraf sıfırdır) ya da $\frac{N-n}{N-1} = 1$, yani $n = 1$ gerekir; $0 < p < 1$ olan asıl durumda eşitlik yalnızca $n = 1$ 'de sağlanır. $n = N$ için çarpan $\frac{N-N}{N-1} = 0$ ve $\text{Var}(X) = 0$ 'dır. Yorum: bütün nesneler çekildiğinde çekilenler arasındaki başarılı sayısı kesinlikle M 'dir; gerçekten $D_X = \{M\}$ ve $f_X(M) = \frac{\binom{M}{M}\binom{N-M}{N-M}}{\binom{N}{N}} = 1$ olur. Sabit bir rastgele değişkenin varyansı sıfırdır. ■

Kesikli dünyada bekleme sürelerini ve iadesiz çekilişi tamamladık. Bir sonraki bölümde sürekli dünyaya geçiyor; hafızasızlığın sürekli karşılığı olan üstel dağılımı, en yalın sürekli dağılım olan düzgün dağılımı ve olasılık teorisinin merkezindeki normal dağılımı inceliyoruz: [Düzgün, Üstel ve Normal Dağılımlar](#).

23. Düzgün, Üstel ve Normal Dağılımlar

Geometrik, Negatif Binom ve Hipergeometrik Dağılımlar bölümüyle kesikli dağılım ailelerini tamamladık. Bernoulli, binom, Poisson, geometrik, negatif binom ve hipergeometrik dağılımların ortak yanı, değer kümelerinin sayılabilir olmasıydı; her birini bir olasılık fonksiyonuyla (Tanım 10.2) tanımladık. Bu bölümde sürekli tarafa geçiyoruz: değer kümesi bir aralık ya da bütün $\mathbb{R}$ olan, dağılımı bir olasılık yoğunluk fonksiyonuyla (Tanım 11.2) verilen üç temel aileyi inceleyeceğiz.

Sürekli düzgün dağılım, bir aralıktan “rastgele nokta seçmek” fikrinin rastgele değişken diliyle ifadesidir; kitabın başındaki geometrik olasılık örnekleri bu dağılımın örnekleriydi. **Üstel dağılım**, bekleme sürelerinin ve ömürlerin dağılımıdır; geometrik dağılımın sürekli karşılığıdır ve onun gibi hafızasızdır. **Normal dağılım** ise olasılık teorisinin en önemli dağılımıdır: ölçüm hatalarını modeller, çok sayıda küçük etkinin toplamı olarak ortaya çıkar ve istatistiğin klasik yöntemlerinin temelinde yatar.

Her dağılım için aynı programı izleyeceğiz: yoğunluğu doğrulamak, dağılım fonksiyonunu bulmak, beklenen değeri, varyansı ve moment üreten fonksiyonu hesaplamak, dağılımı ötekilere bağlayan yapısal özellikleri ispatlamak. Normal dağılımda bu program iki ünlü hesabı içerir: kutupsal koordinatlarla Gauss integrali ve kareye tamamlamayla moment üreten fonksiyon. Bölümün sonunda standart normal tablonun kullanımını sayısal örneklerle göreceğiz. Bu, kitabın son bölümüdür; alıştırılardan sonra kitap boyunca kurduğumuz yapıya toplu olarak bakacağız.

23.1 Sürekli Düzgün Dağılım

Olasılık Uzaı Örnekleri ve Geometrik Olasılık bölümünde bir aralıktan “rastgele” nokta seçme deneyini, alt aralığın olasılığını uzunluğuyla orantılı olarak modellemiştik: (a, b) ’den seçilen noktanın $(c, d) \subseteq (a, b)$ ’ye düşme olasılığı $\frac{d-c}{b-a}$ idi. Bu, Laplace uzayının (Tanım 4.1) sürekli karşılığıdır: her nokta “eşit olasılıklı”dır, ama tek noktanın olasılığı sıfır olduğundan bu eşitlik ancak yoğunluğun sabit olmasıyla ifade edilebilir.

Tanım 23.1 (Sürekli Düzgün Dağılım). $a < b$ reel sayılar olsun. Olasılık yoğunluk fonksiyonu

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{diğer durumlarda} \end{cases}$$

olan sürekli rastgele değişkene (a, b) aralığında **sürekli düzgün dağılıma** sahiptir denir ve $X \sim U(a, b)$ yazılır. $U(0, 1)$ dağılımına **standart düzgün dağılım** denir.

f_X gerçekten bir yoğunluktur: negatif değildir ve

$$\int_{-\infty}^{\infty} f_X(x) dx = \int_a^b \frac{1}{b-a} dx = \frac{b-a}{b-a} = 1.$$

Dağılım fonksiyonunu (Tanım 9.1) yoğunluğu integre ederek buluruz. $x \leq a$ için integrand sıfırdır, $F_X(x) = 0$. $a < x < b$ için

$$F_X(x) = \int_a^x \frac{1}{b-a} du = \frac{x-a}{b-a}.$$

$x \geq b$ için $F_X(x) = 1$. Böylece

$$F_X(x) = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a < x < b \\ 1, & x \geq b. \end{cases}$$

$a \leq c < d \leq b$ için

$$P(c < X < d) = F_X(d) - F_X(c) = \frac{d-c}{b-a};$$

yani bir alt aralığın olasılığı yalnızca uzunluğuna bağlıdır, nerede durduğuna değil. Bu tam olarak geometrik olasılık modelidir. Sürekli bir rastgele değişkende tek noktanın olasılığı sıfır olduğundan (Önerme 11.1) uç noktaların aralığa dâhil olup olmaması hiçbir olasılığı değiştirmez; $U(a, b)$ ile $U[a, b]$ arasında ayırım yapılmaz.

Düzgün dağılımın sayısal özelliklerini hesaplayalım. Simetri nedeniyle beklenen değer aralığın orta noktası olması beklenir; varyans ise aralık uzunluğunun karesiyle ölçeklenmelidir.

Teorem 23.1 (Düzgün Dağılımın Momentleri ve Moment Üreten Fonksiyonu). $X \sim U(a, b)$ olsun. O zaman

$$E(X) = \frac{a+b}{2}, \quad \text{Var}(X) = \frac{(b-a)^2}{12}$$

ve moment üreten fonksiyon her $t \in \mathbb{R}$ için sonludur:

$$M_X(t) = \begin{cases} \frac{e^{bt}-e^{at}}{(b-a)t}, & t \neq 0 \\ 1, & t = 0. \end{cases}$$

İspat.

Beklenen değer tanımından (Tanım 16.1)

$$E(X) = \int_a^b \frac{x}{b-a} dx = \frac{1}{b-a} \cdot \frac{b^2-a^2}{2} = \frac{(b-a)(b+a)}{2(b-a)} = \frac{a+b}{2}.$$

İkinci moment:

$$E(X^2) = \int_a^b \frac{x^2}{b-a} dx = \frac{1}{b-a} \cdot \frac{b^3-a^3}{3} = \frac{(b-a)(b^2+ab+a^2)}{3(b-a)} = \frac{a^2+ab+b^2}{3}.$$

Varyans hesap formülü (Önerme 16.3) ile

$$\text{Var}(X) = E(X^2) - (E(X))^2 = \frac{a^2+ab+b^2}{3} - \frac{(a+b)^2}{4} = \frac{4a^2+4ab+4b^2-3a^2-6ab-3b^2}{12} = \frac{a^2-2ab+b^2}{12} = \frac{(b-a)^2}{12}.$$

Moment üreten fonksiyon (Tanım 17.2) için $t \neq 0$ olsun:

$$M_X(t) = E(e^{tX}) = \int_a^b \frac{e^{tx}}{b-a} dx = \frac{1}{b-a} \left[\frac{e^{tx}}{t} \right]_a^b = \frac{e^{bt}-e^{at}}{(b-a)t}.$$

$t = 0$ için $M_X(0) = E(1) = 1$. İntegral her t için sonludur; dolayısıyla M_X bütün $\mathbb{R}$ 'de vardır ve Teorem 17.4 ile momentler ondan da okunabilir. Üstel fonksiyonun seri açılımıyla

$$e^{bt}-e^{at} = \sum_{n=0}^{\infty} \frac{(b^{n+1}-a^{n+1})t^{n+1}}{(n+1)!}$$

olduğundan, $t \neq 0$ için $(b-a)t$ ile bölüp $(n+1)! = (n+1)n!$ yazarsak

$$M_X(t) = \sum_{n=0}^{\infty} \frac{b^{n+1}-a^{n+1}}{(n+1)(b-a)} \cdot \frac{t^n}{n!}$$

elde ederiz. $\frac{t^n}{n!}$ 'in katsayısı $E(X^n)$ olduğundan

$$E(X^n) = \frac{b^{n+1}-a^{n+1}}{(n+1)(b-a)}, \quad n = 1, 2, \dots$$

$n = 1$ için $\frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2}$, $n = 2$ için $\frac{b^3-a^3}{3(b-a)} = \frac{a^2+ab+b^2}{3}$ bulunur; yukarıdaki hesaplarla uyuyor.

■

Özel olarak $U \sim U(0, 1)$ için $E(U) = \frac{1}{2}$, $\text{Var}(U) = \frac{1}{12}$ ve $M_U(t) = \frac{e^t - 1}{t}$ 'dir. Genel durumda standart sapma $\frac{b-a}{2\sqrt{3}} \approx 0,289(b-a)$, yani aralık uzunluğunun yaklaşık yüzde 29'udur.

23.1.1 Olasılık İntegral Dönüşümü ve Ters Dönüşüm Yöntemi

Standart düzgün dağılımın özel bir önemi vardır: **her** sürekli dağılım ona dönüştürülebilir ve ondan üretilir. Bu iki yönlü ilişki, bilgisayarda rastgele sayı üretmenin temelidir. Bilgisayarlar doğrudan $U(0, 1)$ dağılımlı sayılar üretir; istenen dağılıma sahip bir değer elde etmek için bu sayıyı dağılım fonksiyonunun tersinden geçirmek yeter.

Aşağıdaki önermenin (b) kısmında dağılım fonksiyonu kesin artan olmak zorunda olmadığından sıradan ters fonksiyon yerine **genelleştirilmiş ters** kullanılır: F bir dağılım fonksiyonu ve $0 < u < 1$ için

$$F^{-1}(u) = \inf\{x \in \mathbb{R} : F(x) \geq u\}.$$

F sürekli ve kesin artan olduğunda bu, sıradan ters fonksiyondur.

Önerme 23.1 (Olasılık İntegral Dönüşümü ve Ters Dönüşüm Yöntemi). (a) X , dağılım fonksiyonu F_X sürekli olan bir rastgele değişken olsun. O zaman $F_X(X) \sim U(0, 1)$ 'dir.

(b) $U \sim U(0, 1)$ ve F herhangi bir dağılım fonksiyonu olsun. $X = F^{-1}(U)$ rastgele değişkeninin dağılım fonksiyonu F 'dir.

İspat.

(a) $Y = F_X(X)$ diyelim. F_X 'in değerleri $[0, 1]$ 'de olduğundan Y de $[0, 1]$ 'de değer alır; $u < 0$ için $P(Y \leq u) = 0$ ve $u \geq 1$ için $P(Y \leq u) = 1$ 'dir. Geriye $0 < u < 1$ için $P(Y \leq u) = u$ göstermek kalır; $u = 0$ değeri bundan dağılım fonksiyonunun sağdan sürekliliğiyle çıkar: $F_Y(0) = \lim_{u \downarrow 0} F_Y(u) = \lim_{u \downarrow 0} u = 0$.

$A_u = \{x \in \mathbb{R} : F_X(x) \leq u\}$ kümesini alalım. $\lim_{x \rightarrow -\infty} F_X(x) = 0 < u$ olduğundan A_u boş değildir; $\lim_{x \rightarrow \infty} F_X(x) = 1 > u$ olduğundan üstten sınırlıdır. $x_u = \sup A_u$ olsun. F_X sürekli olduğundan $F_X(x_u) = u$ 'dur: A_u içinde $x_n \rightarrow x_u$ olan bir dizi vardır ve $F_X(x_u) = \lim F_X(x_n) \leq u$; öte yandan $x > x_u$ olan her x için $x \notin A_u$, yani $F_X(x) > u$ olduğundan $F_X(x_u) = \lim_{x \rightarrow x_u^+} F_X(x) \geq u$.

Şimdi $(Y \leq u) = (X \leq x_u)$ olduğunu gösterelim. $X \leq x_u$ ise, F_X azalmayan olduğundan $F_X(X) \leq F_X(x_u) = u$; $X > x_u$ ise $X \notin A_u$, yani $F_X(X) > u$. Öyleyse

$$P(Y \leq u) = P(X \leq x_u) = F_X(x_u) = u.$$

Demek ki Y 'nin dağılım fonksiyonu $U(0, 1)$ 'in dağılım fonksiyonudur; dağılım fonksiyonu dağılımı belirlediğinden $Y \sim U(0, 1)$.

(b) $0 < u < 1$ için $B_u = \{x \in \mathbb{R} : F(x) \geq u\}$ olsun. $\lim_{x \rightarrow \infty} F(x) = 1 > u$ olduğundan B_u boş değildir; $\lim_{x \rightarrow -\infty} F(x) = 0 < u$ olduğundan alttan sınırlıdır. Dolayısıyla $F^{-1}(u) = \inf B_u$ sonlu bir reel sayıdır. U 'nun değerleri 1 olasılıkla $(0, 1)$ 'de olduğundan $X = F^{-1}(U)$ iyi tanımlıdır.

Anahtar gözlem şudur: her $x \in \mathbb{R}$ ve $u \in (0, 1)$ için

$$F^{-1}(u) \leq x \Leftrightarrow u \leq F(x).$$

Sağdan sola: $u \leq F(x)$ ise $x \in B_u$, dolayısıyla $\inf B_u \leq x$. Soldan sağa: $m = \inf B_u$ olsun. B_u içinde $y_n \downarrow m$ olan bir dizi vardır ($m \in B_u$ ise sabit dizi alınır); F sağdan sürekli olduğundan $F(m) = \lim F(y_n) \geq u$, yani $m \in B_u$. F azalmayan olduğundan $m \leq x$ ise $F(x) \geq F(m) \geq u$.

Bu denklikle, her $x \in \mathbb{R}$ için

$$P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x);$$

son eşitlikte $U \sim U(0, 1)$ ve $0 \leq F(x) \leq 1$ kullanıldı. Demek ki X 'in dağılım fonksiyonu F 'dir. ■

(a) kısmına **olasılık integral dönüşümü**, (b) kısmına **ters dönüşüm yöntemi** denir. (b) kısmı F 'nin türü hakkında hiçbir şey varsaymaz; (a) kısmındaki süreklilik koşulu ise atılamaz: X kesikli ise $F_X(X)$ yalnızca sayılabilir çoklukta değer alır ve düzgün dağılamaz.

Örnek 23.1 (Düzgün Dağılım Hesapları). (a) Bir durağa otobüs her 30 dakikada bir gelmektedir. Durağa rastgele bir anda gelen yolcunun bekleme süresi X (dakika), $U(0, 30)$ dağılımlıdır. $P(X > 20)$, $P(5 < X < 15)$, $E(X)$ ve σ_X değerlerini bulunuz.

(b) $X \sim U(0, 1)$ ve $Y = X^2$ olsun. Y 'nin yoğunluk fonksiyonunu ve beklenen değerini bulunuz.

(c) Yoğunluğu $f(x) = 2x$ ($0 < x < 1$) olan bir rastgele değişkeni ve $U(a, b)$ dağılımlı bir rastgele değişkeni standart düzgün bir U 'dan ters dönüşüm yöntemiyle üretiniz.

Çözüm.

(a) $b - a = 30$ olduğundan olasılıklar uzunluk oranlarıdır:

$$P(X > 20) = \frac{30 - 20}{30} = \frac{1}{3}, \quad P(5 < X < 15) = \frac{15 - 5}{30} = \frac{1}{3}.$$

Teorem 23.1 ile $E(X) = \frac{0+30}{2} = 15$ dakika, $\text{Var}(X) = \frac{30^2}{12} = 75$ ve $\sigma_X = \sqrt{75} = 5\sqrt{3} \approx 8,66$ dakika.

(b) $g(x) = x^2$ fonksiyonu $(0, 1)$ üzerinde kesin artan ve türevlenebilirdir; tersi $h(y) = \sqrt{y}$, türevi $h'(y) = \frac{1}{2\sqrt{y}}$ 'dir. Dönüşüm formülü (**Teorem 12.2**) ile $0 < y < 1$ için

$$f_Y(y) = f_X(h(y)) |h'(y)| = 1 \cdot \frac{1}{2\sqrt{y}} = \frac{1}{2\sqrt{y}},$$

öteki y 'ler için $f_Y(y) = 0$. Beklenen değer:

$$E(Y) = \int_0^1 y \cdot \frac{1}{2\sqrt{y}} dy = \frac{1}{2} \int_0^1 y^{1/2} dy = \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}.$$

Bu, $E(X^2) = \frac{0+0+1}{3} = \frac{1}{3}$ ile uyudur.

(c) Birinci dağılımın dağılım fonksiyonu $0 \leq x \leq 1$ için $F(x) = \int_0^x 2u du = x^2$ 'dir; $(0, 1)$ üzerinde sürekli ve kesin artandır, tersi $F^{-1}(u) = \sqrt{u}$ 'dur. **Önerme 23.1** gereği $X = \sqrt{U}$ istenen dağılıma sahiptir. Doğrudan sağlama: $0 < x < 1$ için

$$P(\sqrt{U} \leq x) = P(U \leq x^2) = x^2 = F(x).$$

$U(a, b)$ için $F(x) = \frac{x-a}{b-a}$, tersi $F^{-1}(u) = a + (b-a)u$ 'dur; öyleyse $a + (b-a)U \sim U(a, b)$.

■

23.2 Üstel Dağılım

Bir telefon santraline ilk çağrının gelmesine kadar geçen süre, bir ampulün ömrü, radyoaktif bir atomun bozunmasına kadar geçen süre: bunların hepsi pozitif değerli, “ne kadar beklenirse beklensin gelecek” türünden rastgele büyüklüklerdir. Geometrik dağılım (**Tanım 22.2**) ilk başarıya kadar geçen **deneme sayısını** modelliyordu; üstel dağılım, ilk olaya kadar geçen **süreyi** modeller.

Tanım 23.2 (Üstel Dağılım). $\lambda > 0$ olsun. Olasılık yoğunluk fonksiyonu

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & \text{diğer durumlarda} \end{cases}$$

olan sürekli rastgele değişkene λ parametrelili **üstel dağılıma** sahiptir denir ve $X \sim \text{Üstel}(\lambda)$ yazılır. λ 'ya **oran parametresi** denir.

f_X bir yoğunluktur: negatif değildir ve

$$\int_0^\infty \lambda e^{-\lambda x} dx = \left[-e^{-\lambda x} \right]_0^\infty = 0 - (-1) = 1.$$

Dağılım fonksiyonu $x \leq 0$ için $F_X(x) = 0$, $x > 0$ için

$$F_X(x) = \int_0^x \lambda e^{-\lambda u} du = \left[-e^{-\lambda u} \right]_0^x = 1 - e^{-\lambda x}$$

olur. Uygulamalarda dağılım fonksiyonundan çok **yaşam fonksiyonu** denen tümleyeni kullanılır:

$$P(X > x) = 1 - F_X(x) = e^{-\lambda x}, \quad x > 0.$$

Teorem 23.2 (Üstel Dağılımın Momentleri ve Moment Üreten Fonksiyonu). $X \sim \text{Üstel}(\lambda)$ olsun. Moment üreten fonksiyon $t < \lambda$ için sonludur ve

$$M_X(t) = \frac{\lambda}{\lambda - t}, \quad t < \lambda.$$

Ayrıca

$$E(X) = \frac{1}{\lambda}, \quad \text{Var}(X) = \frac{1}{\lambda^2}, \quad E(X^n) = \frac{n!}{\lambda^n} \quad (n \in \mathbb{N}).$$

İspat.

Tanımdan

$$M_X(t) = \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx = \lambda \int_0^\infty e^{-(\lambda-t)x} dx.$$

$t \geq \lambda$ ise integrand her $x \geq 0$ için 1'den küçük değildir ve integral ıraksar; $t < \lambda$ ise

$$M_X(t) = \lambda \left[\frac{e^{-(\lambda-t)x}}{-(\lambda-t)} \right]_0^\infty = \lambda \cdot \frac{1}{\lambda-t} = \frac{\lambda}{\lambda-t}.$$

M_X , sıfırın $(-\lambda, \lambda)$ komşuluğunda sonlu olduğundan [Teorem 17.4](#) uygulanır. Türevler:

$$M'_X(t) = \frac{\lambda}{(\lambda-t)^2}, \quad M''_X(t) = \frac{2\lambda}{(\lambda-t)^3}.$$

$t = 0$ 'da $E(X) = M'_X(0) = \frac{1}{\lambda}$ ve $E(X^2) = M''_X(0) = \frac{2}{\lambda^2}$; dolayısıyla

$$\text{Var}(X) = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

Genel moment için M_X 'i geometrik seri olarak açalım; $|t| < \lambda$ için

$$M_X(t) = \frac{1}{1-t/\lambda} = \sum_{n=0}^{\infty} \left(\frac{t}{\lambda}\right)^n = \sum_{n=0}^{\infty} \frac{n!}{\lambda^n} \cdot \frac{t^n}{n!}.$$

$\frac{t^n}{n!}$ 'in katsayısı $E(X^n)$ olduğundan $E(X^n) = \frac{n!}{\lambda^n}$; $n = 1, 2$ için yukarıdaki değerler yeniden bulunur.

■

Üstel dağılımda standart sapma beklenen değere eşittir: $\sigma_X = \frac{1}{\lambda} = E(X)$. Oran parametresi büyüdükçe olaylar sıklaşır, bekleme kısalmır.

Üstel dağılımın en önemli yapısal özelliği, geometrik dağılımla paylaştığı hafızasızlıktır: bir parça s zaman çalışmışsa, en az t zaman daha çalışma olasılığı, yeni bir parçanın en az t zaman çalışma olasılığına eşittir. Geçmiş, geleceğe dair hiçbir bilgi taşımaz. Teorem bunun tersini de söyler: bu özelliğe sahip tek sürekli dağılım üsteldir.

Teorem 23.3 (Hafızasızlık Özelliği). (a) $X \sim \text{Üstel}(\lambda)$ ise her $s, t \geq 0$ için

$$P(X > s + t \mid X > s) = P(X > t).$$

(b) Tersine, $X, P(X > 0) = 1$ olan bir rastgele değişken ve her $s, t \geq 0$ için

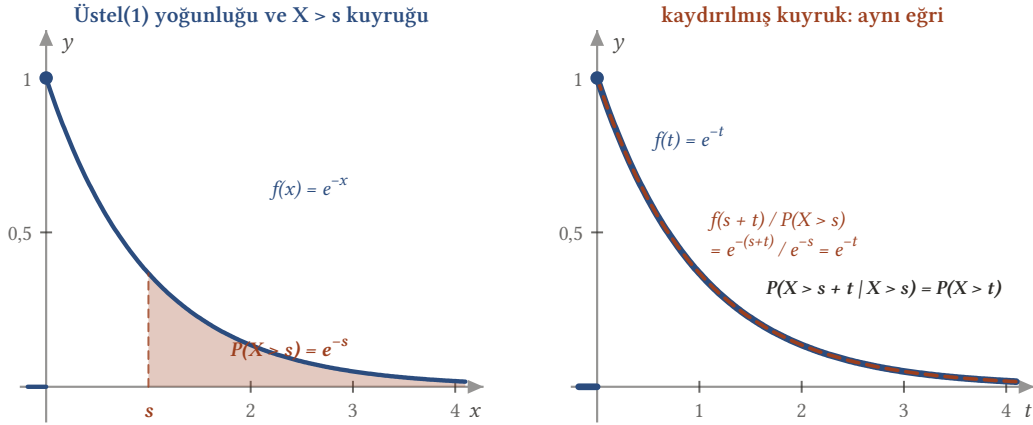
$$P(X > s + t) = P(X > s) P(X > t)$$

ise, bir $\lambda > 0$ için $X \sim \text{Üstel}(\lambda)$ 'dir.

İspat.

(a) $s + t \geq s$ olduğundan $(X > s + t) \subseteq (X > s)$ 'dir. Koşullu olasılık tanımıyla

$$P(X > s + t \mid X > s) = \frac{P(X > s + t, X > s)}{P(X > s)} = \frac{P(X > s + t)}{P(X > s)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = P(X > t).$$



Şekil 23.1: Üstel dağılımın hafızasızlığı ($\lambda = 1$). Solda s 'nin sağında kalan kuyruk, $P(X > s) = e^{-s}$ alanıdır. Sağda bu kuyruk s kadar sola kaydırılıp e^{-s} 'e bölünmüştür: $X > s$ verildiğinde kalan sürenin yoğunluğu (kesikli) orijinal yoğunlukla birebir çakışır. Geçen s süresi, kalan süre hakkında hiçbir bilgi taşımaz: $P(X > s + t \mid X > s) = P(X > t)$.

(b) $G(t) = P(X > t) = 1 - F_X(t)$ yazalım ($t \geq 0$). Varsayım $G(s + t) = G(s)G(t)$ biçimini alır. G artmayan bir fonksiyondur; $G(0) = P(X > 0) = 1$ 'dir ve dağılım fonksiyonunun sonsuzdaki limitinden $G(t) \rightarrow 0$ ($t \rightarrow \infty$).

Adım 1: $0 < G(1) < 1$. Fonksiyonel denklemi ardışık olarak uygulayarak her $n \in \mathbb{N}$ için $G(1) = \left(\frac{1}{n}\right)^n$ ve $G(n) = G(1)^n$ buluruz. $G(1) = 0$ olsaydı her n için $\left(\frac{1}{n}\right)^n = 0$ olurdu; G artmayan olduğundan her $t > 0$ için $G(t) = 0$ olur, dolayısıyla

$$P(X > 0) = P\left(\bigcup_{n=1}^{\infty} \left(X > \frac{1}{n}\right)\right) = \lim_{n \rightarrow \infty} \left(\frac{1}{n}\right)^n = 0$$

çıkardı (artan olaylar dizisinde olasılığın sürekliliği); bu, $P(X > 0) = 1$ ile çelişir. $G(1) = 1$ olsaydı her n için $G(n) = 1$ olur, $G(t) \rightarrow 0$ ile çelişirdi. Öyleyse $0 < G(1) < 1$ ve

$$\lambda = -\ln G(1)$$

sayısı $(0, \infty)$ aralığındadır; $G(1) = e^{-\lambda}$.

Adım 2: rasyonel noktalar. $m, n \in \mathbb{N}$ için $\left(\frac{m}{n}\right)^n = \left(\frac{1}{n}\right)^m$ ve $\left(\frac{1}{n}\right)^n = G(1)^{1/n} = e^{-\lambda/n}$ olduğundan

$$\left(\frac{m}{n}\right)^n = e^{-\lambda m/n}.$$

Yani her pozitif rasyonel q için $G(q) = e^{-\lambda q}$.

Adım 3: bütün pozitif reel sayılar. $t > 0$ irrasyonel olsun. $q_1 < t < q_2$ olan rasyonel q_1, q_2 alalım. G artmayan olduğundan

$$e^{-\lambda q_2} = G(q_2) \leq G(t) \leq G(q_1) = e^{-\lambda q_1}.$$

$q_1 \uparrow t$ ve $q_2 \downarrow t$ alınırda iki uç da $e^{-\lambda t}$ 'ye gider; sıkıştırmayla $G(t) = e^{-\lambda t}$.

Böylece $t > 0$ için $F_X(t) = 1 - e^{-\lambda t}$ ve $P(X > 0) = 1$ nedeniyle $t \leq 0$ için $F_X(t) = 0$ bulundu. Bu, Üstel (λ) dağılımının dağılım fonksiyonudur; dağılım fonksiyonu dağılımı belirlediğinden $X \sim \text{Üstel}(\lambda)$.

■

İ Hafızasızlık yalnızca iki dağılıma özgüdür

Geometrik dağılımda $P(X > n) = (1 - p)^n$, üstel dağılımda $P(X > t) = e^{-\lambda t}$ idi; ikisi de $G(s + t) = G(s)G(t)$ denklemini sağlar ve $1 - p$ ile $e^{-\lambda}$ birbirine karşılık gelir. Yukarıdaki ispat, bu denklemin pozitif reel sayılarda sağlayan, $G(0) = 1$ olan ve sonsuzda sifıra giden tek artmayan fonksiyonun $e^{-\lambda t}$ olduğunu gösterdi; doğal sayılarda aynı akıl yürütme $(1 - p)^n$ 'i verir. Dolayısıyla hafızasız tek sürekli dağılım üstel, hafızasız tek kesikli dağılım geometrik dağılımdır. Yıpranan parçalar için $P(X > s + t | X > s)$, s ile azalır; üstel model yalnızca “yaşlanmayan” parçalara uygundur.

Üstel dağılım, Poisson dağılımının (Tanım 21.1) zaman eksenindeki karşılığıdır. Poisson dağılımı sabit bir zaman aralığındaki **olay sayısını** sayar; üstel dağılım ilk olaya kadar geçen **süreyi** ölçer. Aralarındaki ilişki tek satırlık bir gözleme dayanır: “ilk olay t 'den sonra gelir” ile “ $(0, t]$ aralığında hiç olay yoktur” aynı olaydır.

Önerme 23.2 (Üstel Dağılım ile Poisson Dağılımı Arasındaki İlişki). Zaman içinde gerçekleşen olaylar için her $t > 0$ 'da $(0, t]$ aralığındaki olay sayısı N_t , Poisson (λt) dağılımlı olsun; yani birim zamandaki olay sayısı Poisson (λ) dağılımlıdır. T ilk olayın zamanı ise $T \sim \text{Üstel}(\lambda)$ 'dir.

İspat.

$t > 0$ olsun. İlk olay t 'den sonra gerçekleşir ancak ve ancak $(0, t]$ aralığında hiç olay olmazsa: $(T > t) = (N_t = 0)$. Poisson olasılık fonksiyonuyla

$$P(T > t) = P(N_t = 0) = e^{-\lambda t} \frac{(\lambda t)^0}{0!} = e^{-\lambda t}.$$

Dolayısıyla $t > 0$ için $F_T(t) = 1 - e^{-\lambda t}$; T pozitif olduğundan $t \leq 0$ için $F_T(t) = 0$. Bu, Üstel (λ) dağılımının dağılım fonksiyonudur.

■

Önerme, oran parametresinin anlamını açıklar: birim zamanda ortalama λ olay gerçekleşiyorsa ilk olaya kadar ortalama $\frac{1}{\lambda}$ birim beklenir. Ardışık iki olay arasındaki sürelerin de bağımsız ve Üstel (λ) dağılımlı olduğu, bu kitabın kapsamı dışında kalan stokastik süreçler dersinde ispatlanır; bu yapıya **Poisson süreci** denir.

Örnek 23.2 (Üstel Ömür ve Bekleme Süresi Hesapları). (a) Elektronik bir parçanın ömrü X (saat), beklenen değeri 1000 saat olan üstel dağılıma sahiptir. Parçanın 2000 saatten uzun çalışma olasılığını, 500 saatten önce bozulma olasılığını, ömrün medyanını ve 500 saat çalışmış bir parçanın en az 1000 saat daha çalışma olasılığını bulunuz.

(b) Bir çağrı merkezine çağrılar, saatte ortalama 3 çağrı olacak biçimde, her t saatlik aralıktaki çağrı sayısı Poisson $(3t)$ olacak şekilde gelmektedir. İlk çağrı için yarım saatten uzun bekleme olasılığını, 10 dakikadan kısa bekleme olasılığını ve beklenen bekleme süresini bulunuz.

Çözüm.

(a) $E(X) = \frac{1}{\lambda} = 1000$ olduğundan $\lambda = \frac{1}{1000}$ ve $P(X > x) = e^{-x/1000}$.

$$P(X > 2000) = e^{-2} \approx 0,1353, \quad P(X < 500) = 1 - e^{-1/2} \approx 0,3935.$$

Medyan m , $F_X(m) = \frac{1}{2}$ denkleminin çözümüdür: $1 - e^{-m/1000} = \frac{1}{2}$, yani $e^{-m/1000} = \frac{1}{2}$ ve

$$m = 1000 \ln 2 \approx 693,1 \text{ saat}.$$

Medyan beklenen değerden küçüktür: ömürlerin yarısından fazlası ortalamanın altındadır, az sayıdaki uzun ömür ortalamayı yukarı çeker.

Son olasılık için hafızasızlık (Teorem 23.3) kullanılır:

$$P(X > 1500 \mid X > 500) = P(X > 1000) = e^{-1} \approx 0,3679.$$

Doğrudan hesapla da aynı sonuç çıkar: $\frac{e^{-1500/1000}}{e^{-500/1000}} = e^{-1}$.

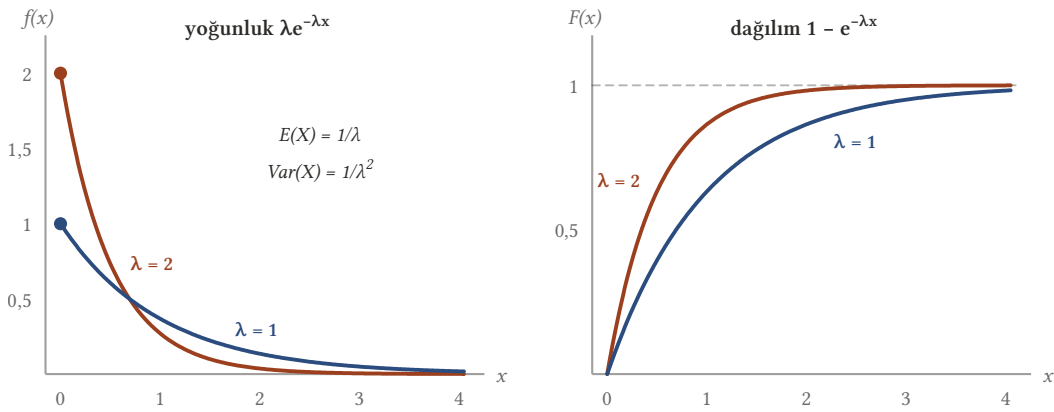
(b) Önerme 23.2 gereği ilk çağrıya kadar geçen süre T (saat), Üstel (3) dağılımlıdır.

$$\left(P_T > \frac{1}{2}\right) = e^{-3 \cdot 1/2} = e^{-1,5} \approx 0,2231.$$

Aynı sonuç Poisson tarafından da çıkar: $P(N_{1/2} = 0) = e^{-1,5}$. 10 dakika $\frac{1}{6}$ saattir:

$$\left(P_T < \frac{1}{6}\right) = 1 - e^{-3/6} = 1 - e^{-1/2} \approx 0,3935.$$

Beklenen bekleme süresi $E(T) = \frac{1}{3}$ saat, yani 20 dakikadır.



Şekil 23.2: $\lambda = 1$ ve $\lambda = 2$ için üstel dağılım. Yoğunluk sıfırda λ değerinden başlayıp geometrik hızla azalır; λ büyüdükçe kütle sıfıra yığılır, ortalama $1/\lambda$ küçülür. Dağılım fonksiyonu $1 - e^{-\lambda x}$ de aynı ölçüde daha hızlı 1'e yaklaşır.

■

23.3 Normal Dağılım

Normal dağılım, olasılık teorisinin merkezindeki dağılımdır. Gauss onu ölçüm hatalarının dağılımı olarak kullandı; bugün önemi, matematiksel istatistikte ispatlanacak olan merkezi limit teoreminden gelir: çok sayıda bağımsız, küçük etkinin toplamı, tek tek etkilerin dağılımı ne olursa olsun, yaklaşık olarak normal dağılır. Bu bölümde teoremi değil, dağılımın kendisini kuracağız.

Tanım 23.3 (Normal (Gauss) Dağılımı). $\mu \in \mathbb{R}$ ve $\sigma > 0$ olsun. Olasılık yoğunluk fonksiyonu

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad x \in \mathbb{R}$$

olan sürekli rastgele değişkene μ ve σ^2 parametrelili **normal dağılım**a (Gauss dağılımına) sahiptir denir ve $X \sim N(\mu, \sigma^2)$ yazılır. $N(0, 1)$ dağılımına **standart normal dağılım** denir; yoğunluğu ve dağılım fonksiyonu

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad \Phi(z) = \int_{-\infty}^z \varphi(u) du$$

ile gösterilir.

Yoğunluğun biçimine bakalım. f_X her yerde pozitiftir, $x = \mu$ noktasına göre simetriktir ve en büyük değerini $\frac{1}{\sigma\sqrt{2\pi}}$ olarak $x = \mu$ 'de alır. Türevleri

$$f'_X(x) = -\frac{x-\mu}{\sigma^2} f_X(x), \quad f''_X(x) = \left(\frac{(x-\mu)^2}{\sigma^4} - \frac{1}{\sigma^2} \right) f_X(x)$$

olduğundan büküm noktaları $x = \mu \pm \sigma$ 'dadır. σ küçüldükçe çan daralır ve yükselir, büyüdüğü genişler ve basıklaşır; μ değişince eğri yalnızca yatay kayar. Tanımda σ^2 yazılmasının nedeni, ikinci parametrenin varyans olmasıdır.

Bu fonksiyonun bir yoğunluk olduğunu göstermek bir satırlık iş değildir; e^{-x^2} 'nin ilkel fonksiyonu elemanter fonksiyonlarla yazılamaz. İntegralin değeri, iki boyutlu integrale geçip kutupsal koordinat kullanan klasik bir hileyle bulunur.

Teorem 23.4 (Normal Yoğunluğun İntegrali Birdir). Standart normal yoğunluğun normalleştirme sabitini veren eşitlik şudur:

$$\int_{-\infty}^{\infty} e^{-z^2/2} dz = \sqrt{2\pi}.$$

Dolayısıyla her $\mu \in \mathbb{R}$ ve $\sigma > 0$ için Tanım 23.3'deki f_X bir olasılık yoğunluk fonksiyonudur.

İspat.

Adım 1: integral yakınsar. $[-1, 1]$ üzerinde integrand sürekli olduğundan oradaki integral sonludur. $|z| \geq 1$ için $z^2 \geq |z|$, dolayısıyla $e^{-z^2/2} \leq e^{-|z|/2}$ 'dir; $\int_1^{\infty} e^{-z/2} dz = 2e^{-1/2}$ sonlu olduğundan karşılaştırmayla sağ kuyruk yakınsar, integrand çift olduğundan sol kuyruk da aynı değere yakınsar. Böylece $\int_{-\infty}^{\infty} e^{-z^2/2} dz$ yakınsar. Değerine I diyelim; integrand pozitif olduğundan $I > 0$.

Adım 2: I^2 'yi iki katlı integral olarak yazma. Değişken adının önemi olmadığından

$$I^2 = \left(\int_{-\infty}^{\infty} e^{-x^2/2} dx \right) \left(\int_{-\infty}^{\infty} e^{-y^2/2} dy \right) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)/2} dx dy;$$

integrand pozitif ve iki ayrı integralin çarpımı sonlu olduğundan iki katlı integral düzlemin tamamı üzerinde yakınsar ve ardışık integral olarak hesaplanabilir.

Adım 3: kutupsal koordinatlar. $x = r \cos \theta$, $y = r \sin \theta$ ($r > 0$, $0 \leq \theta < 2\pi$) dönüşümünün Jacobian'ı r 'dir ve $x^2 + y^2 = r^2$ olur:

$$I^2 = \int_0^{2\pi} \int_0^{\infty} e^{-r^2/2} r dr d\theta.$$

İç integralde $u = \frac{r^2}{2}$, $du = r dr$ alınırsa

$$\int_0^{\infty} e^{-r^2/2} r dr = \int_0^{\infty} e^{-u} du = 1.$$

Öyleyse $I^2 = \int_0^{2\pi} 1 d\theta = 2\pi$ ve $I > 0$ olduğundan $I = \sqrt{2\pi}$.

Adım 4: genel μ , σ . $f_X \geq 0$ açıktır. $z = \frac{x-\mu}{\sigma}$ değişken değişimiyle $dx = \sigma dz$ ve

$$\int_{-\infty}^{\infty} f_X(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} \sigma dz = \frac{\sigma}{\sigma\sqrt{2\pi}} \cdot \sqrt{2\pi} = 1.$$

■

Teorem aynı zamanda φ 'nin bir yoğunluk ve Φ 'nin bir dağılım fonksiyonu olduğunu söyler. Φ sürekli ve kesin artandır, $\Phi(0) = \frac{1}{2}$ 'dir; ama kapalı bir formülü yoktur, değerleri tablodan okunur.

Şimdi parametrelerin anlamını belirleyelim. Önce moment üreten fonksiyonu hesaplamak en kestirme yoldur; beklenen değeri, varyansı ve dağılımın öteki yapısal özelliklerini ondan okuyacağız.

Teorem 23.5 (Normal Dağılımın Moment Üreten Fonksiyonu, Beklenen Değeri ve Varyansı).

$X \sim N(\mu, \sigma^2)$ olsun. Her $t \in \mathbb{R}$ için

$$M_X(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)$$

ve

$$E(X) = \mu, \quad \text{Var}(X) = \sigma^2.$$

İspat.

$t \in \mathbb{R}$ sabit olsun. Tanımdan

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx.$$

$x = \mu + \sigma z$ değişken değişimiyle ($dx = \sigma dz$)

$$M_X(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{t(\mu+\sigma z)} e^{-z^2/2} dz = e^{\mu t} \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(\sigma t z - \frac{z^2}{2}\right) dz.$$

Üsteli kareye tamamlayalım:

$$\sigma t z - \frac{z^2}{2} = -\frac{1}{2}(z^2 - 2\sigma t z) = -\frac{1}{2}((z - \sigma t)^2 - \sigma^2 t^2) = -\frac{(z - \sigma t)^2}{2} + \frac{\sigma^2 t^2}{2}.$$

z 'ye bağlı olmayan $\frac{\sigma^2 t^2}{2}$ terimi integralin dışına çıkar:

$$M_X(t) = e^{\mu t + \sigma^2 t^2/2} \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(z - \sigma t)^2/2} dz.$$

Son integralde $w = z - \sigma t$ alınırsa $\int_{-\infty}^{\infty} e^{-w^2/2} dw = \sqrt{2\pi}$ olur (Teorem 23.4). Dolayısıyla

$$M_X(t) = e^{\mu t + \sigma^2 t^2/2}.$$

İntegral her t için yakınsadığından M_X bütün $\mathbb{R}$ 'de sonludur ve Teorem 17.4 uygulanır. Türevler:

$$M'_X(t) = (\mu + \sigma^2 t) M_X(t), \quad M''_X(t) = \sigma^2 M_X(t) + (\mu + \sigma^2 t)^2 M_X(t).$$

$M_X(0) = 1$ olduğundan $E(X) = M'_X(0) = \mu$ ve $E(X^2) = M''_X(0) = \sigma^2 + \mu^2$; Önerme 16.3 ile

$$\text{Var}(X) = \sigma^2 + \mu^2 - \mu^2 = \sigma^2.$$

■

Demek ki $N(\mu, \sigma^2)$ gösterimindeki iki parametre tam olarak beklenen değer ve varyanstır. Moment üreten fonksiyonun üstelinin t 'de ikinci dereceden bir polinom olması normal dağılımın ayırt edici özelliğidir; aşağıdaki sonuçların hepsi bu biçimin korunmasına dayanır.

23.3.1 Standartlaştırma

Normal dağılım hesaplarında tek bir tablo yeter, çünkü her normal rastgele değişken doğrusal bir dönüşümle standart normale çevrilebilir.

Teorem 23.6 (Standartlaştırma). $X \sim N(\mu, \sigma^2)$ ise

$$Z = \frac{X - \mu}{\sigma} \sim N(0, 1).$$

Tersine, $Z \sim N(0, 1)$ ise $X = \mu + \sigma Z \sim N(\mu, \sigma^2)$. Sonuç olarak her $x \in \mathbb{R}$ ve $a < b$ için

$$F_X(x) = \Phi\left(\frac{x - \mu}{\sigma}\right), \quad P(a < X \leq b) = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right)$$

ve her $z \in \mathbb{R}$ için $\Phi(-z) = 1 - \Phi(z)$.

İspat.

$g(x) = \frac{x-\mu}{\sigma}$ fonksiyonu $\mathbb{R}$ üzerinde kesin artan ve türevlenebilirdir; tersi $h(z) = \mu + \sigma z$, türevi $h'(z) = \sigma > 0$ 'dır. Dönüşüm formülü (Teorem 12.2) ile $Z = g(X)$ 'in yoğunluğu

$$f_Z(z) = f_X(h(z)) |h'(z)| = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\mu + \sigma z - \mu)^2}{2\sigma^2}\right) \sigma = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} = \varphi(z).$$

Demek ki $Z \sim N(0, 1)$. Ters yön aynı formülle: $g(z) = \mu + \sigma z$ kesin artan, tersi $h(x) = \frac{x-\mu}{\sigma}$, $h'(x) = \frac{1}{\sigma}$; $X = g(Z)$ 'nin yoğunluğu

$$f_X(x) = \left(\frac{x-\mu}{\sigma}\right) \cdot \frac{1}{\sigma} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),$$

yani $X \sim N(\mu, \sigma^2)$.

Dağılım fonksiyonu için, g kesin artan olduğundan $(X \leq x) = (Z \leq \frac{x-\mu}{\sigma})$ ve

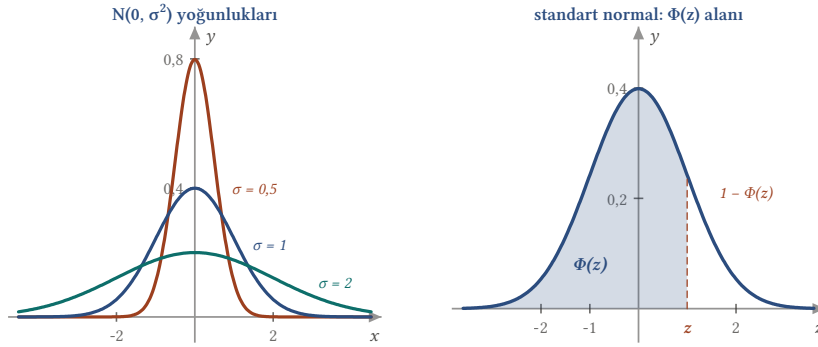
$$F_X(x) = \left(PZ \leq \frac{x-\mu}{\sigma}\right) = \left(\Phi\left(\frac{x-\mu}{\sigma}\right)\right).$$

$P(a < X \leq b) = F_X(b) - F_X(a)$ eşitliği bunu izler; X sürekli olduğundan aralığın uçlarının dâhil olup olmaması önemsizdir.

Simetri özelliği için $\Phi(-z)$ integralinde $u = -v$ değişimini yapalım:

$$\Phi(-z) = \int_{-\infty}^{-z} \varphi(u) du = \int_z^{\infty} \varphi(-v) dv = \int_z^{\infty} \varphi(v) dv = 1 - \Phi(z);$$

burada $\varphi(-v) = \varphi(v)$ (çift fonksiyon) ve $\int_{-\infty}^{\infty} \varphi = 1$ kullanıldı.



Şekil 23.3: Solda sıfır ortalamalı normal yoğunlukları: σ küçüldükçe eğri dikleşir ve kütle ortalamaya toplanır; her eğrinin altındaki alan 1'dir. Sağda standart normal dağılım fonksiyonu $\Phi(z) = P(Z \leq z)$, z 'nin solunda kalan alan olarak; simetri gereği $\Phi(-z) = 1 - \Phi(z)$.

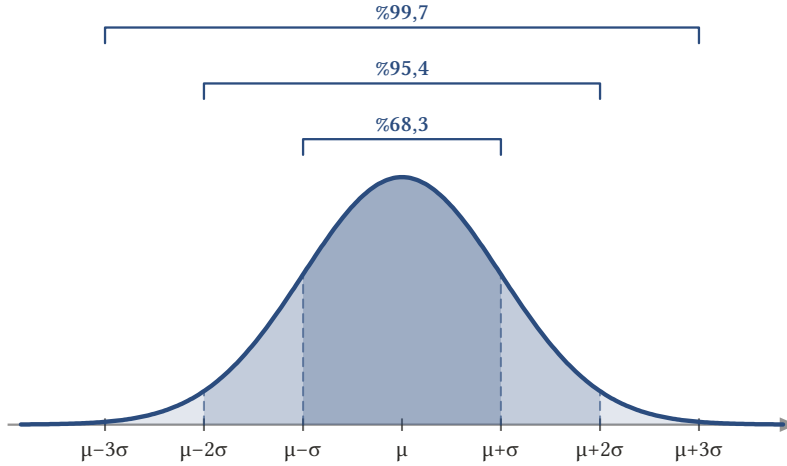
■

💡 Bir, iki, üç standart sapma kuralı

$X \sim N(\mu, \sigma^2)$ için Teorem 23.6 ve $\Phi(-z) = 1 - \Phi(z)$ ile

$$P(|X - \mu| < k\sigma) = P(-k < Z < k) = \Phi(k) - \Phi(-k) = 2\Phi(k) - 1.$$

$k = 1, 2, 3$ için bu olasılıklar yaklaşık 0,6827, 0,9545 ve 0,9973'tür: normal dağılımlı bir büyüklük, ortalamasının bir standart sapma yakınına yaklaşık üçte iki, iki standart sapma yakınına yüzde 95, üç standart sapma yakınına yüzde 99,7 olasılıkla düşer. Chebyshev eşitsizliğinin (Teorem 16.4) $k = 2$ için verdiği $P(|X - \mu| \geq 2\sigma) \leq \frac{1}{4}$ sınırıyla karşılaştırınız: gerçek değer 0,0455'tir.



Şekil 23.4: Bir, iki, üç standart sapma kuralı. $N(\mu, \sigma^2)$ yoğunluğunun altında iç içe üç bant: ortalamadan bir standart sapma yakını yüzde 68,3, iki standart sapma yakını yüzde 95,4, üç standart sapma yakını yüzde 99,7 olasılık taşır. Kuyruklarda kalan toplam olasılık binde üçtür; normal dağılımlı bir büyüklük ortalamasından üç σ uzağa hemen hiç düşmez.

Standartlaştırma, $aX + b$ biçiminde bir dönüşümdür. Genel doğrusal dönüşümün de normal dağılımı koruduğunu bu kez moment üreten fonksiyonla gösterelim; aynı teknik bir sonraki teoremden toplamlar için de işe yarayacak.

Önerme 23.3 (Normal Dağılımın Doğrusal Dönüşümü). $X \sim N(\mu, \sigma^2)$, $a \neq 0$ ve $b \in \mathbb{R}$ ise

$$aX + b \sim N(a\mu + b, a^2\sigma^2).$$

İspat.

$Y = aX + b$ olsun. Her $t \in \mathbb{R}$ için

$$M_Y(t) = E(e^{t(aX+b)}) = e^{bt} E(e^{(at)X}) = e^{bt} M_X(at) = e^{bt} \exp\left(\mu at + \frac{\sigma^2 a^2 t^2}{2}\right) = \exp\left((a\mu + b)t + \frac{a^2 \sigma^2 t^2}{2}\right).$$

Bu, Teorem 23.5 gereği $N(a\mu + b, a^2\sigma^2)$ dağılımının moment üreten fonksiyonudur ve her t için sonludur. Moment üreten fonksiyon dağılımı tek türlü belirlediğinden (Teorem 17.5) $Y \sim N(a\mu + b, a^2\sigma^2)$.

■

Parametreler Teorem 16.2 ile uyumludur; önermenin yeni söylediği, dağılımın **türünün** korunmasıdır. $a = \frac{1}{\sigma}$, $b = -\frac{\mu}{\sigma}$ alınırsa $N(0, 1)$ bulunur; standartlaştırma teoremi böylece ikinci bir ispat kazanır.

23.3.2 Bağımsız Normallerin Toplamı

Bağımsız normal rastgele değişkenlerin toplamı yine normaldir. Bunu konvolüsyonla (Önerme 15.3) göstermek uzun bir kareye tamamlama gerektirir; moment üreten fonksiyon aynı sonucu iki satırda verir.

Teorem 23.7 (Bağımsız Normal Rastgele Değişkenlerin Toplamı). $X_1, \dots, X_n$ bağımsız ve her i için $X_i \sim N(\mu_i, \sigma_i^2)$ olsun. O zaman

$$S = X_1 + \dots + X_n \sim N(\mu_1 + \dots + \mu_n, \sigma_1^2 + \dots + \sigma_n^2).$$

Özel olarak, $X_1, \dots, X_n$ bağımsız ve hepsi $N(\mu, \sigma^2)$ dağılımlı ise

$$S \sim N(n\mu, n\sigma^2) \quad \text{ve} \quad \bar{X} = \frac{S}{n} \sim N\left(\mu, \frac{\sigma^2}{n}\right).$$

İspat.

$X_1, \dots, X_n$ bağımsız olduğundan (Tanım 14.1) $e^{tX_1}, \dots, e^{tX_n}$ de bağımsızdır; bağımsız rastgele değişkenlerin çarpımının beklenen değeri beklenen değerlerin çarpımıdır (Teorem 18.1; n çarpan için, $e^{t(X_1 + \dots + X_{n-1})}$ ile e^{tX_n} bağımsız olduğundan tümevarımla). Dolayısıyla her $t \in \mathbb{R}$ için

$$M_S(t) = E(e^{t(X_1 + \dots + X_n)}) = \left(E \prod_{i=1}^n e^{tX_i}\right) = \prod_{i=1}^n E(e^{tX_i}) = \prod_{i=1}^n \exp\left(\mu_i t + \frac{\sigma_i^2 t^2}{2}\right) = \exp\left(\left(\sum_{i=1}^n \mu_i\right)t + \left(\sum_{i=1}^n \sigma_i^2\right)\frac{t^2}{2}\right).$$

Bu, $(\sum \mu_i, \sum \sigma_i^2)$ dağılımının moment üreten fonksiyonudur; Teorem 17.5 gereği S bu dağılıma sahiptir. Özdeş dağılım durumunda toplamlar $n\mu$ ve $n\sigma^2$ olur. $\bar{X} = \frac{1}{n}S$ için Önerme 23.3'u $a = \frac{1}{n}$, $b = 0$ ile uygularız:

$$\bar{X} \sim N\left(\frac{n\mu}{n}, \frac{n\sigma^2}{n^2}\right) = N\left(\mu, \frac{\sigma^2}{n}\right).$$

■

Varyansların toplanması Sonuç 18.2'nin söylediğinden fazlası değildir; teoremin katkısı yine dağılım türünün korunmasıdır. $\bar{X}$ için bulunan sonuç matematiksel istatistiğin çıkış noktasıdır: n bağımsız ölçümün ortalaması, tek ölçümle aynı merkezde ama $\sqrt{n}$ kat küçük standart sapmayla dağılır.

⚠ Bağımsızlık varsayımı atılamaz

$X \sim N(0, 1)$ ve ε , X 'ten bağımsız, $P(\varepsilon = 1) = P(\varepsilon = -1) = \frac{1}{2}$ olsun; $Y = \varepsilon X$ alalım. Toplam olasılık formülü ve $-X \sim N(0, 1)$ (Önerme 23.3, $a = -1$) ile

$$P(Y \leq y) = \frac{1}{2}P(X \leq y) + \frac{1}{2}P(-X \leq y) = \Phi(y),$$

yani $Y \sim N(0, 1)$. Oysa $X + Y = (1 + \varepsilon)X$, $\varepsilon = -1$ iken tam olarak 0'dır: $P(X + Y = 0) = \frac{1}{2} > 0$. $X + Y$ sürekli bile değildir, normal hiç değildir.

23.3.3 Standart Normal Tablo Hesapları

Φ 'nin kapalı formülü olmadığından değerleri tablodan okunur. Tablolar $z \geq 0$ için verilir; negatif z için $\Phi(-z) = 1 - \Phi(z)$ kullanılır. Çözüm şeması her zaman aynıdır: istenen olay X cinsinden yazılır, Teorem 23.6 ile $Z = \frac{X - \mu}{\sigma}$ cinsine çevrilir, gerekirse simetriyle pozitif z 'lere indirgenir ve tablodan okunur.

Örnek 23.3 (Standart Normal Tablo Hesapları). Aşağıdaki tablo, $\Phi(z)$ değerlerini dört ondalıkla verir; satır z 'nin ilk basamaklarını, sütun eklenecek miktarı gösterir (örneğin $\Phi(1, 25)$, satır 1, 2 ile sütun +0, 05'in kesişimidir).

z	+0, 00	+0, 05	+0, 10	+0, 15
0, 0	0, 5000	0, 5199	0, 5398	0, 5596
0, 2	0, 5793	0, 5987	0, 6179	0, 6368

z	+0,00	+0,05	+0,10	+0,15
0,4	0,6554	0,6736	0,6915	0,7088
0,6	0,7257	0,7422	0,7580	0,7734
0,8	0,7881	0,8023	0,8159	0,8289
1,0	0,8413	0,8531	0,8643	0,8749
1,2	0,8849	0,8944	0,9032	0,9115
1,4	0,9192	0,9265	0,9332	0,9394
1,6	0,9452	0,9505	0,9554	0,9599
1,8	0,9641	0,9678	0,9713	0,9744
2,0	0,9772	0,9798	0,9821	0,9842
2,2	0,9861	0,9878	0,9893	0,9906
2,4	0,9918	0,9929	0,9938	0,9946
2,6	0,9953	0,9960	0,9965	0,9970
2,8	0,9974	0,9978	0,9981	0,9984
3,0	0,9987	0,9989	0,9990	0,9992

Bir sınavdaki puanlar $X \sim N(100, 15^2)$ dağılımlı olsun.

(a) $P(X \leq 115)$, $P(X > 130)$ ve $P(85 < X < 122, 5)$ olasılıklarını bulunuz.

(b) $P(X \leq c) = 0,9332$ olan c puanını ve puanların yalnızca yüzde 5'inin aştığı d puanını bulunuz ($\Phi(1,6449) = 0,9500$ verilmiştir).

(c) Bir üründe iki parçanın uzunlukları bağımsız ve $Y_1 \sim N(50, 3^2)$, $Y_2 \sim N(30, 4^2)$ (mm) dağılımlıdır. Toplam uzunluğun 90 mm'yi aşma olasılığını bulunuz.

(d) Bir büyüklük, beklenen değeri μ ve varyansı 16 olan normal dağılımlı bir aletle 16 kez bağımsız olarak ölçülüyor. Ölçümlerin ortalamasının μ 'ye 1 birimden yakın olma olasılığını bulunuz.

Çözüm.

(a) $\mu = 100$, $\sigma = 15$ ile standartlaştırırız.

$$P(X \leq 115) = \left(\Phi \frac{115 - 100}{15} \right) = \Phi(1) = 0,8413.$$

$$P(X > 130) = 1 - \left(\Phi \frac{130 - 100}{15} \right) = 1 - \Phi(2) = 1 - 0,9772 = 0,0228.$$

$$P(85 < X < 122,5) = \left(\Phi \frac{122,5 - 100}{15} \right) - \left(\Phi \frac{85 - 100}{15} \right) = \Phi(1,5) - \Phi(-1).$$

$\Phi(-1) = 1 - \Phi(1) = 1 - 0,8413 = 0,1587$ olduğundan

$$P(85 < X < 122,5) = 0,9332 - 0,1587 = 0,7745.$$

(b) $P(X \leq c) = \left(\Phi \frac{c - 100}{15} \right) = 0,9332$ olmalıdır. Tabloda $0,9332 = \Phi(1,5)$; Φ kesin artan olduğundan $\frac{c - 100}{15} = 1,5$ ve $c = 100 + 15 \cdot 1,5 = 122,5$.

$P(X > d) = 0,05$, yani $P(X \leq d) = 0,95$ istenir: $\frac{d - 100}{15} = 1,6449$ ve

$$d = 100 + 15 \cdot 1,6449 = 124,67.$$

Puanların yüzde 5'i 124,67'nin üstündedir. (Tabloda $\Phi(1,60) = 0,9452$ ve $\Phi(1,65) = 0,9505$; 0,95 bu ikisinin arasındadır.)

(c) Teorem 23.7 gereği $T = Y_1 + Y_2 \sim N(50 + 30,9 + 16) = N(80,25)$; $\sigma_T = 5$.

$$P(T > 90) = 1 - \Phi\left(\frac{90 - 80}{5}\right) = 1 - \Phi(2) = 0,0228.$$

(d) $\sigma^2 = 16$, $n = 16$; Teorem 23.7 gereği $\bar{X} \sim \left(\mu, \frac{16}{16}\right) = N(\mu, 1)$. Öyleyse

$$P(|\bar{X} - \mu| < 1) = P\left(-1 < \frac{\bar{X} - \mu}{1} < 1\right) = \Phi(1) - \Phi(-1) = 2\Phi(1) - 1 = 2 \cdot 0,8413 - 1 = 0,6826$$

(daha hassas değeri 0,6827'dir). Tek ölçümde aynı olasılık $2\Phi(0,25) - 1 = 0,1974$ olurdu; on altı ölçüm standart sapmayı 4 kat küçültür.

■

23.4 Alıştırmalar

Alıştırma 23.1 (Düzgün, Üstel ve Normal Dağılımlar Üzerine). a) $X \sim U(-1, 3)$ olsun. $E(X)$, $\text{Var}(X)$, $P(|X| < 1)$ ve $P(X > 2 | X > 0)$ değerlerini bulunuz.

b) $U \sim U(0, 1)$ ve $\lambda > 0$ olsun. $Y = -\frac{1}{\lambda} \ln(1 - U)$ rastgele değişkeninin Üstel (λ) dağılımlı olduğunu iki yolla gösteriniz: ters dönüşüm yöntemiyle ve dönüşüm formülüyle. $-\frac{1}{\lambda} \ln U$ 'nun da aynı dağılıma sahip olduğunu açıklayınız.

c) $X_1 \sim \text{Üstel}(\lambda_1)$ ve $X_2 \sim \text{Üstel}(\lambda_2)$ bağımsız olsun. $\min\{X_1, X_2\} \sim \text{Üstel}(\lambda_1 + \lambda_2)$ olduğunu ve $P(X_1 < X_2) = \frac{\lambda_1}{\lambda_1 + \lambda_2}$ olduğunu gösteriniz.

d) $X \sim \text{Üstel}(\lambda)$ ve $0 < p < 1$ olsun. $P(X \leq x_p) = p$ olan x_p sayısını (p -kantili) bulunuz. Beklenen ömrü 1000 saat olan parça için ömürlerin yüzde 10'unun altında kaldığı ve yüzde 10'unun üstünde kaldığı süreleri hesaplayınız.

e) $X \sim N(\mu, \sigma^2)$ için $P(X < 10) = 0,1587$ ve $P(X < 20) = 0,9332$ olduğu biliniyor. μ ve σ 'yı bulunuz; ardından $P(X > 18)$ 'i hesaplayınız.

f) $X \sim N(\mu, \sigma^2)$ için $P(|X - \mu| \geq 3\sigma)$ olasılığını Chebyshev eşitsizliğinden gelen üst sınırla karşılaştırınız.

g) $Z \sim N(0, 1)$ olsun. Moment üreten fonksiyonun seri açılımından $E(Z^3) = 0$ ve $E(Z^4) = 3$ olduğunu gösteriniz. $X \sim N(\mu, \sigma^2)$ için $E(e^X)$ 'i bulunuz.

Çözüm.

a) Teorem 23.1 ile $E(X) = \frac{-1+3}{2} = 1$ ve $\text{Var}(X) = \frac{4^2}{12} = \frac{4}{3}$. $(|X| < 1) = (-1 < X < 1)$ olayının uzunluğu 2, aralığın uzunluğu 4: $P(|X| < 1) = \frac{2}{4} = \frac{1}{2}$. Koşullu olasılık:

$$P(X > 2 | X > 0) = \frac{P(X > 2)}{P(X > 0)} = \frac{1/4}{3/4} = \frac{1}{3}.$$

b) **Ters dönüşüm.** Üstel (λ) dağılım fonksiyonu $F(x) = 1 - e^{-\lambda x}$ ($x > 0$), $(0, \infty)$ üzerinde sürekli ve kesin artandır. $F(x) = u$ denklemi $e^{-\lambda x} = 1 - u$, yani $x = -\frac{1}{\lambda} \ln(1 - u)$ verir; $F^{-1}(u) = -\frac{1}{\lambda} \ln(1 - u)$. Önerme 23.1 gereği $Y = F^{-1}(U) \sim \text{Üstel}(\lambda)$.

Dönüşüm formülü. $g(u) = -\frac{1}{\lambda} \ln(1 - u)$, $(0, 1)$ üzerinde kesin artan ve türevlenebilirdir; görüntüsü $(0, \infty)$, tersi $h(y) = 1 - e^{-\lambda y}$, $h'(y) = \lambda e^{-\lambda y}$. Teorem 12.2 ile $y > 0$ için

$$f_Y(y) = f_U(h(y)) |h'(y)| = 1 \cdot \lambda e^{-\lambda y},$$

ki bu Üstel (λ) yoğunluğudur.

$0 < u < 1$ için $P(1 - U \leq u) = P(U \geq 1 - u) = u$ olduğundan $1 - U \sim U(0, 1)$; dolayısıyla $-\frac{1}{\lambda} \ln U$ da Üstel (λ) dağılımlıdır.

c) $M = \min\{X_1, X_2\}$ olsun. $t > 0$ için $(M > t) = (X_1 > t) \cap (X_2 > t)$ ve bağımsızlıkla

$$P(M > t) = P(X_1 > t) P(X_2 > t) = e^{-\lambda_1 t} e^{-\lambda_2 t} = e^{-(\lambda_1 + \lambda_2)t}.$$

$M > 0$ olduğundan $F_M(t) = 1 - e^{-(\lambda_1 + \lambda_2)t}$ ($t > 0$), $F_M(t) = 0$ ($t \leq 0$): $M \sim \text{Üstel}(\lambda_1 + \lambda_2)$.

İkinci olasılık için bağımsızlık nedeniyle ortak yoğunluk çarpımdır: $f(x_1, x_2) = \lambda_1 e^{-\lambda_1 x_1} \lambda_2 e^{-\lambda_2 x_2}$ ($x_1, x_2 > 0$). $\{x_1 < x_2\}$ bölgesi üzerinde integral alalım; önce x_2 üzerinden:

$$P(X_1 < X_2) = \int_0^\infty \lambda_1 e^{-\lambda_1 x_1} \left(\int_{x_1}^\infty \lambda_2 e^{-\lambda_2 x_2} dx_2 \right) dx_1 = \int_0^\infty \lambda_1 e^{-\lambda_1 x_1} e^{-\lambda_2 x_1} dx_1 = \lambda_1 \int_0^\infty e^{-(\lambda_1 + \lambda_2)x_1} dx_1 = \frac{\lambda_1}{\lambda_1 + \lambda_2}.$$

d) $1 - e^{-\lambda x_p} = p$ denkleminde $e^{-\lambda x_p} = 1 - p$ ve

$$x_p = -\frac{\ln(1-p)}{\lambda} = \frac{1}{\lambda} \ln \frac{1}{1-p}.$$

$p = \frac{1}{2}$ için medyan $\frac{\ln 2}{\lambda}$ bulunur (Örnek 23.2 ile uyumlu). $\lambda = \frac{1}{1000}$ için:

$$x_{0,1} = -1000 \ln 0,9 \approx 105,4 \text{ saat}, \quad x_{0,9} = -1000 \ln 0,1 \approx 2302,6 \text{ saat}.$$

Ömürlerin yüzde 10'u 105,4 saatin altında, yüzde 10'u 2302,6 saatin üstündedir.

e) Standartlaştırma ile $\left(\frac{10-\mu}{\sigma}\right) = 0,1587$ ve $\left(\frac{20-\mu}{\sigma}\right) = 0,9332$. Tablodan $\Phi(1) = 0,8413$, dolayısıyla $\Phi(-1) = 0,1587$; ve $\Phi(1,5) = 0,9332$. Φ kesin artan olduğundan

$$\frac{10-\mu}{\sigma} = -1, \quad \frac{20-\mu}{\sigma} = 1,5.$$

İkinci denklemden birinciyi çıkarınca $\frac{10}{\sigma} = 2,5$, yani $\sigma = 4$; birinci denklemden $\mu = 10 + \sigma = 14$. Sağlama: $\frac{20-14}{4} = 1,5$. O hâlde $X \sim N(14, 16)$ ve

$$P(X > 18) = 1 - \left(\Phi\left(\frac{18-14}{4}\right)\right) = 1 - \Phi(1) = 0,1587.$$

f) Teorem 16.4 ile $P(|X - \mu| \geq 3\sigma) \leq \frac{1}{9} \approx 0,1111$. Gerçek değer, standartlaştırma ve simetriyle,

$$P(|X - \mu| \geq 3\sigma) = P(|Z| \geq 3) = 2(1 - \Phi(3)) = 2(1 - 0,9987) = 0,0026$$

(daha hassas değeri 0,0027'dir). Chebyshev sınırı gerçek olasılığın kırk katından büyüktür; her dağılım için geçerli olmanın bedeli budur.

g) $M_Z(t) = e^{t^2/2}$ (Teorem 23.5, $\mu = 0, \sigma = 1$). Üstel serisiyle

$$M_Z(t) = \sum_{k=0}^{\infty} \frac{(t^2/2)^k}{k!} = 1 + \frac{t^2}{2} + \frac{t^4}{8} + \dots$$

Seride yalnızca çift kuvvetler bulunur; $\frac{t^3}{3!}$ 'in katsayısı 0 olduğundan Teorem 17.4 gereği $E(Z^3) = 0$. $\frac{t^4}{4!}$ 'in katsayısı $E(Z^4)$ 'tür: $\frac{t^4}{8} = E(Z^4)$ eşitliğinden $E(Z^4) = \frac{24}{8} = 3$. Son olarak $E(e^X)$, moment üreten fonksiyonun $t = 1$ 'deki değeridir:

$$E(e^X) = M_X(1) = e^{\mu + \sigma^2/2};$$

örneğin $X \sim N(0, 1)$ için $E(e^X) = e^{1/2} \approx 1,6487$.

■

23.5 Olasılık Teorisi'nin Sonu

Bu bölümle kitap tamamlanıyor. Geriye dönüp bakınca tek bir iplik görünür.

Bir rastgele deneyin bütün olası sonuçlarını toplayan **örnek uzay** Ω ile başladık. Olasılık atayacağımız olayları keyfi bir aile olarak değil, tümleyen ve sayılabilir birleşim altında kapalı bir **σ -cebiri** $\mathcal{U}$ olarak seçtik; sayılabilir işlemler altında kapalılık, sonsuz deneyler ve limitler hakkında konuşabilmenin bedeliydi. Bu aile üzerinde Kolmogorov aksiyomlarını sağlayan bir **olasılık ölçüsü** P tanımladık ve $(\Omega, \mathcal{U}, P)$ üçlüsüne olasılık uzayı dedik. Laplace uzayı, sayma teknikleri ve geometrik olasılık bu yapının ilk örnekleri oldu; koşullu olasılık, Bayes teoremi ve bağımsızlık, bir olay hakkındaki bilginin ötekilerin olasılığını nasıl değiştirdiğini anlattı.

İkinci adım, Ω 'yı sayı doğrusuna taşıyan **rastgele değişken** kavramıydı: $(X \leq a)$ kümelerinin olay olması koşuluyla $X : \Omega \rightarrow \mathbb{R}$. Bu koşul, P ölçüsünü Borel cebirine taşıyıp X 'in **dağılımını** üretti; dağılım fonksiyonu F_X bu dağılımı tek bir azalmayan, sağdan sürekli fonksiyonda özetledi. Kesikli değişkenlerde olasılık fonksiyonu, sürekli değişkenlerde yoğunluk fonksiyonu F_X 'i yeniden kurdu; dönüşüm formülü bir

değişkenin fonksiyonunun dağılımını verdi. Rastgele vektörler aynı yapıyı $\mathbb{R}^n$ 'e genişletti: ortak, marjinal ve koşullu dağılımlar, bağımsızlık ve konvolüsyon.

Üçüncü adım **beklenen değer** oldu: dağılımın tek bir sayıya sıkıştırılmış özeti. Varyans, kovaryans ve korelasyon yayılımı ve birlikte hareketi ölçtü; Chebyshev eşitsizliği varyanstan olasılık sınırı çıkardı. Moment üreten fonksiyon ve karakteristik fonksiyon bütün momentleri tek bir fonksiyonda topladı ve teklik teoremi sayesinde dağılımları tanımanın en güçlü aracı oldu. Koşullu beklenen değer, bilginin beklentiye nasıl güncellediğini söyledi.

Son kısımda bu araçları **dağılım ailelerine** uyguladık. Bernoulli denemelerinden binom, geometrik ve negatif binom; nadir olayların limitinden Poisson; iadesiz çekilişten hipergeometrik dağılım doğdu. Bu bölümde sürekli tarafta düzgün, üstel ve normal dağılımları kurduk: her sürekli dağılım düzgünden üretilir, üstel dağılım geometrik dağılımın sürekli kardeşi ve Poisson dağılımının zaman eksenindeki karşılığıdır, normal dağılım ise doğrusal dönüşüm ve bağımsız toplam altında kapalı kalan çan eğrisidir. Her dağılım için aynı soruları sorduk ve her yanıt önceki kısımlardaki genel teoremlerin bir uygulaması oldu.

Buradan iki ders devam eder. **Matematiksel İstatistik** soruyu tersine çevirir: dağılımı bilip örneklemin davranışını hesaplamak yerine, örnekleme görüp dağılımın bilinmeyen parametrelerini tahmin etmek. Büyük sayılar yasası ve merkezi limit teoremi normal dağılımın neden her yerde ortaya çıktığını açıklar; örneklem ortalaması için bulduğumuz $(\bar{y}, \frac{\sigma^2}{n})$ sonucu güven aralıklarının ve hipotez testlerinin çıkış noktasıdır. **Stokastik Süreçler** ise zamanla değişen rastgele değişken ailelerini inceler: Poisson süreci, ilk olayın zamanı için gösterdiğimiz üstel dağılımın bütün olaylar dizisine genişletilmiş hâlidir; Markov zincirleri koşullu olasılık ve hafızasızlık fikrini zamana yayar; Brown hareketi normal dağılımlı, bağımsız artışı sürekli bir süreçtir. Bu kitapta kurulan olasılık uzayı, rastgele değişken ve dağılım dili her ikisinin de ortak temelidir.

Kitabın bütününe ve öteki bölümlere dönmek için: [Müfredat ve Giriş](#).