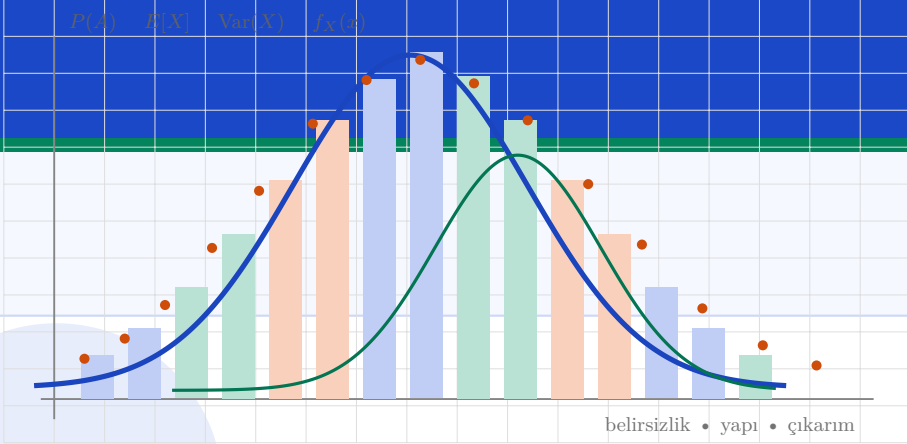


TEMELLER • DAĞILIMLAR • UYGULAMALI MODELLEME

OLASILIK TEORİSİ

Temeller, Dağılımlar ve Uygulamalı Modelleme



İbrahim GÜNEY

Istanbul Sabahattin Zaim University



2026

Olasılık Teorisi

Araştırma Yöntemleri Serisi

Temeller, Dağılımlar ve Uygulamalı Modelleme

Prof. Dr. İbrahim GÜNEY

Eser adı	Olasılık Teorisi
Seri	Araştırma Yöntemleri Serisi
Alt başlık	Temeller, Dağılımlar ve Uygulamalı Modelleme
Seri no	İstanbul Sabahattin Zaim Üniversitesi Yayınları: 112
ISBN	978-625-7558-54-9
DOI	10.5281/zenodo.22140995
Birinci baskı	Ağustos 2026
Yazar	Prof. Dr. İbrahim GÜNEY
Yayın koordinatörü	Mustafa GENÇ
Kapak tasarımı	Erhan YALUR
Yayıncı	İstanbul Sabahattin Zaim Üniversitesi, Halkalı Cad. No: 281, Küçükçekmece/İstanbul 34303
Baskı	1. Baskı



İstanbul, 2026

Önsöz

Bu kitap, lisans ve lisansüstü düzeyde ders kitabı veya yardımcı ders kitabı olarak kullanılmak üzere hazırlanmıştır. Araştırmacılar için ise olasılığın aksiyomatik temellerinden ileri olasılık konularına uzanan düzenli bir başvuru kaynağı sunmayı amaçlamaktadır.

Konular örnek uzay, olay ve olasılık aksiyomlarından başlayarak sayma yöntemleri, rassal değişkenler, kesikli ve sürekli dağılımlar, ortak dağılımlar, koşullu beklenti, örnekleme teorisi, yakınsama, limit teoremleri ve Monte Carlo yöntemleri doğrultusunda ilerlemektedir. Bu sıralama, temel kavramlar ile ileri sonuçlar arasındaki ilişkinin aşamalı biçimde kurulmasına olanak verir.

Kitapta matematiksel tanımlar ve teoremler, sezgisel açıklamalar ve uygulamalarla birlikte ele alınmıştır. Çözümlü örnekler, bölüm sonu problemleri ve farklı alanlardan seçilen uygulamalar, kuramsal bilginin kullanımını göstermeyi amaçlar. Python ve R kodları ise hesaplamaların yeniden üretilmesine, dağılımların incelenmesine ve simülasyon çalışmalarının yürütülmesine yardımcı olur.

Metin, olasılığı birbirinden bağımsız formüllerin toplamı olarak değil, belirsizlik altında düşünmeyi sağlayan bütünlüklü bir matematiksel yapı olarak sunmaktadır. Bu nedenle yalnızca işlemlerin nasıl yapıldığına değil, kullanılan modelin neden uygun olduğuna ve elde edilen sonuçların nasıl yorumlanması gerektiğine de yer verilmiştir.

Kitabın istatistik, veri bilimi, matematik, mühendislik, ekonomi, eğitim ve benzeri alanlarda çalışan öğrenci ve araştırmacılara yararlı olması; derslerde ve bireysel çalışmalarda güvenilir bir kaynak olarak kullanılması amaçlanmaktadır.

Prof. Dr. İbrahim Güney
İstanbul, 2026

Seri İçindeki Yeri

Bu kitap, *Olasılık ve İstatistik Teorisi* dizisinin ilk cildidir. Dizinin amacı, belirsizliğin matematiksel temellerinden istatistiksel çıkarım ve modellemeye uzanan bütünlüklü bir öğrenme yolu sunmaktır.

Birinci cilt	<i>Olasılık Teorisi</i> , olaylar ve olasılık aksiyomları, rassal değişkenler, dağılımlar, koşulluluk, bağımlılık, yakınsama, limit teoremleri ve simülasyon konularını ele alır.
İkinci cilt	<i>İstatistik Teorisi</i> , örnekleme, tahmin, olabilirlik, güven aralıkları, hipotez testleri, asimptotik çıkarım ve model değerlendirme konularına odaklanır.

Bu ciltte incelenen beklenen değer, varyans, ortak dağılımlar, koşullu beklenti, örnekleme dağılımları ve merkezi limit teoremi gibi kavramlar, istatistiksel çıkarımın matematiksel altyapısını oluşturur. İlk bölümler temel olasılık dilini, orta bölümler dağılım ve bağımlılık yapılarını, son bölümler ise yakınsama ve simülasyon yöntemlerini geliştirir.

Yazar Hakkında

Prof. Dr. İbrahim Güney, olasılık, istatistik ve uygulamalı modelleme alanlarında çalışan bir akademisyendir. Bu kitap, belirsizlik altında düşünme, matematiksel model kurma ve istatistiksel çıkarıma hazırlık konularında Türkçe kaynak ihtiyacına katkı sunmak amacıyla hazırlanmıştır.

Çalışmalarında matematiksel tanımlar, kuramsal sonuçlar, hesaplama yöntemleri ve uygulamalar arasındaki ilişkinin açık biçimde kurulmasına önem vermektedir. Olasılık ve istatistik konularının öğrenciler ve araştırmacılar için anlaşılır, düzenli ve yeniden üretilebilir bir yapıda sunulması temel akademik ilgi alanları arasındadır.

Teşekkür

Bu kitabın hazırlanmasında derslerde yöneltilen sorular, öğrencilerin konuya ilişkin değerlendirmeleri ve meslektaşların görüşleri yol gösterici olmuştur. Metnin daha açık, düzenli ve uygulamaya dönük bir yapıya kavuşmasına katkı sağlayan tüm öğrenci ve meslektaşlarıma teşekkür ederim.

Çalışma sürecinde desteğini ve anlayışını esirgemeyen aileme ayrıca teşekkür ederim.

Simgeler ve Kısaltmalar

Bu bölümde kitap boyunca kullanılan temel simgeler ve kısaltmalar toplu olarak verilmiştir.

Simge	Anlamı
Ω	Örnek uzay
ω	Örnek nokta
$\mathcal{F}$	Olaylar σ -cebiri
$P(A)$	A olayının olasılığı
A^c	A olayının tümleyeni
$A \cup B, A \cap B$	Olayların birleşimi ve kesişimi
$\emptyset$	Boş küme (imkânsız olay)
$P(A B)$	B bilindiğinde A 'nın koşullu olasılığı
X, Y, Z	Rassal değişkenler
$F_X(x)$	Kümülatif dağılım fonksiyonu (CDF)
$p_X(x)$	Olasılık kütle fonksiyonu (PMF)
$f_X(x)$	Olasılık yoğunluk fonksiyonu (PDF)
$F^{-1}(u)$	Kantil fonksiyonu
$F_{X,Y}, p_{X,Y}, f_{X,Y}$	Ortak dağılım, ortak PMF ve ortak PDF
$E[X], \mu$	Beklenen değer (ortalama)
$\text{Var}(X), \sigma^2$	Varyans
σ	Standart sapma
$\text{Cov}(X, Y)$	Kovaryans
$\text{Corr}(X, Y), \rho$	Korelasyon katsayısı
$E[X Y]$	Koşullu beklenti
$\text{Var}(X Y)$	Koşullu varyans
$I_A, \mathbf{1}\{\cdot\}$	Gösterge (indikatör) değişkeni
$G_X(s)$	Olasılık üreten fonksiyon (PGF)
$M_X(t)$	Moment üreten fonksiyon (MGF)
$\varphi_X(t)$	Karakteristik fonksiyon
$\bar{X}_n$	Örneklem ortalaması
S^2	Örneklem varyansı
$\hat{p}$	Örneklem oranı
$X_{(k)}$	k 'ncı sıralı istatistik
SE	Standart hata

Devamı sonraki sayfada

Simge	Anlamı
$\xrightarrow{p}, \xrightarrow{d}, \xrightarrow{a.s.}$	Olasılıkta, dağılımda, hemen hemen kesin yakınsama
$O_p(\cdot), o_p(\cdot)$	Stokastik merteye gösterimleri
$\Phi(z)$	Standart normal dağılım fonksiyonu
$\Gamma(\alpha)$	Gamma fonksiyonu
$n!$	Faktöriyel
$\binom{n}{r}$	Binom katsayısı
$\sim$	“Dağılımına sahiptir”

Kısaltma	Açılımı
CDF	Kümülatif dağılım fonksiyonu
PMF	Olasılık kütle fonksiyonu
PDF	Olasılık yoğunluk fonksiyonu
PGF	Olasılık üreten fonksiyon
MGF	Moment üreten fonksiyon
LOTUS	Bilinçsiz istatistikçi yasası
BSY	Büyük sayılar yasası
MLT	Merkezi limit teoremi
MSE	Ortalama karesel hata
MLE	En çok olabilirlik tahmin edicisi
IQR	Çeyrekler arası açıklık
CV	Değişim katsayısı
MC	Monte Carlo
i.i.d.	Bağımsız ve aynı dağılımlı
CI	Güven aralığı

İçindekiler

Önsöz	v
Seri İçindeki Yeri	vii
Yazar Hakkında	ix
Teşekkür	xi
Simgeler ve Kısaltmalar	xiii
1 Olasılığın Temelleri ve Aksiyomatik Yapı	1
1.1 Gerçek Dünya Problemi (Giriş)	1
1.2 Rassal Deney, Örnek Uzay ve Olay	2
1.3 Olay Cebiri ve Küme İşlemleri	2
1.4 Olasılık Ölçüsü ve Kolmogorov Aksiyomları	4
1.5 Sonlu ve Eş Olasılıklı Modeller	8
1.6 Koşullu Olasılık	9
1.7 Toplam Olasılık Yasası ve Bayes Teoremi	11
1.8 Bağımsızlık	13
1.9 Ürün Uzayları ve Bağımsız Deneyler	14
1.10 Bağımsızlığın Mühendislik Uygulaması: Sistem Güvenilirliği	15
1.11 Olasılığın Yorumu ve Modelleme Disiplini	16
1.12 Örnek Veri Tablosu ve Yazılım Çözümleri	17
1.13 Görselleştirme ve Grafikler	18
1.14 Bölüm Sonu Çözümlü Problemler	18
1.15 Bölüm Sonu Alıştırmaları	20
1.16 Kaynak Notu	21
2 Kombinatorik Analiz ve Sayma Yöntemleri	23
2.1 Gerçek Dünya Problemi (Giriş)	23
2.2 Sayma Probleminin Olasılıktaki Yeri	24
2.3 Toplama ve Çarpma Kuralları	24

2.4	Faktöriyel ve Permütasyon	25
2.5	Tekrarlı Permütasyon	26
2.6	Kombinasyon	27
2.7	Sıralı-Sırasız ve Tekrarlı-Tekrarsız Seçim	29
2.8	Binom Katsayıları ve Binom Teoremi	30
2.9	Multinom Katsayıları	31
2.10	İçerme-Dışlama ile Sayma	31
2.11	Yerine Koymalı ve Yerine Koymasız Örnekleme	32
2.12	Olasılık Problemlerinde Sayma	32
2.13	Dağılımlara Hazırlık: Binom ve Hipergeometrik Yapı	33
2.14	Örnek Veri Tablosu ve Yazılım Çözümleri	34
2.15	Görselleştirme ve Grafikler	35
2.16	Bölüm Sonu Çözümlü Problemler	36
2.17	Bölüm Sonu Alıştırmaları	37
2.18	Kaynak Notu	38
3	Rassal Değişkenler ve Dağılımlar	39
3.1	Gerçek Dünya Problemi (Giriş)	39
3.2	Rassal Değişken Kavramı	40
3.3	Dağılım Fonksiyonu	40
3.4	Kesikli Rassal Değişkenler	42
3.5	Sürekli Rassal Değişkenler	43
3.6	Karma Tip Dağılımlar ve Model Seçimi	44
3.7	Beklenen Değer, Varyans ve Momentler	45
3.8	Gösterge Değişkenleri ve Sayma	47
3.9	Olasılık Üreten Fonksiyon	48
3.10	Moment Üreten Fonksiyon	49
3.11	Temel Kesikli Dağılımlar	50
3.12	Kantiller ve Olasılık İntegral Dönüşümü	52
3.13	Dağılımın Bilgiye Bağlı Yorumu	53
3.14	Örnek Veri Tablosu ve Yazılım Çözümleri	54
3.15	Görselleştirme ve Grafikler	55
3.16	Bölüm Sonu Çözümlü Problemler	56
3.17	Bölüm Sonu Alıştırmaları	56
3.18	Kaynak Notu	58
4	Sürekli Dağılımlar, Dönüşümler ve Momentler	59
4.1	Gerçek Dünya Problemi (Giriş)	59
4.2	Sürekli Yoğunluk ve Dağılım Fonksiyonu	60
4.3	Kantiller ve Ters Dönüşüm	62
4.4	Dönüşümle Yeni Dağılım Elde Etme	63
4.5	Momentler ve Dağılımın Şekli	65
4.6	Moment Üreten Fonksiyon	66
4.7	Karakteristik Fonksiyon	68
4.8	Temel Sürekli Dağılım Aileleri	69

4.9	Örnek Veri Tablosu ve Yazılım Çözümleri	72
4.10	Görselleştirme ve Grafikler	73
4.11	Bölüm Sonu Çözümlü Problemler	73
4.12	Bölüm Sonu Alıştırmaları	74
4.13	Kaynak Notu	75
5	Ortak Dağılımlar ve Bağımlılık	77
5.1	Gerçek Dünya Problemi (Giriş)	77
5.2	Ortak Dağılım Fonksiyonu	78
5.3	Kesikli Ortak Dağılımlar ve Marjinaller	79
5.4	Sürekli Ortak Dağılımlar ve Marjinaller	80
5.5	Rassal Değişkenlerin Bağımsızlığı	81
5.6	İki Değişkenli Beklenti	82
5.7	Kovaryans	82
5.8	Korelasyon Katsayısı	84
5.9	Doğrusal Birleşimlerin ve Toplamların Varyansı	85
5.10	İki Değişkenli Normal Dağılım	86
5.11	Multinom Dağılım	87
5.12	Örnek Veri Tablosu ve Yazılım Çözümleri	88
5.13	Görselleştirme ve Grafikler	89
5.14	Bölüm Sonu Çözümlü Problemler	89
5.15	Bölüm Sonu Alıştırmaları	91
5.16	Kaynak Notu	92
6	Örnekleme Teorisi ve Örnekleme Dağılımları	93
6.1	Gerçek Dünya Problemi (Giriş)	93
6.2	Anakütle, Örneklem ve İstatistik	94
6.3	Rassal Örneklem	95
6.4	Örnekleme Dağılımı	96
6.5	Örneklem Ortalamasının Dağılımı	97
6.6	Örneklem Varyansı	98
6.7	Sonlu Anakütleden Örnekleme	99
6.8	Normal Anakütlerde Kesin Örnekleme Dağılımları	100
6.9	Örneklem Oranı ve Binom Modeli	101
6.10	Sıralı İstatistikler	102
6.11	Merkezi Limit Teoremi ve Büyük Örneklem Yaklaşımı	103
6.12	Örnek Veri Tablosu ve Yazılım Çözümleri	105
6.13	Görselleştirme ve Grafikler	106
6.14	Bölüm Sonu Çözümlü Problemler	107
6.15	Bölüm Sonu Alıştırmaları	108
6.16	Kaynak Notu	108

7	Asimptotik Teori ve Limit Teoremleri	111
7.1	Gerçek Dünya Problemi (Giriş)	111
7.2	Asimptotik Düşünce	112
7.3	Yakınsama Türleri	113
7.4	Büyük Sayılar Yasası	115
7.5	Merkezi Limit Teoremi	117
7.6	Normal Yaklaşım ve Süreklilik Düzeltmesi	119
7.7	Sürekli Dönüşüm ve Slutsky Teoremi	120
7.8	Delta Yöntemi	121
7.9	Tutarlılık ve Asimptotik Normallik	122
7.10	Örnek Veri Tablosu ve Yazılım Çözümleri	123
7.11	Görselleştirme ve Grafikler	124
7.12	Bölüm Sonu Çözümlü Problemler	125
7.13	Bölüm Sonu Alıştırmaları	126
7.14	Kaynak Notu	127
8	Özel Dağılımlar ve Dağılım Aileleri	129
8.1	Gerçek Dünya Problemi (Giriş)	129
8.2	Dağılım Ailesi ve Parametre Mantığı	130
8.3	Bernoulli ve Binom Dağılımları	130
8.4	Geometrik ve Negatif Binom Dağılımları	132
8.5	Hipergeometrik Dağılım	133
8.6	Poisson Dağılımı	134
8.7	Kesikli Düzgün Dağılım	136
8.8	Sürekli Düzgün Dağılım	136
8.9	Üstel ve Gamma Dağılımları	136
8.10	Beta Dağılımı	138
8.11	Normal Dağılım	139
8.12	Normal Dağılımdan Türeyen Dağılımlar	140
8.13	Dağılımlar Arası İlişkiler	141
8.14	Dağılım Seçim Kontrol Listesi	142
8.15	Örnek Veri Tablosu ve Yazılım Çözümleri	142
8.16	Görselleştirme ve Grafikler	143
8.17	Bölüm Sonu Çözümlü Problemler	144
8.18	Bölüm Sonu Alıştırmaları	145
8.19	Kaynak Notu	145
9	Koşullu Dağılımlar ve Beklenti	147
9.1	Gerçek Dünya Problemi (Giriş)	147
9.2	Koşullama Fikrinin Temeli	148
9.3	Kesikli Koşullu Dağılımlar	149
9.4	Sürekli Koşullu Yoğunluk	150
9.5	Bağımsızlık ve Koşullama	151
9.6	Koşullu Beklenti	152
9.7	Toplam Beklenti Yasası	153

9.8	Koşullu Varyans ve Toplam Varyans Yasası	154
9.9	Bayes Kuralı ve Güncelleme	155
9.10	Koşullu Beklenti ve En İyi Tahmin	156
9.11	Kule Özelliği ve Bilgi Düzeyleri	158
9.12	Koşullama ile Parçalama Stratejisi	158
9.13	Örnek Veri Tablosu ve Yazılım Çözümleri	159
9.14	Görselleştirme ve Grafikler	160
9.15	Bölüm Sonu Çözümlü Problemler	161
9.16	Bölüm Sonu Alıştırmaları	162
9.17	Kaynak Notu	163
10	Olasılık Eşitsizlikleri ve Yakınsama	165
10.1	Gerçek Dünya Problemi (Giriş)	165
10.2	Neden Olasılık Eşitsizlikleri?	166
10.3	Markov Eşitsizliği	166
10.4	Çebişev Eşitsizliği	167
10.5	Moment Sınırları ve Genelleştirme	169
10.6	Jensen Eşitsizliği	169
10.7	Cauchy-Schwarz Eşitsizliği	170
10.8	Boole Eşitsizliği ve Birleşim Sınırı	171
10.9	Yakınsama Türleri	172
10.10	Yakınsama İlişkileri	172
10.11	Borel-Cantelli Lemması	173
10.12	Yoğunlaşma: Hoeffding Eşitsizliği	174
10.13	Eşitsizliklerden Büyük Sayılar Yasasına	175
10.14	Örnek Veri Tablosu ve Yazılım Çözümleri	175
10.15	Görselleştirme ve Grafikler	176
10.16	Bölüm Sonu Çözümlü Problemler	177
10.17	Bölüm Sonu Alıştırmaları	178
10.18	Kaynak Notu	179
11	Simülasyon ve Monte Carlo Yöntemleri	181
11.1	Gerçek Dünya Problemi (Giriş)	181
11.2	Monte Carlo Düşüncesi	182
11.3	Sözde Rassal Sayılar ve Tekrarlanabilirlik	183
11.4	Kesikli Dağılım Simülasyonu	184
11.5	Ters Dönüşüm Yöntemi	184
11.6	Kabul-Red Yöntemi	185
11.7	Olay Olasılığı Tahmini	186
11.8	Monte Carlo Standart Hatası	187
11.9	Monte Carlo Güven Aralığı	187
11.10	Monte Carlo İntegrasyonu	188
11.11	Varyans Azaltma Fikirleri	189
11.12	Rassal Süreç Simülasyonuna Giriş	190
11.13	Simülasyon Sonucunu Raporlama	192

11.14	Örnek Veri Tablosu ve Yazılım Çözümleri	192
11.15	Görselleştirme ve Grafikler	193
11.16	Bölüm Sonu Çözümlü Problemler	194
11.17	Bölüm Sonu Alıştırmaları	195
11.18	Kaynak Notu	195
A	Ekler	197
A.1	Standart Normal Dağılım Tablosu	197
A.2	t Dağılımı Kritik Değerleri	198
A.3	z Dağılımı Kritik Değerleri	199
A.4	F Dağılımı Kritik Değerleri	199
A.5	Ki-Kare Dağılımı Kritik Değerleri	199
B	Cevap Anahtarı	201
B.1	Bölüm 1: Olasılığın Temelleri ve Aksiyomatik Yapı	201
B.2	Bölüm 2: Kombinatorik Analiz ve Sayma Yöntemleri	202
B.3	Bölüm 3: Rassal Değişkenler ve Dağılımlar	203
B.4	Bölüm 4: Sürekli Dağılımlar ve Dönüşümler	204
B.5	Bölüm 5: Ortak Dağılımlar ve Bağımlılık	205
B.6	Bölüm 6: Örneklem Teorisi ve Örneklem Dağılımları	206
B.7	Bölüm 7: Asimptotik Teori ve Limit Teoremleri	206
B.8	Bölüm 8: Özel Dağılımlar ve Dağılım Aileleri	206
B.9	Bölüm 9: Koşullu Dağılımlar ve Beklenti	207
B.10	Bölüm 10: Olasılık Eşitsizlikleri ve Yakınsama	208
B.11	Bölüm 11: Simülasyon ve Monte Carlo Yöntemleri	209
C	Hızlı Başvuru Kartı	211
D	Modelleme Rehberi	213
	Kaynakça	217
	Dizin	221

Bölüm 1

Olasılığın Temelleri ve Aksiyomatik Yapı

1.1 Gerçek Dünya Problemi (Giriş)

Bir hastanede yeni bir tanı testinin pozitif çıkması, bir fabrikada seçilen parçanın kusurlu olması, bir internet servisinde aynı dakika içinde çok sayıda bağlantı isteği gelmesi veya bir sigorta müşterisinin hasar bildirmesi ilk bakışta farklı problemlerdir. Ortak nokta şudur: Sonuç kesin değildir; fakat sonuçlar kümesi, olaylar ve bu olaylara atanan olasılıklar sistemli biçimde kurulursa belirsizlik hesaplanabilir hâle gelir.

Bu Bölüm Ne Sağlar?

Bu bölüm olasılığın temel dilini kurar: rassal deney, örnek uzay, olay, olay cebiri, olasılık ölçüsü, koşullu olasılık, bağımsızlık, toplam olasılık yasası ve Bayes teoremi. Anlatım önce sezgisel ve sonlu örneklerle başlar; ardından Kolmogorov aksiyomları ve ölçü temelli yapı ile matematiksel olgunluğa taşınır.

Amaç ve Öğrenme Çıktıları

Bu bölümün sonunda okuyucu

- bir rassal deneyi örnek uzay ve olaylar aracılığıyla modelleyebilir,
- Kolmogorov aksiyomlarını kullanarak temel olasılık sonuçlarını ispatlayabilir,
- sonlu ve eş olasılıklı modellerde sayma ilkelerini uygulayabilir,
- koşullu olasılık, bağımsızlık, toplam olasılık ve Bayes teoremini doğru bağlamda kullanabilir,
- olasılık modelinin fiziksel, frekansçı ve bilgi temelli yorumları arasındaki farkı açıklayabilir.

1.2 Rassal Deney, Örnek Uzay ve Olay

Olasılık hesabı bir formülle değil, model kurma ile başlar. Model kurmanın ilk adımı, deneyin ne olduğunu ve hangi sonuçların mümkün sayıldığını açıkça belirlemektir. Bu bakış, klasik olasılık kitaplarında ve mühendislik uygulamalarında temel başlangıç noktasıdır (Hsu, 1997; Ross, 2019; Soong, 2004).

Tanım 1.1 (Rassal Deney) Sonucu önceden kesin olarak bilinemeyen; fakat olası sonuçları önceden tanımlanabilen işleme rassal deney denir.

Tanım 1.2 (Örnek Uzay) Bir rassal deneyin tüm mümkün sonuçlarının oluşturduğu kümeye örnek uzay denir ve genellikle Ω ile gösterilir.

Tanım 1.3 (Örnek Nokta ve Olay) Örnek uzayın her bir elemanına örnek nokta denir. Örnek uzayın alt kümelerine olay denir. Bir olay, deney sonucunda gerçekleşip gerçekleşmediği sorgulanan önermedir.

Örnek 1.1 (Zar Deneyi). *Bir zar bir kez atıldığında örnek uzay*

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

olur. Çift gelme olayı $A = \{2, 4, 6\}$, beşten büyük gelme olayı $B = \{6\}$ biçiminde yazılır.

Örnek 1.2 (Mühendislikte Ömür Deneyi). *Bir elektronik parçanın arızalanma süresi ölçülüyorsa örnek uzay $\Omega = [0, \infty)$ olarak alınabilir. “Parça 100 saatten fazla çalışır” olayı $A = (100, \infty)$ biçiminde ifade edilir.*

Olasılık problemlerinde en yaygın hata, örnek uzay kurulmadan formül uygulamaktır. Deneyin ne ürettiği net değilse olay da, olayın olasılığı da net değildir.

1.3 Olay Cebiri ve Küme İşlemleri

Olaylar kümelerle ifade edildiği için olasılığın dili küme işlemlerine dayanır. Birleşim “veya”, kesişim “ve”, tümleyen ise “olmama” anlamını taşır. Bu dil, aksiyomatik olasılığın temelidir (Capiński & Kopp, 2004; Kolmogorov, 1956).

Tanım 1.4 (Birleşim, Kesişim ve Tümleyen) A ve B olayları için

$$A \cup B = \{\omega : \omega \in A \text{ veya } \omega \in B\},$$

$$A \cap B = \{\omega : \omega \in A \text{ ve } \omega \in B\},$$

$$A^c = \{\omega : \omega \notin A\}$$

olarak tanımlanır.

Tanım 1.5 (Ayrık Olaylar) $A \cap B = \emptyset$ ise A ve B olaylarına ayrık olaylar denir. Ayrık olaylar aynı deney sonucunda birlikte gerçekleşemez.

Tanım 1.6 (σ -Cebiri) Ω üzerinde bir $\mathcal{F}$ olay sınıfı şu koşulları sağlıyorsa σ -cebiri olarak adlandırılır:

1. $\Omega \in \mathcal{F}$,
2. $A \in \mathcal{F}$ ise $A^c \in \mathcal{F}$,
3. $A_1, A_2, \dots \in \mathcal{F}$ ise $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Örnek 1.3 (Basit Olay Cebiri). $\Omega = \{1, 2, 3, 4, 5, 6\}$ ve $A = \{2, 4, 6\}$ olsun. Bu durumda $A^c = \{1, 3, 5\}$ olur. A çift gelmeyi, A^c tek gelmeyi temsil eder.

Küme işlemleri arasındaki en kullanışlı ilişki, “veya” ile “ve” işlemlerinin tümleyen altında yer değiştirmesidir. De Morgan kuralları, olasılık hesaplarında bir olayın tümleyenini daha kolay hesaplanan bir biçime çevirmek için sürekli kullanılır (Ross, 2019; Billingsley, 1995).

Teorem 1.1 (De Morgan Kuralları) Herhangi bir $\{A_i\}_{i \in I}$ olay ailesi için

$$\left(\bigcup_{i \in I} A_i \right)^c = \bigcap_{i \in I} A_i^c, \quad \left(\bigcap_{i \in I} A_i \right)^c = \bigcup_{i \in I} A_i^c$$

geçerlidir.

Kanıt. İlk eşitliği gösterelim. $\omega \in (\bigcup_i A_i)^c$ olması, ω 'nın hiçbir A_i 'ye ait olmaması, yani her i için $\omega \in A_i^c$ olması demektir; bu da $\omega \in \bigcap_i A_i^c$ ile aynıdır. İki kümenin elemanları birebir örtüştüğünden eşitlik sağlanır. İkinci eşitlik, ilk eşitlik A_i^c ailesine uygulanıp tümleyen alınarak elde edilir. $\square$

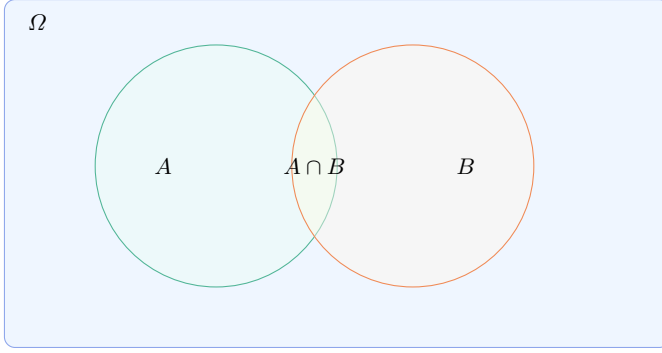
Tanım 1.7 (Artan ve Azalan Olay Dizileri) $A_1, A_2, \dots$ olay dizisi için her n 'de $A_n \subseteq A_{n+1}$ ise dizi *artan*, her n 'de $A_{n+1} \subseteq A_n$ ise *azalan* olarak adlandırılır. Artan dizinin limiti $\bigcup_n A_n$, azalan dizinin limiti $\bigcap_n A_n$ olarak tanımlanır ve $A_n \uparrow A$ ya da $A_n \downarrow A$ ile gösterilir.

Tanım 1.8 (Olay Dizisinin Üst ve Alt Limiti) Bir $A_1, A_2, \dots$ olay dizisi için

$$\limsup_n A_n = \bigcap_{n=1}^{\infty} \bigcup_{k \geq n} A_k, \quad \liminf_n A_n = \bigcup_{n=1}^{\infty} \bigcap_{k \geq n} A_k$$

tanımlanır. $\limsup A_n$, “sonsuz çok A_n gerçekleşir” olayını; $\liminf A_n$ ise “bir indisten sonra tüm A_n gerçekleşir” olayını ifade eder.

Örnek 1.4 (Monoton Olay Dizisi). $A_n = (1/n, 1]$ aralıkları $[0, 1]$ üzerinde tanımlı olsun. Dizi artandır ve $\bigcup_n A_n = (0, 1]$ olur. Benzer biçimde $B_n = [0, 1/n)$ azalan bir dizidir ve $\bigcap_n B_n = \{0\}$ değerine iner. Bu tür limitler, sürekli olasılık modellerinde olayların “sınıra yaklaşması” fikrini somutlaştırır.



Şekil 1.1 İki olayın Venn şeması. Ortadaki kesişim bölgesi $A \cap B$, toplama kuralında bir kez fazladan sayıldığı için $P(A \cup B)$ hesabında çıkarılır.

1.4 Olasılık Ölçüsü ve Kolmogorov Aksiyomları

Modern olasılık kuramında olasılık, olaylara sayı atayan bir ölçü olarak tanımlanır. Bu yaklaşım Kolmogorov'un aksiyomatik çerçevesiyle standart hâle gelmiştir (Kolmogorov, 1956). Aksiyomlar kısa görünür; fakat olasılık teorisinin neredeyse bütün temel sonuçları bu üç ilkedен türetilir.

Tanım 1.9 (Olasılık Uzayı) Bir olasılık modeli

$$(\Omega, \mathcal{F}, P)$$

üçlüsünden oluşur. Burada Ω örnek uzay, $\mathcal{F}$ olayların σ -cebiri, P ise olaylara olasılık atayan fonksiyondur.

Tanım 1.10 (Kolmogorov Aksiyomları) $P : \mathcal{F} \rightarrow [0, 1]$ fonksiyonu aşağıdaki koşulları sağlıyorsa olasılık ölçüsüdür:

1. **Negatif olmama:** Her $A \in \mathcal{F}$ için $P(A) \geq 0$.
2. **Normallik:** $P(\Omega) = 1$.
3. **Sayılabilir toplamsallık:** $A_1, A_2, \dots$ karşılıklı ayrık olaylar ise

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

Teorem 1.2 (Boş Kümenin Olasılığı) Her olasılık uzayında

$$P(\emptyset) = 0$$

olur.

Kanıt. Ω ve $\emptyset$ ayrık olaylardır ve $\Omega \cup \emptyset = \Omega$ olur. Toplamsallıktan

$$P(\Omega) = P(\Omega) + P(\emptyset)$$

yazılır. Her iki taraftan $P(\Omega)$ çıkarılırsa $P(\emptyset) = 0$ elde edilir. $\square$

Teorem 1.3 (Tümleyen Kuralı) Her $A \in \mathcal{F}$ için

$$P(A^c) = 1 - P(A)$$

olur.

Kanıt. A ve A^c ayrık olaylardır ve birleşimleri Ω 'dır. Bu nedenle

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

Düzenleme sonucu verir. $\square$

Teorem 1.4 (Monotonluk) $A \subseteq B$ ise

$$P(A) \leq P(B)$$

olur.

Kanıt. B olayı ayrık olarak A ve $B \cap A^c$ parçalarına ayrılır. Dolayısıyla

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

Olasılıklar negatif olmadığından $P(B \cap A^c) \geq 0$ ve sonuç $P(A) \leq P(B)$ olur. $\square$

Teorem 1.5 (İki Olay İçin Toplama Kuralı) Her $A, B \in \mathcal{F}$ için

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

olur.

Kanıt. $A \cup B$, ayrık biçimde A ve $B \cap A^c$ olaylarının birleşimidir. Bu nedenle

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

Ayrıca $B = (A \cap B) \cup (B \cap A^c)$ ayrık birleşimidir. Buradan

$$P(B \cap A^c) = P(B) - P(A \cap B)$$

yazılır. Yerine koyunca istenen eşitlik elde edilir. $\square$

Örnek 1.5 (Çift veya Büyük Sayı). *Adil bir zar atıldığında $A = \{2, 4, 6\}$ çift gelme, $B = \{5, 6\}$ en az 5 gelme (yani 4'ten büyük gelme) olayı olsun. $A \cap B = \{6\}$ olduğundan*

$$P(A \cup B) = \frac{3}{6} + \frac{2}{6} - \frac{1}{6} = \frac{4}{6}$$

bulunur.

İki olay için toplama kuralı, ikiden çok olaya doğal biçimde genişler. Bu genelleme *içerme-dışlama ilkesi* olarak bilinir ve eşleşme, kapsama ve örtme problemlerinin temelidir (Feller, 1968; Ross, 2019).

Teorem 1.6 (Sonlu Toplamsallık) $A_1, \dots, A_n$ karşılıklı ayrık olaylar ise

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

olur.

Kanıt. Sayılabilir toplamsallık aksiyomunda $A_{n+1} = A_{n+2} = \dots = \emptyset$ alınır. $P(\emptyset) = 0$ olduğundan sonsuz toplam ilk n terime iner. $\square$

Teorem 1.7 (Boole Eşitsizliği (Alt Toplamsallık)) Herhangi $A_1, A_2, \dots$ olayları için (ayrık olmaları gerekmez)

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

geçerlidir.

Kanıt. Olayları ayrıklaştıralım: $B_1 = A_1$ ve $n \geq 2$ için $B_n = A_n \setminus (A_1 \cup \dots \cup A_{n-1})$ tanımlansın. Bu durumda B_n 'ler karşılıklı ayrıktır, $\bigcup_n B_n = \bigcup_n A_n$ ve $B_n \subseteq A_n$ olur. Monotonluktan $P(B_n) \leq P(A_n)$ ve toplamsallıktan

$$P\left(\bigcup_n A_n\right) = \sum_n P(B_n) \leq \sum_n P(A_n).$$

$\square$

Teorem 1.8 (İçerme-Dışlama İlkesi) $A_1, \dots, A_n$ olayları için

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

Kanıt. $n = 2$ durumu toplama kuralıdır. Tümevarım varsayımıyla $C = A_1 \cup \dots \cup A_{n-1}$ için formül doğru kabul edilir. $P(C \cup A_n) = P(C) + P(A_n) - P(C \cap A_n)$ yazılır; $C \cap A_n = \bigcup_{i < n} (A_i \cap A_n)$ olduğundan bu terime de tümevarım uygulanır. Terimler düzenlendiğinde n olay için formül elde edilir. $\square$

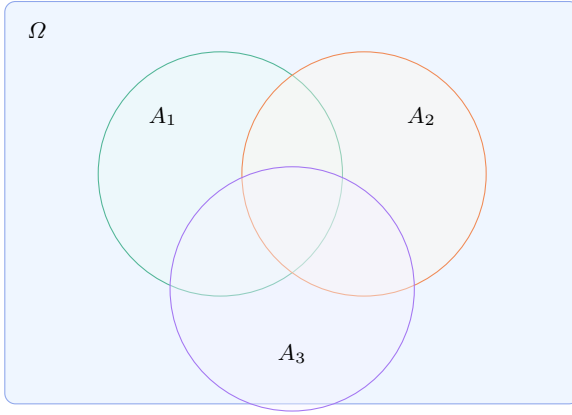
Teorem 1.9 (Bonferroni Eşitsizlikleri) İçerme-dışlama toplamının ilk teriminde durulursa üst, ilk iki teriminde durulursa alt sınır elde edilir:

$$\sum_i P(A_i) - \sum_{i < j} P(A_i \cap A_j) \leq P\left(\bigcup_i A_i\right) \leq \sum_i P(A_i).$$

Kant. Üst sınır Boole eşitsizliğidir. Alt sınır, içme–dışlama serisinin kısmi toplamlarının ardışık olarak gerçek değeri alttan ve üstten kuşatması (terimlerin işaret değiştirmesi) ilkesinden çıkar. $\square$

Örnek 1.6 (Üç Olay İçin İçme–Dışlama). $P(A_i) = 0.3$, *ikişerli kesişimler* 0.1 ve *üçlü kesişim* 0.02 olsun. O hâlde

$$P(A_1 \cup A_2 \cup A_3) = 3(0.3) - 3(0.1) + 0.02 = 0.62.$$



Şekil 1.2 Üç olayda içme–dışlama: tek tek toplamlar kesişimleri fazla sayar; ikişerli kesişimler çıkarılır, üçlü kesişim geri eklenir.

Aksiyomların belki de en güçlü sonucu, olasılık ölçüsünün monoton olay dizilerinde *sürekli* davranmasıdır. Bu süreklilik, sonlu hesapları sonsuz/sürekli modellere taşıyan köprüdür (Billingsley, 1995; Çınlar, 2011).

Teorem 1.10 (Olasılığın Sürekliliği) $A_n \uparrow A$ ise $P(A_n) \uparrow P(A)$; $A_n \downarrow A$ ise $P(A_n) \downarrow P(A)$ olur.

Kant. Artan durum için $B_1 = A_1$, $n \geq 2$ 'de $B_n = A_n \setminus A_{n-1}$ ayrık olaylarını alalım. $\bigcup_{k=1}^n B_k = A_n$ ve $\bigcup_k B_k = A$ olur. Sayılabilir toplamsallıktan

$$P(A) = \sum_{k=1}^{\infty} P(B_k) = \lim_{n \rightarrow \infty} \sum_{k=1}^n P(B_k) = \lim_{n \rightarrow \infty} P(A_n).$$

Azalan durum, $A_n^c \uparrow A^c$ olduğu gözlemiyle tümleyen kuralından elde edilir. $\square$

1.5 Sonlu ve Eş Olasılıklı Modeller

Klasik olasılıkta tüm sonuçlar eş olasılıklıysa olay olasılığı sayma problemine indirgenir. Bu yaklaşım, zar, kart, kura ve basit örnekleme problemlerinin doğal temelidir (Grinstead & Snell, 1997; Ross, 2019).

Tanım 1.11 (Klasik Olasılık) Ω sonlu ve tüm örnek noktalar eş olasılıklıysa her $A \subseteq \Omega$ için

$$P(A) = \frac{|A|}{|\Omega|}$$

olur.

Tanım 1.12 (Çarpma İlkesi) Bir işlem ardışık iki aşamada yapılıyor, ilk aşama m farklı biçimde, ikinci aşama her ilk seçimden sonra n farklı biçimde yapılabilirse toplam mn farklı sonuç vardır.

Tanım 1.13 (Permütasyon ve Kombinasyon) n farklı nesneden sıralı r seçim sayısı

$$P(n, r) = \frac{n!}{(n - r)!},$$

sırasız r seçim sayısı ise

$$\binom{n}{r} = \frac{n!}{r!(n - r)!}$$

olur.

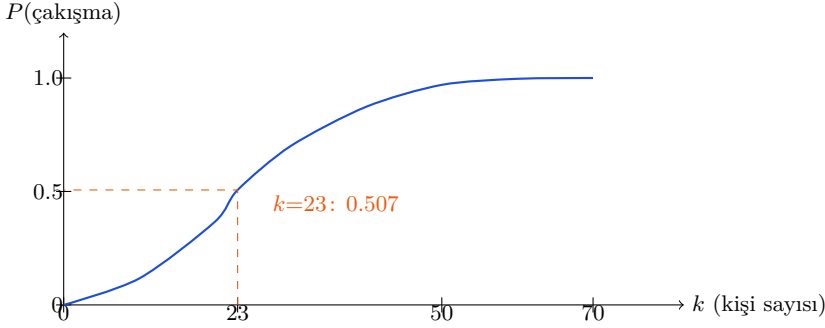
Örnek 1.7 (Kart Seçimi). 52 kartlık desteden 5 kartlık el seçme sayısı $\binom{52}{5}$ olur. Eğer belirli bir el tipi sayılabiliyorsa, bu sayı $\binom{52}{5}$ değerine bölünerek olasılık bulunur.

Örnek 1.8 (Basit Rassal Örnekleme). N elemanlı bir kitleden n eleman yerine koymadan ve eş olasılıkla seçiliyorsa, belirli bir birimin örneğe girme olasılığı n/N olur. Çünkü bu birimi içeren örneklem sayısı $\binom{N-1}{n-1}$, tüm örneklem sayısı $\binom{N}{n}$ 'dir.

Örnek 1.9 (Doğum Günü Problemi). k kişilik bir grupta en az iki kişinin aynı doğum gününe sahip olma olasılığını hesaplayalım (yıl 365 gün, doğumlar bağımsız ve eş olasılıklı varsayalım). Tümleyen üzerinden gitmek kolaydır: tüm doğum günlerinin farklı olma olasılığı

$$P(\text{hepsi farklı}) = \frac{365 \cdot 364 \cdots (365 - k + 1)}{365^k} = \prod_{j=0}^{k-1} \left(1 - \frac{j}{365}\right).$$

Aranan olasılık $1 - P(\text{hepsi farklı})$ olur. Sezgiye aykırı biçimde $k = 23$ için bu olasılık 0.507'yi geçer; $k = 50$ 'de ise 0.97'ye ulaşır. Küçük bir grupta çakışmanın bu kadar olası olması, $\binom{k}{2}$ çift sayısının k ile karesel büyümesinden kaynaklanır.



Şekil 1.3 Doğum günü problemi: çakışma olasılığı kişi sayısı ile hızla 1'e yaklaşır ve $k = 23$ 'te $1/2$ 'yi geçer.

Örnek 1.10 (Eşleşme (Şapka) Problemi ve Dağınklıklar). n konuk şapkalarını vestiyere bırakıp çıkışta rastgele birer şapka alsın. Hiçbir konunun kendi şapkasını almaması olayının olasılığını arıyoruz. Hiçbir sabit noktası olmayan permütasyona dağınklık (derangement) denir ve sayısı D_n ile gösterilir. İçerme-dışlama ilkesiyle

$$D_n = n! \sum_{j=0}^n \frac{(-1)^j}{j!}, \quad P(\text{hiç eşleşme yok}) = \frac{D_n}{n!} = \sum_{j=0}^n \frac{(-1)^j}{j!} \xrightarrow{n \rightarrow \infty} e^{-1} \approx 0.368.$$

Konuk sayısı büyüdükçe bu olasılık konuk sayısından bağımsız biçimde e^{-1} 'e yakınsar; bu, içerme-dışlama ilkesinin zarif bir uygulamasıdır (Feller, 1968; Grinstead & Snell, 1997).

1.6 Koşullu Olasılık

Yeni bilgi geldiğinde örnek uzay fiilen daralır. Koşullu olasılık, bu bilgi güncellemesini sayısal biçimde ifade eder. Koşullu olasılık Bayes teoreminin, bağımsızlık kavramının ve istatistiksel çıkarımın ana kapısıdır (Hogg et al., 2019; Hsu, 1997; Ross, 2019).

Tanım 1.14 (Koşullu Olasılık) $P(B) > 0$ ise B bilindiğinde A olayının koşullu olasılığı

$$P(A | B) = \frac{P(A \cap B)}{P(B)}$$

olarak tanımlanır.

Teorem 1.11 (Çarpım Kuralı) $P(B) > 0$ ise

$$P(A \cap B) = P(A | B)P(B)$$

olur. Ayrıca $P(A \cap B) > 0$ ise

$$P(A \cap B \cap C) = P(C \mid A \cap B)P(B \mid A)P(A)$$

yazılabilir.

Kanıt. İlk eşitlik koşullu olasılık tanımının her iki tarafının $P(B)$ ile çarpılmasıyla elde edilir. Üç olay için aynı tanım önce C olayına, sonra B olayına uygulanır:

$$P(A \cap B \cap C) = P(C \mid A \cap B)P(A \cap B) = P(C \mid A \cap B)P(B \mid A)P(A).$$

□

Örnek 1.11 (Kalite Kontrol). *Bir ürünün kusurlu olma olasılığı 0.04 ve kusurlu ürünün testi geçememe olasılığı 0.90 olsun. Ürünün kusurlu ve testi geçememiş olma olasılığı*

$$0.04 \times 0.90 = 0.036$$

olur.

Koşullu olasılığın aksiyomları bozmadığını görmek önemlidir: bir olay sabitlendiğinde geri kalan dünya hâlâ tutarlı bir olasılık modelidir. Bu, koşullandırmanın neden “örnek uzayı daraltmak” olarak okunabileceğini açıklar.

Teorem 1.12 (Koşullu Olasılık Bir Olasılık Ölçüsüdür) $P(B) > 0$ olsun. $Q(\cdot) = P(\cdot \mid B)$ ile tanımlanan fonksiyon $(\Omega, \mathcal{F})$ üzerinde Kolmogorov aksiyomlarını sağlar.

Kanıt. Her A için $P(A \cap B) \geq 0$ ve $P(B) > 0$ olduğundan $Q(A) \geq 0$. Ayrıca $Q(\Omega) = P(B)/P(B) = 1$. Son olarak $A_1, A_2, \dots$ ayrık ise $(A_i \cap B)$ da ayrıktır ve

$$Q\left(\bigcup_i A_i\right) = \frac{P(\bigcup_i (A_i \cap B))}{P(B)} = \frac{\sum_i P(A_i \cap B)}{P(B)} = \sum_i Q(A_i).$$

Üç aksiyom da sağlandığından Q bir olasılık ölçüsüdür. □

Örnek 1.12 (Monty Hall Problemi). *Üç kapıdan birinin arkasında araba, ikisinin arkasında keçi vardır. Yarışmacı bir kapı seçer; sunucu kalan iki kapıdan arkasında keçi olan birini açar ve seçimi değiştirme fırsatı verir. Değiştirmek mi avantajlıdır? İlk seçimin doğru olma olasılığı 1/3, yanlış olma olasılığı 2/3’tür. Yarışmacı ilk seçiminde yanlış kapıyı seçtiyse (olasılık 2/3), sunucunun zorunlu açışı sonrası değiştirmek mutlaka arabayı kazandırır. Dolayısıyla*

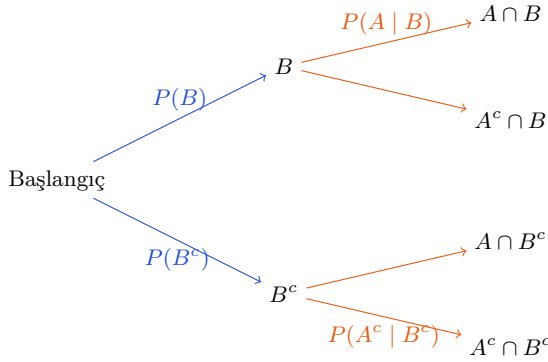
$$P(\text{değiştirerek kazanma}) = \frac{2}{3}, \quad P(\text{kalarak kazanma}) = \frac{1}{3}.$$

Sezgiye aykırı sonuç, sunucunun bilgili davranarak örnek uzaya yeni bilgi katmasından kaynaklanır (Grinstead & Snell, 1997).

Örnek 1.13 (İki Çocuk Problemi). *Bir ailenin iki çocuğu olsun ve her birinin kız/erkek olma olasılığı eşit ve bağımsız kabul edilsin. “En az biri kız” bilindiğinde her ikisinin de kız olma olasılığı, örnek uzay $\{KK, KE, EK, EE\}$ üzerinde*

$$P(KK \mid \text{en az bir } K) = \frac{P(KK)}{P(\{KK, KE, EK\})} = \frac{1/4}{3/4} = \frac{1}{3}$$

olur. Buna karşılık “ilk çocuk kız” bilindiğinde olasılık $1/2$ 'ye çıkar. Aradaki fark, verilen bilginin örnek uzayı nasıl farklı daralttığını gösterir.



Şekil 1.4 Olasılık ağacı: dallar boyunca çarpım kuralı, dalların toplamı ise toplam olasılık yasası uygulanır.

1.7 Toplam Olasılık Yasası ve Bayes Teoremi

Bir olayın olasılığı çoğu zaman alt gruplara ayrılarak hesaplanır. Toplam olasılık yasası bu parçalama işlemini, Bayes teoremi ise gözleminden nedene doğru ters güncellemeyi sağlar.

Tanım 1.15 (Ayrışım) $B_1, \dots, B_n$ olayları karşılıklı ayrık, $P(B_i) > 0$ ve

$$\Omega = B_1 \cup \dots \cup B_n$$

ise bu olaylar Ω 'nın bir ayrışımını oluşturur.

Teorem 1.13 (Toplam Olasılık Yasası) $B_1, \dots, B_n$ örnek uzayın bir ayrışımını ise her A olayı için

$$P(A) = \sum_{i=1}^n P(A \mid B_i)P(B_i)$$

olur.

Kanıt. A olayı ayrık biçimde

$$A = (A \cap B_1) \cup \dots \cup (A \cap B_n)$$

olarak yazılır. Toplamsallıktan

$$P(A) = \sum_{i=1}^n P(A \cap B_i)$$

olur. Koşullu olasılık tanımıyla $P(A \cap B_i) = P(A | B_i)P(B_i)$ yazılır ve sonuç elde edilir. $\square$

Teorem 1.14 (Bayes Teoremi) $B_1, \dots, B_n$ örnek uzayın bir ayrışımı ve $P(A) > 0$ ise

$$P(B_k | A) = \frac{P(A | B_k)P(B_k)}{\sum_{i=1}^n P(A | B_i)P(B_i)}$$

olur.

Kanıt. Koşullu olasılıktan

$$P(B_k | A) = \frac{P(A \cap B_k)}{P(A)}$$

yazılır. Pay $P(A | B_k)P(B_k)$ biçimine dönüştürülür. Payda için toplam olasılık yasası uygulanır. Böylece Bayes formülü elde edilir. $\square$

Teorem 1.15 (Bayes Teoreminin Olasılık Oranı Biçimi) $P(A)$, $P(A^c)$, $P(B | A)$ ve $P(B | A^c)$ pozitif ise

$$\frac{P(A | B)}{P(A^c | B)} = \frac{P(A)}{P(A^c)} \cdot \frac{P(B | A)}{P(B | A^c)}.$$

Kanıt. Bayes formülü A ve A^c için ayrı ayrı yazılır:

$$P(A | B) = \frac{P(B | A)P(A)}{P(B)}, \quad P(A^c | B) = \frac{P(B | A^c)P(A^c)}{P(B)}.$$

İlk eşitlik ikinciye bölündüğünde $P(B)$ sadeleşir ve sonuç elde edilir. $\square$

Örnek 1.14 (Tanı Testi). *Bir hastalığın toplumdaki oranı 0.01, testin duyarlılığı 0.95 ve yanlış pozitif oranı 0.05 olsun. Pozitif test sonrası hastalık olasılığı*

$$P(H | +) = \frac{0.95(0.01)}{0.95(0.01) + 0.05(0.99)} \approx 0.161$$

olur. Test güçlü görünse bile hastalık nadir olduğu için pozitif sonuç hastalığı kesinleştirmez.

Örnek 1.15 (Bahis Oranı (Odds) Biçimi). *Aynı problemi olasılık oranı biçiminde çözmek hesabı sadeleştirir. Önsel oran $P(H)/P(H^c) = 0.01/0.99$ ve olabilirlik oranı (likelihood ratio) $0.95/0.05 = 19$ olduğundan sonsal oran*

$$\frac{P(H | +)}{P(H^c | +)} = \frac{0.01}{0.99} \times 19 \approx 0.192, \quad P(H | +) = \frac{0.192}{1 + 0.192} \approx 0.161.$$

Oran biçimi, paydadaki toplam olasılığı tekrar hesaplamadan güncelleme yapmayı sağlar; ardışık kanıtlarda olabilirlik oranları yalnızca çarpılır.

Örnek 1.16 (Ardışık Bayes Güncellemesi). *Aynı hastaya bağımsız ikinci bir test daha uygulansın ve yine pozitif çıksın. Birinci testin sonsalı (≈ 0.161) ikinci testin önseli olur. Oran biçiminde sonsal oran tekrar 19 ile çarpılır:*

$$\frac{0.161}{0.839} \times 19 \approx 3.65, \quad P(H | ++)= \frac{3.65}{1 + 3.65} \approx 0.785.$$

Tek başına ikna edici olmayan iki kanıt birikince inanç 0.16'dan 0.79'a çıkar. Bu, Bayesçi güncellenmenin kanıt biriktirme mantığını gösterir (Jaynes, 2003).

1.8 Bağımsızlık

Bağımsızlık, iki olaydan birinin gerçekleşmesinin diğerinin olasılığını değiştirmemesi anlamına gelir. Bu kavram hem sezgisel hem de matematiksel olarak dikkat ister; ayrık olaylarla bağımsız olaylar karıştırılmamalıdır (Bertsekas & Tsitsiklis, 2008; Grinstead & Snell, 1997).

Tanım 1.16 (Bağımsız Olaylar) A ve B olayları

$$P(A \cap B) = P(A)P(B)$$

koşulunu sağlıyorsa bağımsızdır.

Teorem 1.16 (Koşullu Olasılık ve Bağımsızlık) $P(B) > 0$ olsun. A ve B bağımsız ise

$$P(A | B) = P(A)$$

olur. Tersisi de doğrudur: $P(A | B) = P(A)$ ise A ve B bağımsızdır.

Kanıt. Bağımsızlık varsa

$$P(A | B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A)$$

olur. Ters yönde $P(A | B) = P(A)$ eşitliğinin iki tarafı $P(B)$ ile çarpılırsa $P(A \cap B) = P(A)P(B)$ elde edilir. $\square$

Tanım 1.17 (Karşılıklı Bağımsızlık) $A_1, \dots, A_n$ olayları için her alt küme seçimi altında

$$P\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} P(A_j)$$

sağlanıyorsa olaylar karşılıklı bağımsızdır.

Örnek 1.17 (İki Zar). *İki adil zar atıldığında birinci zarın çift gelmesi ile ikinci zarın 6 gelmesi bağımsızdır. Çünkü deneyin mekanizması iki zar sonucunu birbirinden ayırır ve ortak olasılık çarpım biçiminde hesaplanır.*

Ayrık olaylar genellikle bağımsız değildir. $A \cap B = \emptyset$ ve $P(A), P(B) > 0$ ise $P(A \cap B) = 0$, fakat $P(A)P(B) > 0$ olur. Bu nedenle ayrıklık “birlikte olamaz”, bağımsızlık ise “birbirini etkilemez” anlamına gelir.

Teorem 1.17 (Tümleyenlerin Bağımsızlığı) A ve B bağımsız ise A ile B^c , A^c ile B , ve A^c ile B^c çiftleri de bağımsızdır.

Kanıt. $A = (A \cap B) \cup (A \cap B^c)$ ayrık birleşiminden $P(A \cap B^c) = P(A) - P(A \cap B) = P(A) - P(A)P(B) = P(A)(1 - P(B)) = P(A)P(B^c)$ olur. Diğer çiftler simetri ve aynı argümanla elde edilir. $\square$

Tanım 1.18 (Koşullu Bağımsızlık) $P(C) > 0$ olsun. A ve B olayları

$$P(A \cap B \mid C) = P(A \mid C)P(B \mid C)$$

koşulunu sağlıyorsa C verildiğinde *koşullu bağımsız* adını alır. Koşulsuz bağımsızlık koşullu bağımsızlığı gerektirmez; tersi de doğru değildir.

Örnek 1.18 (İkişerli Bağımsız Ama Karşılıklı Bağımlı). *İki adil madeni para atılsın. $A = \text{“ilk yazı”}$, $B = \text{“ikinci yazı”}$, $C = \text{“iki sonuç farklı”}$ olaylarını tanımlayalım. Her biri $1/2$ olasılıklıdır ve*

$$P(A \cap B) = P(A \cap C) = P(B \cap C) = \frac{1}{4} = P(A)P(B)$$

olduğundan olaylar ikişerli bağımsızdır. Ancak $A \cap B \cap C = \emptyset$ iken $P(A)P(B)P(C) = 1/8 \neq 0$ olur; dolayısıyla karşılıklı bağımsız değildirler. Bu klasik Bernstein örneği, ikişerli bağımsızlığın karşılıklı bağımsızlığı gerektirmediğini gösterir (Feller, 1968).

1.9 Ürün Uzayları ve Bağımsız Deneyler

Birden fazla deney birlikte modellenmek istendiğinde ürün uzayı kullanılır. Bu yapı bağımsız ölçümlerin ve çok değişkenli olasılığın başlangıç noktasıdır (Capiński & Kopp, 2004; Shao, 2003).

Tanım 1.19 (Ürün σ -Cebiri) $(\Omega_1, \mathcal{F}_1)$ ve $(\Omega_2, \mathcal{F}_2)$ ölçülebilir uzayları için ürün σ -cebiri, $A_1 \times A_2$ biçimindeki ölçülebilir dikdörtgenlerin ürettiği en küçük σ -cebidir ve $\mathcal{F}_1 \otimes \mathcal{F}_2$ ile gösterilir.

Teorem 1.18 (Bağımsız Ürün Olasılığı) P_1 ve P_2 olasılık ölçüleriye, dikdörtgenlerde

$$P(A_1 \times A_2) = P_1(A_1)P_2(A_2)$$

biçiminde tanımlanan ölçü, iki bağımsız deneyin ortak olasılık ölçüsünü verir.

Kanıt. Ölçü kuramında ürün ölçü varlık teoremi, dikdörtgenler üzerinde verilen çarpım biçimli tanımı $\mathcal{F}_1 \otimes \mathcal{F}_2$ üzerine tekil biçimde genişletir. Bu genişletme, dikdörtgen olaylarda bağımsızlık çarpımını korur. $\square$

Örnek 1.19 (İki Bağımsız Ölçüm). *Bir sıcaklık ölçümü ve bir basınç ölçümü bağımsız modelleniyorsa, ortak olay $A \times B$ için olasılık $P_T(A)P_P(B)$ olur.*

1.10 Bağımsızlığın Mühendislik Uygulaması: Sistem Güvenilirliği

Bağımsızlık kavramının en somut uygulamalarından biri, bileşenlerden oluşan sistemlerin çalışma olasılığını hesaplamaktır. Her bileşenin çalışma olasılığına *güvenilirlik* denir ve bileşen arızaları bağımsız varsayılır (Soong, 2004; Montgomery & Runger, 2018).

Teorem 1.19 (Seri ve Paralel Sistem Güvenilirliği) *Bağımsız bileşenlerin güvenilirlikleri $r_1, \dots, r_n$ olsun. Tüm bileşenlerin çalışmasını gerektiren seri sistemin güvenilirliği*

$$R_{\text{seri}} = \prod_{i=1}^n r_i,$$

en az bir bileşenin çalışmasının yeterli olduğu paralel sistemin güvenilirliği ise

$$R_{\text{paralel}} = 1 - \prod_{i=1}^n (1 - r_i)$$

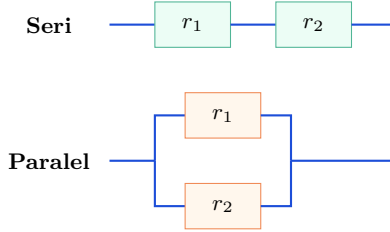
olur.

Kanıt. Seri sistemde sistem yalnız tüm bileşenler çalışırsa çalışır; bağımsızlıkla bu olayın olasılığı güvenilirliklerin çarpımıdır. Paralel sistemde sistem yalnız tüm bileşenler arızalanırsa arızalanır; bu olayın olasılığı $\prod_i (1 - r_i)$ olduğundan tümleyeni alınır. $\square$

Örnek 1.20 (Yedekli Tasarım). *Güvenilirliği 0.9 olan tek bir bileşen yerine, aynı güvenilirlikte iki bileşen paralel bağlanırsa sistem güvenilirliği*

$$1 - (1 - 0.9)^2 = 1 - 0.01 = 0.99$$

olur. Yedeklilik, tek bileşendeki 0.1'lik arıza riskini 0.01'e indirir. Bu, kritik sistemlerde paralel (yedekli) tasarımın temel gerekçesidir.



Şekil 1.5 Seri sistemde tüm bileşenler gereklidir ($R = \prod r_i$); paralel sistemde bir bileşen yeterlidir ($R = 1 - \prod (1 - r_i)$).

1.11 Olasılığın Yorumu ve Modelleme Disiplini

Olasılık yalnız hesap aracı değildir; belirsizliği nasıl yorumladığımızı da belirler. Frekansçı bakış olasılığı uzun dönem görelî sıklıkla, Bayesçi bakış ise bilgi altında tutarlı inanç derecesiyle ilişkilendirir. Jaynes'in vurguladığı gibi olasılık, eksik bilgi altında tutarlı akıl yürütmenin nicel dili olarak da okunabilir (Jaynes, 2003). Ölçüm problemlerinde ise yoğunlukların kullanılan hacim elemanı ile birlikte düşünülmesi gerekir; koordinat değişse bile fiziksel bölgenin olasılığı değişmemelidir (Tarantola, 2001).

Tanım 1.20 (Olasılık Yoğunluğu ve Hacim Elemanı) Sürekli modellerde bir bölgenin olasılığı, yoğunluğun uygun hacim elemanı üzerinde integraliyle hesaplanır. Bu nedenle yoğunluk değeri tek başına değil, hangi ölçüye göre tanımlandığıyla birlikte anlam taşır.

Teorem 1.20 (Bölge Olasılığının Koordinattan Bağımsızlığı) Bir koordinat dönüşümünde yoğunluk Jacobian ile uygun biçimde dönüştürülürse aynı fiziksel bölgenin olasılığı değişmez.

Kanıt. x koordinatlarında bölge olasılığı $\int_A f_X(x) dx$ olsun. $y = g(x)$ dönüşümünde $dy = |J_g(x)| dx$ ve yeni yoğunluk $f_Y(y) = f_X(x)/|J_g(x)|$ olur. Böylece

$$\int_{g(A)} f_Y(y) dy = \int_A \frac{f_X(x)}{|J_g(x)|} |J_g(x)| dx = \int_A f_X(x) dx.$$

□

Örnek 1.21 (Birim Değişimi). *Uzunluk metre yerine santimetreyle yazıldığında yoğunluğun sayısal değeri değişir; fakat belirli bir fiziksel aralığın olasılığı değişmez.*

1.12 Örnek Veri Tablosu ve Yazılım Çözümleri

Kavram	Matematiksel araç	Tipik kullanım
Örnek uzay	Ω	Deney sonuçlarını belirleme
Olay	$A \subseteq \Omega$	İlgilenilen sonucu ifade etme
Olasılık ölçüsü	$P(A)$	Belirsizliği sayısallaştırma
Koşullu olasılık	$P(A B)$	Yeni bilgiyle güncelleme
Bayes teoremi	$P(B_i A)$	Gözlemden nedene gitme
Bağımsızlık	$P(A \cap B) = P(A)P(B)$	Etkisiz deneyleri ayırma

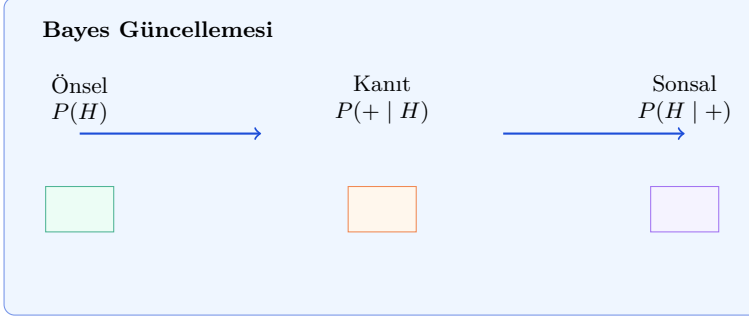
Python Çözümü

```
pH = 0.01
sens = 0.95
falsePos = 0.05
posterior = sens * pH / (sens * pH + falsePos * (1 - pH))
print(posterior)
```

R ile Hesaplama:

```
pH <- 0.01
sens <- 0.95
falsePos <- 0.05
posterior <- sens * pH / (sens * pH + falsePos * (1 - pH))
posterior
```

1.13 Görselleştirme ve Grafikler



Şekil 1.6 Bayes teoreminde önsel bilgi, kanıt ile birleşerek sonsal olasılığa dönüşür.

Dikkat: Kavram Yanılgısı

Ayrık olaylar ile bağımsız olaylar aynı şey değildir. Ayrıklık, iki olayın aynı anda gerçekleşmemesi; bağımsızlık ise bir olayın bilgisinin diğerinin olasılığını değiştirmemesi demektir. $P(A) > 0$ ve $P(B) > 0$ iken ayrık olaylar bağımsız olamaz.

Kendini Dene:

Ara Kontrol: Bir testin duyarlılığı çok yüksek olsa bile hastalık çok nadirse pozitif sonucun güvenilirliği neden düşük kalabilir? İpucu: Bayes paydasındaki yanlış pozitif terimini ayrı düşünün.

Mini-Proje / Vaka Çalışması:

COVID-19 hızlı test yorumu: Bir tarama testinin duyarlılığı 0.90, yanlış pozitif oranı 0.04 ve toplumdaki yaygınlık 0.02 olsun. Pozitif test sonrası hastalık olasılığını Bayes teoremiyle hesaplayın; sonra yaygınlık 0.20 olduğunda sonucu yeniden bulun. Aynı testin farklı toplumlarda neden farklı yorumlanacağını kısa bir paragrafta açıklayın.

1.14 Bölüm Sonu Çözümlü Problemler

1. $(\Omega, \mathcal{F}, P)$ bir olasılık uzayı olsun. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ eşitliğini ispatlayınız.

Çözüm. $A \cup B = A \cup (B \cap A^c)$ ayrık birleşimdir. Bu nedenle $P(A \cup B) = P(A) + P(B \cap A^c)$ olur. Ayrıca $B = (A \cap B) \cup (B \cap A^c)$ ayrık birleşiminden $P(B \cap A^c) = P(B) - P(A \cap B)$ elde edilir. Yerine yazınca sonuç çıkar.

2. Bir zar atıldığında çift veya 5'ten büyük gelme olasılığı nedir?

Çözüm. $A = \{2, 4, 6\}$, $B = \{6\}$ ve $A \cup B = \{2, 4, 6\}$ olduğundan olasılık $3/6 = 1/2$ olur.

3. Bir hastalığın önsel olasılığı 0.02, test duyarlılığı 0.90 ve yanlış pozitif oranı 0.10 ise pozitif test sonrası hastalık olasılığı nedir?

Çözüm.

$$P(H | +) = \frac{0.90(0.02)}{0.90(0.02) + 0.10(0.98)} = \frac{0.018}{0.116} \approx 0.155.$$

4. $P(A) = 0.4$, $P(B) = 0.5$ ve A, B bağımsız ise $P(A \cup B)$ değerini bulunuz.

Çözüm. Bağımsızlıktan $P(A \cap B) = 0.4 \cdot 0.5 = 0.2$ olur. Dolayısıyla $P(A \cup B) = 0.4 + 0.5 - 0.2 = 0.7$.

5. $A \subseteq B$ ise $P(A) \leq P(B)$ olduğunu gösteriniz.

Çözüm. $B = A \cup (B \cap A^c)$ ayrık birleşimdir. Bu nedenle $P(B) = P(A) + P(B \cap A^c) \geq P(A)$ olur.

6. Üç olay için $P(A \cup B \cup C)$ açılımını yazınız ve $P(A_i) = 0.5$, ikişerli kesişimler 0.2, üçlü kesişim 0.05 iken değerini bulunuz.

Çözüm. İçerme-dışlama ile $P(A \cup B \cup C) = 3(0.5) - 3(0.2) + 0.05 = 0.95$. Üç olay örnek uzayın büyük bölümünü kaplamaktadır.

7. 30 kişilik bir sınıfta en az iki öğrencinin aynı doğum gününe sahip olma olasılığının 0.5'ten büyük olduğunu doğum günü formülüyle açıklayınız.

Çözüm. $P(\text{çakışma}) = 1 - \prod_{j=0}^{29} (1 - j/365) \approx 1 - 0.294 = 0.706 > 0.5$. Karesel büyüyen çift sayısı çakışmayı olası kılar.

8. Boole eşitsizliğini kullanarak, her biri 0.001 hata olasılığına sahip 100 bağımsız bileşenden en az birinde hata olma olasılığına bir üst sınır veriniz.

Çözüm. $P(\bigcup_i A_i) \leq \sum_i P(A_i) = 100 \times 0.001 = 0.1$. Gerçek değer $1 - (0.999)^{100} \approx 0.095$ olup sınır ile uyumludur.

9. Güvenilirliği 0.8 olan üç bileşen önce seri, sonra paralel bağlanırsa sistem güvenilirliklerini karşılaştırınız.

Çözüm. Seri: $0.8^3 = 0.512$. Paralel: $1 - (0.2)^3 = 0.992$. Paralel tasarım güvenilirliği belirgin biçimde artırır.

10. Bir torbada 5 kırmızı, 3 mavi top vardır. İadesiz iki top çekiliyor. İkinci topun mavi olma olasılığını toplam olasılık yasasıyla bulunuz.

Çözüm. İlk top mavi (3/8) ya da kırmızı (5/8) olabilir:

$$P(2. \text{ mavi}) = \frac{3}{8} \cdot \frac{2}{7} + \frac{5}{8} \cdot \frac{3}{7} = \frac{6 + 15}{56} = \frac{21}{56} = \frac{3}{8}.$$

İkinci çekilişin kenar olasılığı, ilk çekilişin marjinaliyle aynıdır; bu, değiştirilebilirlik (exchangeability) ilkesinin bir örneğidir.

11. Bir nadir hastalık için ($P(H) = 0.005$) duyarlılık 0.99 ve yanlış pozitif oranı 0.02'dir. Pozitif test sonrası hastalık olasılığını oran biçiminde hesaplayınız.

Çözüm. Önsel oran $0.005/0.995$, olabilirlik oranı $0.99/0.02 = 49.5$. Sonsal oran ≈ 0.2487 , dolayısıyla $P(H | +) \approx 0.2487/1.2487 \approx 0.199$.

12. $A_n \downarrow A$ ve $P(A_1) < \infty$ ise $P(A) = \lim_n P(A_n)$ olduğunu gösteriniz.

Çözüm. $A_n^c \uparrow A^c$ olur; alttan süreklilikle $P(A_n^c) \uparrow P(A^c)$. Tümleyen kurallıyla $1 - P(A_n) \uparrow 1 - P(A)$, yani $P(A_n) \downarrow P(A)$.

1.15 Bölüm Sonu Alıştırmaları

Bu alıştırmalar, bölümdeki temel kavramları kademeli biçimde pekiştirmek için üç düzeyde düzenlenmiştir. Cevaplar ve kısa çözüm ipuçları Ekler bölümündeki *Cevap Anahtarı*'nda verilmiştir.

Temel Düzey

- B1. Bir madeni para iki kez atılıyor. Örnek uzayı yazınız ve en az bir yazı gelme olayını küme olarak gösteriniz.
- B2. Bir zar atılıyor. $A = \{1, 3, 5\}$ ve $B = \{4, 5, 6\}$ için $A \cup B$, $A \cap B$ ve A^c kümelerini bulunuz.
- B3. $P(A) = 0.35$, $P(B) = 0.50$ ve $P(A \cap B) = 0.20$ ise $P(A \cup B)$ değerini hesaplayınız.
- B4. $A \subseteq B$, $P(A) = 0.28$ ve $P(B) = 0.62$ ise $P(B \cap A^c)$ değerini bulunuz.
- B5. Bir sınıfta öğrencilerin 0.40'ı istatistik, 0.25'i programlama, 0.10'u her iki dersi almaktadır. Rastgele seçilen bir öğrencinin bu derslerden en az birini alma olasılığı nedir?

Orta Düzey

- I1. Bir kutuda 4 beyaz ve 6 siyah top vardır. İadesiz iki top çekiliyor. İki topun da beyaz olma olasılığını bulunuz.
- I2. $P(A) = 0.6$, $P(B) = 0.5$ ve $P(A \cup B) = 0.8$ ise A ve B bağımsız mıdır? Gerekçenizi yazınız.
- I3. Bir test için duyarlılık 0.92, yanlış pozitif oranı 0.06 ve hastalık yaygınlığı 0.03 olsun. Pozitif test sonrası hastalık olasılığını hesaplayınız.
- I4. Üç olay için $P(A) = 0.40$, $P(B) = 0.35$, $P(C) = 0.30$, ikişerli kesişimler sırasıyla 0.12, 0.10, 0.08 ve üçlü kesişim 0.03 ise $P(A \cup B \cup C)$ değerini bulunuz.
- I5. Bir sistem iki bağımsız bileşenin paralel bağlanmasıyla çalışıyor. Bileşen güvenilirlikleri 0.90 ve 0.80 ise sistem güvenilirliğini hesaplayınız.

İleri Düzey

- A1. $A_n \uparrow A$ ise $P(A) = \lim_{n \rightarrow \infty} P(A_n)$ olduğunu Kolmogorov aksiyomlarından hareketle gösteriniz.
- A2. $P(A) > 0$ ve $P(B) > 0$ olan iki ayrık olayın bağımsız olamayacağını ispatlayınız.
- A3. 24 kişilik bir grupta en az iki kişinin aynı doğum gününe sahip olma olasılığını doğum günü formülüyle yaklaşık hesaplayınız. Artık yılları ihmal ediniz.
- A4. Bir nadir olayın önsel olasılığı 0.01'dir. Kanıtın olabilirlik oranı $LR = 20$ ise sonsal olasılığı oran biçimiyle hesaplayınız.
- A5. $P(A_i) \leq 0.002$ olan 250 olay için $P(\bigcup_{i=1}^{250} A_i)$ değerine Boole eşitsizliğiyle üst sınır veriniz. Bu sınırın bağımsızlık varsayımı gerektirip gerektirmediğini belirtiniz.

1.16 Kaynak Notu

Bu bölümdeki temel aksiyomatik yapı Kolmogorov'un ölçü temelli olasılık yaklaşımı üzerine kurulmuştur (Kolmogorov, 1956). Sonlu örnek uzay, sayma, koşullu olasılık ve Bayes anlatımı klasik giriş metinleriyle uyumlu biçimde düzenlenmiştir (Grinstead & Snell, 1997; Hogg et al., 2019; Hsu, 1997; Ross, 2019). İçerme-dışlama, doğum günü ve eşleşme problemleri ile ikişerli/karşılıklı bağımsızlık ayrımı Feller (1968) çizgisindedir; olasılığın sürekliliği ve σ -cebiri limitleri Billingsley (1995) ve Çınlar (2011) ölçü-kuramsal çerçevesiyle uyumludur. Mühendislik ve güvenilirlik yorumları Soong (2004) ile Montgomery ve Runger (2018), ölçüm ve koordinat duyarlılığı Tarantola (2001), bilgi temelli olasılık yorumu Jaynes (2003), ürün uzayı ve ölçü dili ise Capiński ve Kopp (2004) ile Shao (2003) çizgisinde verilmiştir.

Özet:

Olasılık hesabının temeli, rassal deneyi doğru tanımlamak ve olayları küme diliyle ifade etmektir. Kolmogorov aksiyomları olasılık ölçüsünün tutarlı kurallarını verir. Koşullu olasılık yeni bilgiyle güncellemeyi, Bayes teoremi gözlemden nedene ters akıl yürütmeyi, bağımsızlık ise olaylar arasında etkisizlik ilişkisini açıklar.

Bölüm 2

Kombinatorik Analiz ve Sayma Yöntemleri

2.1 Gerçek Dünya Problemi (Giriş)

Bir şifre sisteminde kaç farklı parola üretilebileceğini, bir kalite kontrol sürecinde kaç farklı örneklem seçilebileceğini, bir kart oyununda belirli bir elin kaç biçimde oluşacağını veya bir kurulun kaç farklı üyeden kurulabileceğini hesaplamak olasılık modelinin ilk adımıdır. Eğer örnek uzaydaki sonuç sayısı doğru sayılmazsa olasılık da doğru hesaplanamaz.

Kombinatorik analiz, sonlu örnek uzaylarda sonuçları sistemli biçimde sayma yöntemidir. Olasılıkta özellikle eş olasılıklı modeller, kart ve zar problemleri, binom katsayıları, hipergeometrik dağılım, binom dağılımı ve örnekleme teorisi bu sayma ilkelerine dayanır.

Bu Bölüm Ne Sağlar?

Bu bölüm toplama ve çarpma kurallarıyla başlar; faktöriyel, permütasyon, tekrarlı permütasyon, kombinasyon, binom katsayıları, binom teoremi, multinom katsayıları, yerine koymalı ve yerine koymasız seçim modellerini açıklar. Amaç, sayma formüllerini ezberletmek değil; sıralı-sırasız ve tekrarlı-tekrarsız seçim ayırımını olasılık modelinin merkezine yerleştirmektir.

Amaç ve Öğrenme Çıktıları

Bu bölümün sonunda okuyucu

- toplama ve çarpma kurallarını örnek uzay sayımında kullanabilir,
- permütasyon ve kombinasyon arasındaki farkı açıklayabilir,
- tekrarlı ve tekrarsız seçim modellerini ayırt edebilir,
- binom ve multinom katsayılarını yorumlayabilir,
- eş olasılıklı sonlu modellerde olasılıkları sayma yoluyla hesaplayabilir,
- binom ve hipergeometrik dağılımların kombinatorik temelini kurabilir.

2.2 Sayma Probleminin Olasılıktaki Yeri

Sonlu ve eş olasılıklı bir örnek uzayda olasılık, uygun sonuç sayısının tüm sonuç sayısına oranıdır. Bu nedenle klasik olasılık problemlerinde asıl güçlük çoğu zaman olasılık formülünde değil, doğru sayma yöntemini seçmektir.

Tanım 2.1 (Eş Olasılıklı Sonlu Model) Ω sonlu ve her sonuç eş olasılıklı ise $A \subseteq \Omega$ olayı için

$$P(A) = \frac{|A|}{|\Omega|}$$

olur. Burada $|A|$ olayın içerdiği sonuç sayısını, $|\Omega|$ ise örnek uzayın toplam sonuç sayısını gösterir.

Örnek 2.1 (Zar Atışı). *Adil bir zar iki kez atılsın. Örnek uzay büyüklüğü $6 \cdot 6 = 36$ olur. Toplamın 7 olması olayı*

$$(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)$$

sonuçlarından oluşur. Bu nedenle olasılık $6/36 = 1/6$ olur.

Sayma problemlerinde en yaygın hata, sıralı sonuçlarla sırasız seçimleri karıştırmaktır. (A, B) ile (B, A) farklı kabul ediliyorsa sıralı; aynı kabul ediliyorsa sırasız sayma yapılmalıdır.

2.3 Toplama ve Çarpma Kuralları

Kombinatorik analizin iki temel ilkesi toplama ve çarpma kurallarıdır. Toplama kuralı alternatif durumları, çarpma kuralı ise ardışık seçimleri sayar.

Teorem 2.1 (Toplama Kuralı). *Ayrık iki işlemden biri yapılacaksa ve birinci işlem m , ikinci işlem n farklı biçimde yapılabiliyorsa toplam seçenek sayısı*

$$m + n$$

olur. Daha genel olarak ayrık durumların sayıları toplanır.

Kanıt. Ayrık durumlar birlikte gerçekleşmediği için toplam seçenek kümesi ayrık birleşimdir. Ayrık kümelerin eleman sayısı, eleman sayılarının toplamıdır. $\square$

Teorem 2.2 (Çarpma Kuralı). *Bir işlem önce m farklı biçimde, ardından her bir ilk seçim için n farklı biçimde tamamlanabiliyorsa toplam seçenek sayısı*

$$mn$$

olur. Birden fazla aşama için aşama sayıları çarpılır.

Kanıt. Her ilk seçim, ikinci aşamadaki n seçenekle eşleşir. Böylece m grup ve her grupta n sonuç vardır; toplam mn olur. $\square$

Örnek 2.2 (Parola Sayısı). *Bir parola üç harf ve iki rakamdan oluşsun. Harfler için 26, rakamlar için 10 seçenek ve tekrar serbest olsun. Biçim sabit olarak HHH RR ise seçenek sayısı*

$$26^3 \cdot 10^2$$

olur.

2.4 Faktöriyel ve Permütasyon

Permütasyon, nesnelerin sıralanmasını sayar. Sıra önemliyse permütasyon mantığı kullanılır. Faktöriyel, tüm nesnelerin sıralanma sayısını ifade eder.

Tanım 2.2 (Faktöriyel) Pozitif tam sayı n için

$$n! = n(n-1)(n-2) \cdots 2 \cdot 1$$

olarak tanımlanır. Ayrıca $0! = 1$ kabul edilir.

Faktöriyel çok hızlı büyür ve büyük n için doğrudan hesaplanması zorlaşır. Asimptotik davranışını veren Stirling yaklaşımı, olasılıkta binom katsayılarının ve normal yaklaşımın analizinde merkezî rol oynar (Feller, 1968).

Teorem 2.3 (Stirling Yaklaşımı). $n \rightarrow \infty$ iken

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n,$$

yani iki tarafın oranı 1'e yaklaşır. Daha keskin biçimde $n! = \sqrt{2\pi n} (n/e)^n e^{\theta_n}$ olup $\frac{1}{12n+1} < \theta_n < \frac{1}{12n}$ sağlanır.

Örnek 2.3 (Stirling ile Tahmin). $10! = 3.6288 \times 10^6$ 'dır. Stirling tahmini $\sqrt{20\pi} (10/e)^{10} \approx 3.5987 \times 10^6$ olup yaklaşık %0.8 hata verir. Hata n büyüdükçe hızla küçülür.

Tanım 2.3 (Permütasyon) n farklı nesnenin tamamının sıralanma sayısı

$$n!$$

olur. n farklı nesneden r tanesi seçilip sıralanacaksa permütasyon sayısı

$$P(n, r) = \frac{n!}{(n-r)!}$$

şeklindedir.

Kanıt. İlk konuma n , ikinci konuma $n - 1$, r 'inci konuma $n - r + 1$ seçenek vardır. Çarpma kuralıyla

$$n(n - 1) \cdots (n - r + 1) = \frac{n!}{(n - r)!}$$

elde edilir. $\square$

Örnek 2.4 (Yarış Sıralaması). *10 koşucudan ilk üç derece sıralı biçimde belirlenecekse sonuç sayısı*

$$P(10, 3) = 10 \cdot 9 \cdot 8 = 720$$

olur.

Bazı sıralama problemlerinde nesneler bir doğru üzerinde değil, bir çember etrafında dizilir. Bu durumda tüm dizilişi döndürmek yeni bir sıralama üretmediği için sayım farklılaşır.

Teorem 2.4 (Dairesel Permütasyon). *n farklı nesnenin bir çember etrafında ayırt edilebilir diziliş sayısı*

$$(n - 1)!$$

olur. Diziliş aynalama (saat yönü/tersi) altında da aynı sayılıyorsa sayı $(n - 1)!/2$ 'ye iner.

Kanıt. n nesnenin doğrusal $n!$ dizilişinin her biri, çember üzerinde n farklı döndürme ile aynı dizilişi verir. Bu nedenle $n!/n = (n - 1)!$ ayırt edilebilir dairesel diziliş kalır. Aynalama da serbestse her diziliş ek olarak yansımısıyla eşleşir ve sayı ikiye bölünür. $\square$

Örnek 2.5 (Yuvarlak Masa). *8 kişi yuvarlak bir masaya oturacaksa farklı oturma düzeni sayısı $(8 - 1)! = 5040$ olur. Eğer iki belirli kişinin yan yana oturması isteniyorsa, bu ikiliyi tek bir blok sayıp blok içinde $2!$ sıralama eklenir: $(7 - 1)! \cdot 2! = 1440$.*

2.5 Tekrarlı Permütasyon

Bazı sıralama problemlerinde nesneler birbirinden ayırt edilemez. Aynı türden nesnelerin yer değiştirmesi yeni bir sıralama üretmez. Bu durumda tekrarlı permütasyon kullanılır.

Teorem 2.5 (Tekrarlı Permütasyon). *Toplam n nesnenin n_1 tanesi birinci tür, n_2 tanesi ikinci tür, ..., n_k tanesi k 'ıncı tür ve*

$$n_1 + n_2 + \cdots + n_k = n$$

olsun. Ayırt edilebilir sıralama sayısı

$$\frac{n!}{n_1!n_2!\cdots n_k!}$$

şeklindedir.

Kant. Tüm nesneler ayırt edilebilir olsaydı $n!$ sıralama olurdu. Aynı türden n_i nesnenin kendi aralarındaki $n_i!$ yer değiştirmesi aynı sıralamayı verdiği için bu tekrarlar bölünür. $\square$

Örnek 2.6 (KELİME Sıralaması). “ANANAS” kelimesinde 6 harf vardır; A harfi 3, N harfi 2 kez tekrar eder. Farklı harf dizilişi sayısı

$$\frac{6!}{3!2!} = 60$$

olur.

2.6 Kombinasyon

Kombinasyon, sırasız seçimleri sayar. Bir kurul, komite, el veya örneklem seçiliyorsa ve seçim sırası önemli değilse kombinasyon kullanılır.

Tanım 2.4 (Kombinasyon) n farklı nesneden r tanesinin sırasız seçilme sayısı

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

şeklindedir.

Kant. Önce r nesne sıralı seçilirse $P(n, r) = n!/(n-r)!$ sonuç vardır. Ancak seçilen aynı r nesne kendi arasında $r!$ farklı sırayla sayılmıştır. Bu tekrarlar bölünür ve sonuç elde edilir. $\square$

Örnek 2.7 (Komite Seçimi). 12 kişiden 4 kişilik bir komite seçilecekse sıra önemli değildir. Seçim sayısı

$$\binom{12}{4} = 495$$

olur.

Binom katsayıları, sayma argümanlarıyla ispatlanan zengin bir özdeşlik ailesine sahiptir. Bu özdeşlikler hem cebirsel sadeleştirmede hem de olasılık toplamalarında kullanılır (Feller, 1968; Ross, 2019).

Theorem 2.6 (Simetri ve Pascal Kuralı). $0 \leq r \leq n$ için

$$\binom{n}{r} = \binom{n}{n-r}, \quad \binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r}.$$

Kant. Simetri, r nesne seçmenin dışarıda bırakılacak $n - r$ nesneyi seçmeye denk olmasından çıkar. Pascal kuralı için belirli bir nesneyi işaretleyelim: r 'li seçimler ya bu nesneyi içerir (kalan $r - 1$ nesne $n - 1$ arasından: $\binom{n-1}{r-1}$) ya da içermez (r nesne $n - 1$ arasından: $\binom{n-1}{r}$). İki ayrık durum toplanır. $\square$

Teorem 2.7 (Alt Küme Sayımı ve Vandermonde Özdeşliği).

$$\sum_{k=0}^n \binom{n}{k} = 2^n, \quad \sum_{k=0}^r \binom{m}{k} \binom{n}{r-k} = \binom{m+n}{r}.$$

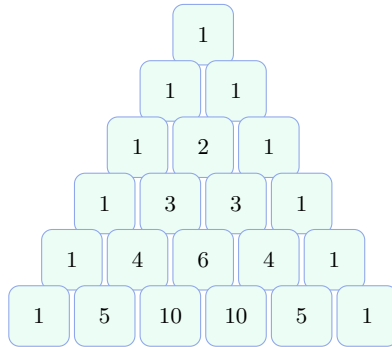
Kant. İlk eşitlik, n elemanlı kümenin tüm alt kümelerinin sayısını iki biçimde saymaktan gelir: her eleman için “içinde/dışında” ikili seçimi 2^n , büyüklüğe göre toplam ise $\sum_k \binom{n}{k}$ verir. İkincisi (Vandermonde): $m + n$ kişilik gruptan r kişi seçmenin bir yolu, ilk m 'den k ve kalan n 'den $r - k$ kişi seçmektir; k üzerinden toplanır. $\square$

Teorem 2.8 (Hokey Sopası Özdeşliği). $0 \leq r \leq n$ için

$$\sum_{k=r}^n \binom{k}{r} = \binom{n+1}{r+1}.$$

Kant. Pascal kuralı $\binom{k}{r} = \binom{k+1}{r+1} - \binom{k}{r+1}$ biçiminde yazılırsa toplam teleskopiktir; ardışık terimler sadeleşir ve geriye $\binom{n+1}{r+1} - \binom{r}{r+1} = \binom{n+1}{r+1}$ kalır. $\square$

Örnek 2.8 (Pascal Kuralının Kullanımı). $\binom{6}{2} = \binom{5}{1} + \binom{5}{2} = 5 + 10 = 15$ olur. Bu yineleme, binom katsayılarını çarpan hesabı yapmadan satır satır üretmenin temelidir.



Şekil 2.1 Pascal üçgeni: her sayı üstündeki iki sayının toplamıdır (Pascal kuralı). n 'inci satır $(a + b)^n$ açılımının katsayılarını verir.

2.7 Sıralı–Sırasız ve Tekrarlı–Tekrarsız Seçim

Sayma problemlerinde dört temel seçim türü vardır. Doğru formül, sıra ve tekrar durumuna göre seçilir.

	Tekrar Yok	Tekrar Var
Sıralı	$\frac{n!}{(n-r)!}$	n^r
Sırasız	$\binom{n}{r}$	$\binom{n+r-1}{r}$

Tanım 2.5 (Tekrarlı Kombinasyon) n tür nesneden tekrar serbest olacak biçimde r tane sırasız seçim yapma sayısı

$$\binom{n+r-1}{r}$$

şeklindedir.

Örnek 2.9 (Dondurma Seçimi). *5 farklı dondurma türünden toplam 3 top seçilecek ve aynı türden birden fazla alınabilecekse seçim sayısı*

$$\binom{5+3-1}{3} = \binom{7}{3} = 35$$

olur.

Tekrarlı kombinasyon formülünün arkasındaki sayma argümanı, “yıldız ve çubuk” yöntemi olarak bilinir ve birçok dağıtım/bölüştürme problemine genelleşir.

Teorem 2.9 (Yıldız ve Çubuk Yöntemi). $x_1 + x_2 + \cdots + x_n = r$ denkleminin negatif olmayan tam sayı çözümlerinin sayısı

$$\binom{n+r-1}{r} = \binom{n+r-1}{n-1}$$

olur.

Kanıt. r adet özdeş yıldız, n kategoriye ayırmak için $n-1$ çubuk kullanılır. Yıldızlar ve çubuklar bir satırda dizilir; toplam $r + (n-1)$ konumdan çubukların yerleştirileceği $n-1$ konum seçmek bir çözümü belirler. Bu seçim $\binom{n+r-1}{n-1}$ biçiminde sayılır. $\square$

Örnek 2.10 (Bonbon Dağıtımı). *9 özdeş bonbon 4 çocuğa, her çocuğa hiç verilmeyebilecek biçimde kaç türlü dağıtılır? Bu, $x_1 + x_2 + x_3 + x_4 = 9$ denkleminin çözüm sayısıdır:*

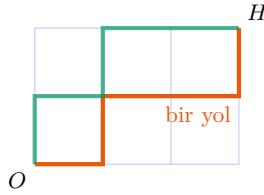
$$\binom{9+4-1}{4-1} = \binom{12}{3} = 220.$$

Her çocuğa en az bir bonbon verilmesi istenirse önce her birine birer bonbon verilir ($x_i \geq 1$), kalan 5 bonbon için $\binom{5+4-1}{3} = \binom{8}{3} = 56$ bulunur.

Örnek 2.11 (Kafes Yolları). Bir ızgarada sol alttan (O) sağ üste (H) yalnız sağa ($\rightarrow$) ve yukarı ($\uparrow$) adımlarla gidiliyor. m sağ ve k yukarı adım gerekiyorsa toplam yol sayısı, $m + k$ adımın hangilerinin “sağ” olduğunu seçmektir:

$$\binom{m+k}{m}.$$

Aşağıdaki 3×2 ızgarada ($m = 3, k = 2$) yol sayısı $\binom{5}{3} = 10$ olur.



Şekil 2.2 Kafes yolları: yalnız sağa ve yukarı adımlarla $O \rightarrow H$ yol sayısı $\binom{m+k}{m}$. İki örnek yol gösterilmiştir.

2.8 Binom Katsayıları ve Binom Teoremi

Binom katsayıları, iki seçenekli seçimlerde kaç yol olduğunu gösterir. Bernoulli denemeleri ve binom dağılımı doğrudan bu katsayılara dayanır.

Teorem 2.10 (Binom Teoremi). n negatif olmayan tam sayı olmak üzere

$$(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}$$

geçerlidir.

Kanıt. $(a + b)^n$ ifadesi n parantezin çarpımıdır. Açılımda $a^k b^{n-k}$ terimi, n parantezin k tanesinden a , kalanlarından b seçilmesiyle oluşur. Bu seçim sayısı $\binom{n}{k}$ olduğundan sonuç elde edilir. $\square$

Örnek 2.12 (Binom Dağılımının Normalizasyonu). $0 \leq p \leq 1$ için

$$\sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} = (p + 1 - p)^n = 1$$

olur. Bu, binom olasılıklarının toplamının 1 olduğunu gösterir.

2.9 Multinom Katsayıları

İkiden fazla kategori olduğunda binom katsayıları multinom katsayılarına genişler. Bu yapı çok kategorili sayım ve multinom dağılım için temel oluşturur.

Tanım 2.6 (Multinom Katsayısı) $n_1 + \dots + n_k = n$ ise

$$\binom{n}{n_1, n_2, \dots, n_k} = \frac{n!}{n_1! n_2! \dots n_k!}$$

ifadesine multinom katsayısı denir.

Teorem 2.11 (Multinom Teoremi).

$$(x_1 + x_2 + \dots + x_k)^n = \sum_{n_1 + \dots + n_k = n} \frac{n!}{n_1! \dots n_k!} x_1^{n_1} \dots x_k^{n_k}.$$

Kanıt. n parantezin her birinden $x_1, \dots, x_k$ terimlerinden biri seçilir. Belirli sayılarda $n_1, \dots, n_k$ seçim yapılmışsa bu düzenlemelerin sayısı tekrarlı permütasyon formülüyle $n!/(n_1! \dots n_k!)$ olur. $\square$

Örnek 2.13 (Üç Kategorili Anket). *Bir ankette cevaplar evet, hayır ve kararsız olsun. 20 kişide sırasıyla 9, 7 ve 4 cevap görülmesinin kaç farklı sıralı cevap dizisiyle oluşabileceği*

$$\frac{20!}{9!7!4!}$$

ile sayılır.

2.10 İçerme–Dışlama ile Sayma

İçerme–dışlama ilkesi yalnız olasılıkta değil, “en az bir kısıtın ihlal edilmediği” düzenlemeleri saymakta da güçlüdür. Tipik uygulaması, örten (surjektif) fonksiyonların ve yasaklı konum problemlerinin sayımıdır (Feller, 1968).

Teorem 2.12 (Örten Fonksiyon Sayısı). n elemanlı kümeden k elemanlı kümeye tanımlı örten (her hedefin en az bir kez kullanıldığı) fonksiyon sayısı

$$\sum_{j=0}^k (-1)^j \binom{k}{j} (k-j)^n$$

olur.

Kanıt. Tüm fonksiyonlar k^n tanedir. A_i , i 'inci hedefin hiç kullanılmadığı fonksiyonlar olsun. Örten fonksiyonlar $\bigcap_i A_i^c$ 'dir. İçerme–dışlama ile $|\bigcap_i A_i^c| =$

$\sum_j (-1)^j \binom{k}{j} (k-j)^n$ elde edilir; çünkü belirli j hedefi dışlayan fonksiyonlar $(k-j)^n$ tanedir. $\square$

Örnek 2.14 (Hediye Dağıtımı). *5 farklı hediye 3 çocuğa, her çocuk en az bir hediye alacak biçimde kaç türlü dağıtılır? Bu, 5 elemandan 3 elemana örten fonksiyon sayısıdır:*

$$3^5 - \binom{3}{1} 2^5 + \binom{3}{2} 1^5 = 243 - 96 + 3 = 150.$$

2.11 Yerine Koymalı ve Yerine Koymasız Örneklemeye

Örneklemeye problemlerinde seçilen birimin tekrar seçilip seçilemeyeceği önemlidir. Yerine koymalı seçim bağımsızlığı korur; yerine koymasız seçim bağımlılık yaratır.

Tanım 2.7 (Yerine Koymalı Seçim) Her seçimden sonra seçilen birim ana-kütleye geri konuyorsa yerine koymalı seçim yapılır. n seçenekten r sıralı seçim sayısı n^r olur.

Tanım 2.8 (Yerine Koymasız Seçim) Seçilen birim tekrar seçilemiyorsa yerine koymasız seçim yapılır. Sıralı seçim sayısı $P(n, r)$, sırasız seçim sayısı $\binom{n}{r}$ olur.

Örnek 2.15 (Kart Çekimi). *52 kartlık desteden 5 kartlık el seçiliyorsa sıra önemli değildir ve yerine koymasız seçim vardır. Tüm el sayısı*

$$\binom{52}{5}$$

olur.

2.12 Olasılık Problemlerinde Sayma

Kombinatorik analiz olasılık hesabına doğrudan uygulanır. Önce örnek uzay sayılır, sonra olayın uygun sonuçları sayılır ve oran alınır.

Örnek 2.16 (Pokerde Aynı Türden Beş Kart). *52 kartlık desteden 5 kartlık el seçilsin. Tüm el sayısı $\binom{52}{5}$ olur. Aynı türden beş kart için önce 4 türden biri seçilir, sonra o türden 5 kart seçilir:*

$$4 \binom{13}{5}.$$

Olasılık

$$\frac{4 \binom{13}{5}}{\binom{52}{5}}$$

şeklindedir.

Örnek 2.17 (Hipergeometrik Mantık). *Bir kutuda 8 sağlam ve 2 kusurlu ürün olsun. İadesiz 3 ürün seçildiğinde tam 1 kusurlu gelme olasılığı*

$$\frac{\binom{2}{1} \binom{8}{2}}{\binom{10}{3}}$$

olur. Bu yapı hipergeometrik dağılımın temelidir.

Örnek 2.18 (Full House Olasılığı). *Pokerde “full house” (bir değerden üç, başka bir değerden iki kart) olasılığını hesaplayalım. Üçlünün değeri 13 türden, kartları $\binom{4}{3}$; ikilinin değeri kalan 12 türden, kartları $\binom{4}{2}$ biçiminde seçilir:*

$$P(\text{full house}) = \frac{13 \binom{4}{3} \cdot 12 \binom{4}{2}}{\binom{52}{5}} = \frac{3744}{2598960} \approx 0.00144.$$

Örnek 2.19 (Sayısal Loto). *49 sayıdan 6’sının çekildiği bir loto oyununda, oynanan tek kuponun tam k sayı tutturma olasılığı hipergeometrik biçimdedir:*

$$P(k \text{ tutturma}) = \frac{\binom{6}{k} \binom{43}{6-k}}{\binom{49}{6}}.$$

Altı sayının tamamını tutturma olasılığı $1/\binom{49}{6} = 1/13\,983\,816$ olur.

2.13 Dağılımlara Hazırlık: Binom ve Hipergeometrik Yapı

Kombinatorik analiz bazı özel dağılımların neden bu formda olduğunu açıklar. Binom dağılımı bağımsız Bernoulli dizilerindeki başarı konumlarını, hipergeometrik dağılım ise iadesiz seçimde başarıların örneklem içindeki yerleşimini sayar.

Teorem 2.13 (Binom Olasılığının Sayma Yorumu). *Bağımsız n Bernoulli denemesinde başarı olasılığı p olsun. Tam k başarı olasılığı*

$$\binom{n}{k} p^k (1-p)^{n-k}$$

şeklindedir.

Kant. Tam k başarı içeren dizilerin sayısı $\binom{n}{k}$ olur. Her belirli dizinin olasılığı bağımsızlıkla $p^k (1-p)^{n-k}$ olduğundan sonuç elde edilir. $\square$

Teorem 2.14 (Hipergeometrik Olasılığın Sayma Yorumu). *N birimde K başarı bulunsun. İadesiz n seçimde x başarı elde etme olasılığı*

$$\frac{\binom{K}{x} \binom{N-K}{n-x}}{\binom{N}{n}}$$

şeklindedir.

Kanıt. Payda tüm n birimlik örneklem sayısıdır. Payda ise K başarıdan x tanesini ve $N - K$ başarısızdan $n - x$ tanesini seçme sayısı vardır. Eş olasılıklı seçimde oran olasılığı verir. $\square$

Üçüncü temel sayma temelli yapı, sabit sayıda başarıya ulaşana kadar gereken deneme sayısını sayar. Bu, negatif binom (Pascal) dağılımının kombinatorik temelidir.

Teorem 2.15 (Negatif Binom Olasılığının Sayma Yorumu). *Bağımsız Bernoulli denemelerinde başarı olasılığı p olsun. r 'inci başarının tam n 'inci denemede gerçekleşmesi olasılığı*

$$\binom{n-1}{r-1} p^r (1-p)^{n-r}$$

olur.

Kanıt. n 'inci deneme başarı olmalıdır (olasılık p). Önceki $n - 1$ deneme içinde tam $r - 1$ başarı bulunmalıdır; bu başarıların konumları $\binom{n-1}{r-1}$ biçiminde seçilir ve her belirli dizinin olasılığı $p^{r-1}(1-p)^{n-r}$ 'dir. Çarpım sonucu verir. $\square$

Örnek 2.20 (Üçüncü Başarıya Kadar). *Başarı olasılığı $p = 0.4$ olan denemelerde üçüncü başarının tam 5. denemede gelmesi olasılığı*

$$\binom{4}{2} (0.4)^3 (0.6)^2 = 6 \cdot 0.064 \cdot 0.36 \approx 0.138.$$

2.14 Örnek Veri Tablosu ve Yazılım Çözümleri

Aşağıdaki tablo, dört temel seçim modelini küçük bir örnek üzerinden karşılaştırır. $n = 5$ nesneden $r = 3$ seçim yapılsın.

Seçim Türü	Formül	Değer
Sıralı, tekrarsız	$P(5, 3)$	60
Sıralı, tekrarlı	5^3	125
Sırasız, tekrarsız	$\binom{5}{3}$	10
Sırasız, tekrarlı	$\binom{7}{3}$	35

```

choose(5, 3)
factorial(5) / factorial(5 - 3)
5^3
choose(5 + 3 - 1, 3)

```

Python Çözümü

```

from math import comb, factorial

n, r = 5, 3
siralı_tekrarsiz = factorial(n) // factorial(n-r)
siralı_tekrarlı = n**r
sirasız_tekrarsiz = comb(n, r)
sirasız_tekrarlı = comb(n+r-1, r)

print(siralı_tekrarsiz, siralı_tekrarlı)
print(sirasız_tekrarsiz, sirasız_tekrarlı)

```

2.15 Görselleştirme ve Grafikler

Binom katsayıları Pascal üçgeni üzerinde görülebilir. Her satırdaki katsayılar $(a + b)^n$ açılımının katsayılarıdır.

```

from math import comb

for n in range(8):
    row = [comb(n, k) for k in range(n + 1)]
    print(row)

```

Dikkat: Kavram Yanılgısı

Sayma problemlerinde en sık hata, sıra ve tekrar bilgisini aynı anda kontrol etmemektir. “Seçiliyor” ifadesi tek başına kombinasyon anlamına gelmez; seçim sırası kaydediliyorsa permütasyon, seçilen nesne tekrar havuza dönüşürse tekrarlı seçim modeli gerekir.

Kendini Dene:

Ara Kontrol: Beş farklı kitap üç rafa yerleştiriliyor. Rafların boş kalmasına izin veriliyorsa bu problem neden yıldız-çubuk değil, önce kitapların raf seçimi problemidir?

2.16 Bölüm Sonu Çözümlü Problemler

1. 8 kişiden başkan, başkan yardımcısı ve sekreter seçilecektir. Kaç farklı seçim yapılabilir?

Çözüm. Görevler farklı olduğu için sıra önemlidir:

$$P(8, 3) = 8 \cdot 7 \cdot 6 = 336.$$

2. 10 kişiden 4 kişilik bir ekip seçilecektir. Kaç farklı ekip vardır?

Çözüm. Sıra önemli değildir:

$$\binom{10}{4} = 210.$$

3. 6 harfli, her konumunda 29 harften biri kullanılabilen ve tekrar serbest olan kaç parola vardır?

Çözüm. Her konum için 29 seçenek vardır:

$$29^6.$$

4. “MARMARA” kelimesindeki harflerin kaç farklı dizilişi vardır?

Çözüm. Toplam 7 harf vardır. A harfi 3, M harfi 2, R harfi 2 kez geçer:

$$\frac{7!}{3!2!2!} = 210.$$

5. 52 kartlık desteden 5 kart seçiliyor. Tam 2 as gelme olasılığını yazın.

Çözüm. Toplam el sayısı $\binom{52}{5}$ olur. 4 astan 2'si ve as olmayan 48 karttan 3'ü seçilir:

$$\frac{\binom{4}{2} \binom{48}{3}}{\binom{52}{5}}.$$

6. 12 ürünün 3'ü kusurludur. İadesiz 4 ürün seçildiğinde hiç kusurlu gelmeme olasılığı nedir?

Çözüm. 9 sağlam üründen 4'ü seçilmelidir:

$$\frac{\binom{9}{4}}{\binom{12}{4}}.$$

7. 10 özdeş bilgisayar 4 farklı bölüme, bir bölüm boş kalabilecek biçimde kaç türlü dağıtılır?

Çözüm. Yıldız-çubuk: $x_1 + \dots + x_4 = 10$ çözüm sayısı

$$\binom{10+4-1}{3} = \binom{13}{3} = 286.$$

8. $\binom{n}{0} + \binom{n}{1} + \dots + \binom{n}{n} = 2^n$ olduğunu sayma argümanı ile açıklayın.

Çözüm. n elemanlı kümenin tüm alt kümeleri büyüklüğe göre $\sum_k \binom{n}{k}$, eleman bazında “içinde/dışında” seçimiyle 2^n sayılır. İki sayım eşittir.

9. Bir 4×3 ızgarada sol alttan sağ üste yalnız sağa ve yukarı adımlarla kaç yol vardır?

Çözüm. 4 sağ ve 3 yukarı adım gerekir:

$$\binom{4+3}{4} = \binom{7}{4} = 35.$$

10. 7 kişi yuvarlak bir masaya oturacak; iki belirli kişi yan yana oturmayacaktır. Kaç düzen vardır?

Çözüm. Tüm dairesel düzen $(7-1)! = 720$. İki kişinin yan yana olduğu düzen $(6-1)! \cdot 2! = 240$. İstenen: $720 - 240 = 480$.

11. 6 farklı mektup 3 farklı posta kutusuna, her kutuya en az bir mektup düşecek biçimde kaç türlü atılır?

Çözüm. Örten fonksiyon sayısı:

$$3^6 - \binom{3}{1}2^6 + \binom{3}{2}1^6 = 729 - 192 + 3 = 540.$$

2.17 Bölüm Sonu Alıştırmaları

Bu alıştırmalar, sıra-tekrar ayrımını, temel sayma ilkelerini ve kombinatorik olasılık hesaplarını pekiştirmek için üç düzeyde düzenlenmiştir. Cevaplar ve kısa çözüm ipuçları Ekler bölümündeki *Cevap Anahtarı*’nda verilmiştir.

Temel Düzey

- B1. 7 farklı kitaptan 3’ü bir rafa sıralanacaktır. Kaç farklı sıralama yapılabilir?
 B2. 12 öğrenciden 5 kişilik bir çalışma grubu seçilecektir. Kaç farklı grup seçilebilir?
 B3. 4 basamaklı ve her basamağı 0–9 arasından seçilen, tekrarın serbest olduğu kaç şifre vardır?
 B4. “ANANAS” kelimesindeki harflerin kaç farklı dizilişi vardır?
 B5. Bir menüde 3 çorba, 5 ana yemek ve 4 tatlı vardır. Her kategoriden bir seçim yapılarak kaç farklı menü oluşturulabilir?

Orta Düzey

- I1. 52 kartlık desteden 5 kart seçiliyor. Tam 3 kupa gelme olasılığını yazınız.
- I2. 9 farklı oyuncudan 4 kişilik bir ekip seçilip ekip içinde bir kaptan belirlenecektir. Kaç farklı sonuç vardır?
- I3. 8 özdeş kalem 3 farklı öğrenciye, her öğrenci en az bir kalem alacak biçimde kaç türlü dağıtılır?
- I4. Bir 5×4 ızgarada sol alttan sağ üste yalnız sağa ve yukarı adımlarla kaç yol vardır?
- I5. 10 ürünün 2'si kusurludur. İadesiz 3 ürün seçildiğinde tam 1 kusurlu ürün gelme olasılığını bulunuz.

İleri Düzey

- A1. 8 kişi yuvarlak bir masaya oturacaktır. İki belirli kişi yan yana oturacaksa kaç dairesel düzen vardır?
- A2. 7 farklı görev 3 farklı kişiye, her kişiye en az bir görev düşecek biçimde kaç türlü dağıtılır?
- A3. $(x + y + z)^6$ açılımında x^2y^3z teriminin katsayısını bulunuz.
- A4. 20 kişilik bir gruptan 6 kişilik komite seçiliyor. İki belirli kişinin birlikte komitede yer alma olasılığını hesaplayınız.
- A5. $\sum_{k=0}^n k \binom{n}{k} = n2^{n-1}$ özdeşliğini sayma argümanı ile açıklayınız.

2.18 Kaynak Notu

Bu bölümdeki sayma ilkeleri, permütasyon, kombinasyon, binom teoremi, multinom katsayıları ve olasılığa uygulamaları; Feller (1968), Ross (2019), Grinstead ve Snell (1997), Hogg, McKean ve Craig (2019) ve Montgomery ve Runger (2018) gibi klasik olasılık ve istatistik kaynaklarında kullanılan standart kombinatorik çerçeveye uyumludur.

Özet:

Kombinatorik analiz, sonlu örnek uzayları doğru saymanın yöntemidir. Toplama kuralı alternatifleri, çarpma kuralı ardışık seçimleri sayar. Permütasyon sıralı seçimlerde, kombinasyon sırasız seçimlerde kullanılır. Tekrarlı seçimler ayrı formüller gerektirir. Binom ve multinom katsayıları çok sayıda olasılık modelinin temelidir. Binom ve hipergeometrik dağılımlar, kombinatorik saymanın doğrudan olasılık dağılımlarına dönüşmüş biçimleridir.

Bölüm 3

Rassal Değişkenler ve Dağılımlar

3.1 Gerçek Dünya Problemi (Giriş)

Bir üretim hattında bir günde çıkan kusurlu parça sayısı, bir web sitesinin dakika başına ziyaretçi sayısı, bir sigorta müşterisinin yıllık hasar tutarı veya bir öğrencinin sınav puanı farklı türden belirsizliklerdir. Bu belirsizlikleri doğrudan örnek uzay üzerinde tartışmak çoğu zaman zordur. Rassal değişken, deney sonucunu sayısal bir büyüklüğe dönüştürerek olasılık modelini hesaplanabilir hâle getirir.

Bu Bölüm Ne Sağlar?

Bu bölüm, olay temelli olasılık dilinden rassal değişken ve dağılım diline geçişi sağlar. Önce rassal değişken ve dağılım fonksiyonu tanımlanır; ardından kesikli ve sürekli değişkenler, beklenen değer, varyans, momentler, gösterge değişkenleri, olasılık üreten fonksiyonlar ve temel kesikli dağılımlar sistematik biçimde ele alınır.

Amaç ve Öğrenme Çıktıları

Bu bölümün sonunda okuyucu

- rassal değişkeni ölçülebilir bir fonksiyon olarak yorumlayabilir,
- dağılım fonksiyonunun temel özelliklerini ve aralık olasılıklarını kullanabilir,
- kesikli ve sürekli dağılımları ayırt edebilir,
- beklenen değer, varyans ve momentleri hesaplayabilir,
- gösterge değişkenleri ve olasılık üreten fonksiyonlarla sayma problemlerini çözebilir,
- Bernoulli, binom, geometrik, hipergeometrik ve Poisson gibi temel kesikli modelleri uygun bağlama yerleştirebilir.

3.2 Rassal Değişken Kavramı

Rassal değişken, rassal deneyin sonucunu sayısal bir değere dönüştüren fonksiyondur. Böylece olayların dili sayısal değişkenlerin dağılım diline çevrilir. Bu yaklaşım hem klasik olasılık metinlerinde hem de modern ölçü temelli anlatımda merkezi konumdadır (Capiński & Kopp, 2004; Hogg et al., 2019; Kolmogorov, 1956; Ross, 2019).

Tanım 3.1 (Rassal Değişken) $(\Omega, \mathcal{F}, P)$ bir olasılık uzayı olsun. Gerçek değerli bir fonksiyon $X : \Omega \rightarrow \mathbb{R}$ için her $x \in \mathbb{R}$ altında

$$\{\omega : X(\omega) \leq x\} \in \mathcal{F}$$

ise X bir rassal değişkendir.

Tanım 3.2 (Rassal Değişkenin Ürettiği Olay) X rassal değişkeni için $\{X \leq x\}$, $\{a < X \leq b\}$ ve $\{X \in B\}$ gibi kümeler olaydır. Olasılık hesabı bu olaylar üzerinden yapılır.

Örnek 3.1 (İki Zarın Toplamı). *İki zar atılsın ve X gelen sayıların toplamı olsun. Örnek uzay 36 sıralı çiftten oluşur; fakat X yalnız 2, 3, ..., 12 değerlerini alır. Böylece deney sonucu bir sayısal değişkene dönüştürülür.*

Örnek 3.2 (Mühendislik Ölçümü). *Bir makinede üretilen milin çapı ölçülüyorsa X çap değerini temsil eden rassal değişkendir. $\{X > 10.05\}$ olayı, milin üst şartname sınırını aşmasıdır.*

3.3 Dağılım Fonksiyonu

Bir rassal değişkenin dağılımı, hangi değerleri hangi olasılıklarla aldığını belirtir. Bütün rassal değişkenler için ortak ve en genel temsil dağılım fonksiyonudur (Hsu, 1997; Kolmogorov, 1956).

Tanım 3.3 (Dağılım Fonksiyonu) Bir rassal değişken X için

$$F_X(x) = P(X \leq x), \quad x \in \mathbb{R}$$

fonksiyonuna X 'in kümülatif dağılım fonksiyonu denir.

Teorem 3.1 (Dağılım Fonksiyonunun Temel Özellikleri) *Her dağılım fonksiyonu F için şu özellikler geçerlidir:*

1. F azalmayandır.
2. $\lim_{x \rightarrow -\infty} F(x) = 0$ ve $\lim_{x \rightarrow \infty} F(x) = 1$.
3. F sağdan süreklidir.

Kanıt. $x < y$ ise $\{X \leq x\} \subseteq \{X \leq y\}$ olduğundan monotonluk $F(x) \leq F(y)$ verir. Uç limitler, $\{X \leq x\}$ olaylarının $x \rightarrow -\infty$ altında boş olaya, $x \rightarrow \infty$ altında tüm örnek uzaya yaklaşmasından gelir. Sağdan süreklilik ise olasılık ölçüsünün azalan olay dizileri altında süreklilik özelliğinin uygulanmasıdır. $\square$

Teorem 3.2 (Aralık Olasılığı) Her $a < b$ için

$$P(a < X \leq b) = F_X(b) - F_X(a)$$

olur.

Kanıt. $\{X \leq b\}$ olayı ayrık biçimde $\{X \leq a\}$ ve $\{a < X \leq b\}$ olaylarının birleşimidir. Dolayısıyla

$$F_X(b) = F_X(a) + P(a < X \leq b).$$

Düzenleme sonucu verir. $\square$

Tanım 3.4 (Sıçrama Büyüklüğü) Dağılım fonksiyonunun x noktasındaki sıçrama büyüklüğü

$$F_X(x) - F_X(x^-)$$

ile tanımlanır ve bu değer $P(X = x)$ olasılığına eşittir.

Örnek 3.3 (Sağdan Kapalı Aralık). $F(1) = 0.40$ ve $F(3) = 0.85$ ise

$$P(1 < X \leq 3) = 0.85 - 0.40 = 0.45$$

olur.

Bekleme süreleri ve güvenilirlik analizinde, dağılım fonksiyonu yerine onun tümleyeni olan *sağ kalım (survival) fonksiyonu* daha doğal bir araçtır.

Tanım 3.5 (Sağ Kalım Fonksiyonu) Bir rassal değişken X için

$$S_X(x) = P(X > x) = 1 - F_X(x)$$

fonksiyonuna sağ kalım fonksiyonu denir. Azalmayan F' 'ye karşılık S azalandır; $S(-\infty) = 1$, $S(\infty) = 0$ olur.

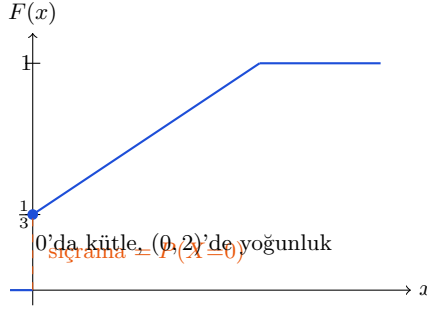
Teorem 3.3 (Negatif Olmayan Değişkende Kuyruk Toplamı Formülü) $X \geq 0$ tam sayı değerli ise $E[X] = \sum_{k=1}^{\infty} P(X \geq k)$; $X \geq 0$ sürekli ise $E[X] = \int_0^{\infty} P(X > x) dx$ olur.

Kanıt. Tam sayı durumunda $X = \sum_{k=1}^{\infty} \mathbf{1}\{X \geq k\}$ yazılır; beklenti alınır $E[\mathbf{1}\{X \geq k\}] = P(X \geq k)$ kullanılır. Sürekli durumda $X = \int_0^{\infty} \mathbf{1}\{X > x\} dx$ özdeşliğinin beklentisi alınır ve Tonelli teoremiyle integral ile beklenti yer değiştirir. $\square$

Örnek 3.4 (Karma Tip Dağılım Fonksiyonu). *Bir bekleme süresi X , $1/3$ olasılıkla anında ($X = 0$) hizmet alsın; aksi hâlde $\text{Uniform}(0, 2)$ dağılsın. CDF*

$$F(x) = \begin{cases} 0, & x < 0, \\ \frac{1}{3} + \frac{2}{3} \cdot \frac{x}{2}, & 0 \leq x < 2, \\ 1, & x \geq 2 \end{cases}$$

olup $x = 0$ 'da $1/3$ büyüklüğünde sıçrama, $(0, 2)$ üzerinde ise düz bir yoğunluk parçası içerir.



Şekil 3.1 Karma tip dağılım fonksiyonu: bir noktada sıçrama (kütle) ve bir aralıkta sürekli artış (yoğunluk) bir arada bulunur.

3.4 Kesikli Rassal Değişkenler

Kesikli rassal değişkenler sonlu veya sayılabilir sayıda değer alır. Bu durumda dağılım, olasılık kütle fonksiyonu ile verilir (Grinstead & Snell, 1997; Hogg et al., 2019; Ross, 2019).

Tanım 3.6 (Olasılık Kütle Fonksiyonu) X sayılabilir bir S kümesinde değer alıyorsa

$$p_X(x) = P(X = x), \quad x \in S$$

fonksiyonuna olasılık kütle fonksiyonu denir. Bu fonksiyon

$$p_X(x) \geq 0, \quad \sum_{x \in S} p_X(x) = 1$$

koşullarını sağlar.

Teorem 3.4 (Kesikli Değişkende Olay Olasılığı) X kesikli ve $A \subseteq S$ ise

$$P(X \in A) = \sum_{x \in A} p_X(x)$$

olur.

Kanıt. $\{X \in A\}$ olayı ayrık biçimde $\bigcup_{x \in A} \{X = x\}$ olarak yazılır. Sayılabilir toplamsallık doğrudan sonucu verir. $\square$

Örnek 3.5 (İki Zarın Toplamı). *İki zarın toplamı X olsun. $P(X = 7) = 6/36$ olur; çünkü 7 toplamını veren altı sıralı sonuç vardır: $(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)$.*

3.5 Sürekli Rassal Değişkenler

Sürekli rassal değişkenlerde tek tek nokta olasılıkları çoğu zaman sıfırdır; olasılık aralıklar üzerinde yoğunluk integraliyle hesaplanır (Hogg et al., 2019; Ross, 2019).

Tanım 3.7 (Olasılık Yoğunluk Fonksiyonu) X rassal değişkeni için negatif olmayan bir f_X fonksiyonu

$$F_X(x) = \int_{-\infty}^x f_X(t) dt$$

koşulunu sağlıyorsa f_X fonksiyonuna X 'in olasılık yoğunluk fonksiyonu denir.

Teorem 3.5 (Sürekli Değişkende Aralık Olasılığı) *Yoğunluğu f_X olan sürekli bir X için*

$$P(a < X \leq b) = \int_a^b f_X(x) dx$$

olur.

Kanıt. Dağılım fonksiyonu tanımından

$$P(a < X \leq b) = F_X(b) - F_X(a)$$

yazılır. Yoğunluk temsili yerine konursa

$$F_X(b) - F_X(a) = \int_{-\infty}^b f_X(x) dx - \int_{-\infty}^a f_X(x) dx = \int_a^b f_X(x) dx$$

elde edilir. $\square$

Teorem 3.6 (Sürekli Değişkende Tek Nokta Olasılığı) *Yoğunluğu olan bir rassal değişken için her a değerinde*

$$P(X = a) = 0$$

olur.

Kanıt. Tek nokta olasılığı

$$P(X = a) = P(a \leq X \leq a) = \int_a^a f_X(x) dx = 0$$

olarak hesaplanır. □

Örnek 3.6 (Üstel Bekleme Süresi). $f(x) = 2e^{-2x}$, $x \geq 0$ yoğunluğuna sahip bir bekleme süresi için

$$P(1 < X \leq 2) = \int_1^2 2e^{-2x} dx = e^{-2} - e^{-4}$$

olur.

3.6 Karma Tip Dağılımlar ve Model Seçimi

Gerçek veri her zaman tamamen kesikli veya tamamen sürekli olmayabilir. Bir değişken belirli bir noktada pozitif olasılık taşıyıp kalan bölgede sürekli yoğunluğa sahip olabilir. Mühendislik ve güvenilirlik problemlerinde bu tür karma modeller sık görülür (Soong, 2004).

Tanım 3.8 (Karma Tip Dağılım) Bir rassal değişkenin dağılımı hem bazı noktalarda sıçrama olasılığı hem de bazı aralıklarda yoğunluk içeriyorsa karma tip dağılım söz konusudur.

Örnek 3.7 (Sıfırda Bekleme ve Sürekli Gecikme). *Bir sistemde müşterilerin bir kısmı hiç beklemeden hizmete alınabilir; bekleyen müşteriler için bekleme süresi sürekli olabilir. Bu durumda $X = 0$ noktasında pozitif olasılık, $x > 0$ için yoğunluk vardır.*

Veri yapısı	Uygun dağılım dili	Örnek
İki sonuç	Bernoulli	Kusurlu/kusursuz
Başarı sayısı	Binom	20 üründe kusurlu sayısı
İlk başarıya kadar bekleme	Geometrik	İlk satışa kadar arama
Yerine koymadan seçim	Hipergeometrik	Partiden kusurlu seçimi
Nadir olay sayısı	Poisson	Dakikadaki çağrı sayısı
Sürekli ölçüm	Yoğunluk fonksiyonu	Mil çapı veya ömür

3.7 Beklenen Değer, Varyans ve Momentler

Dağılımın tamamı ayrıntılı bilgi verir; fakat uygulamada çoğu zaman merkezi eğilim ve yayılım özetleri gerekir. Beklenen değer dağılımın ağırlık merkezini, varyans ise ortalama etrafındaki yayılımı ölçer (Hogg et al., 2019; Ross, 2019).

Tanım 3.9 (Beklenen Değer) Kesikli durumda

$$E[X] = \sum_x xp_X(x)$$

ve sürekli durumda

$$E[X] = \int_{-\infty}^{\infty} xf_X(x) dx$$

olarak tanımlanır; toplam veya integral mutlak yakınsak olduğunda beklenti vardır.

Tanım 3.10 (Varyans ve Standart Sapma) $E[X] = \mu$ ise

$$\text{Var}(X) = E[(X - \mu)^2]$$

olarak tanımlanır. Standart sapma $\sqrt{\text{Var}(X)}$ değeridir.

Teorem 3.7 (Varyansın Hesaplama Formülü) İkinci moment sonlu ise

$$\text{Var}(X) = E[X^2] - \{E[X]\}^2$$

olur.

Kanıt. $\mu = E[X]$ olsun. Kare açılımı ile

$$E[(X - \mu)^2] = E[X^2 - 2\mu X + \mu^2] = E[X^2] - 2\mu E[X] + \mu^2 = E[X^2] - \mu^2$$

sonucu elde edilir. $\square$

Tanım 3.11 (Moment) k pozitif tam sayı olmak üzere $E[X^k]$ değerine k 'inci ham moment denir. $E[(X - \mu)^k]$ ise k 'inci merkezi momenttir.

Örnek 3.8 (Bernoulli Değişkeninin Ortalama ve Varyansı). $X \sim \text{Bernoulli}(p)$ ise $E[X] = p$ ve $E[X^2] = p$ olur. Bu nedenle

$$\text{Var}(X) = p - p^2 = p(1 - p)$$

bulunur.

Beklentinin iki temel özelliği, hesapların büyük kısmını taşır: bir fonksiyonun beklentisini dağılımı yeniden hesaplamadan bulmayı sağlayan “bilinçsiz istatistikçi yasası” ve toplamların beklentisini ayırıştıran doğrusallık.

Teorem 3.8 (Bilinçsiz İstatistikçi Yasası (LOTUS)) g bir fonksiyon olsun. Kesikli durumda

$$E[g(X)] = \sum_x g(x) p_X(x),$$

sürekli durumda

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$$

olur (ilgili toplam/integral mutlak yakınsadığında).

Kanıt. Kesikli durumda $g(X)$ 'in aldığı her y değeri için $P(g(X) = y) = \sum_{x:g(x)=y} p_X(x)$ yazılır. Beklenti tanımında $\sum_y y P(g(X) = y)$ ifadesi, içteki toplamla birleştirilince $\sum_x g(x) p_X(x)$ biçimine gelir. Sürekli durum benzer biçimde ölçü değişimiyle elde edilir. $\square$

Teorem 3.9 (Beklentinin Doğrusallığı ve Varyansın Ölçeklenmesi) *Beklentileri sonlu X, Y ve sabitler a, b için*

$$E[aX + bY] = aE[X] + bE[Y], \quad \text{Var}(aX + b) = a^2 \text{Var}(X).$$

Doğrusallık bağımsızlık gerektirmez.

Kanıt. Doğrusallık, toplam/integralin doğrusallığından çıkar ve ortak dağılım üzerinden $aX + bY$ değerlerinin ağırlıklı toplamı olarak yazılır. Varyans için $\text{Var}(aX + b) = E[(aX + b - a\mu - b)^2] = E[a^2(X - \mu)^2] = a^2 \text{Var}(X)$ olur; sabit b yayılımı değiştirmez. $\square$

Örnek 3.9 (Standartlaştırma). $E[X] = \mu$, $\text{Var}(X) = \sigma^2 > 0$ olsun. $Z = (X - \mu)/\sigma$ değişkeni için doğrusallık ve ölçekleme kuralından $E[Z] = 0$ ve $\text{Var}(Z) = \sigma^2/\sigma^2 = 1$ bulunur. Standartlaştırma, farklı ölçekteki değişkenleri ortak bir tabana taşır.

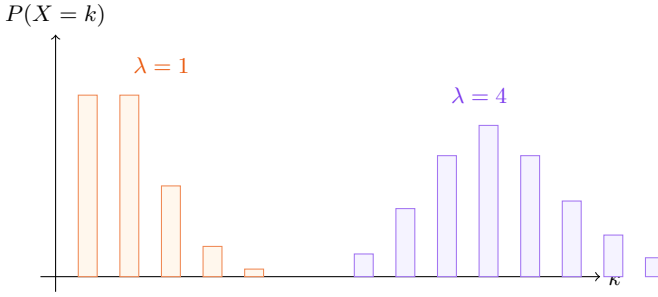
Ortalama ve varyans bir dağılımı tam betimlemez. Dağılımın asimetrisi ve kuyruk ağırlığı, üçüncü ve dördüncü standartlaştırılmış momentlerle ölçülür.

Tanım 3.12 (Çarpıklık ve Basıklık) $\mu = E[X]$ ve $\sigma^2 = \text{Var}(X)$ olsun. Çarpıklık ve basıklık

$$\gamma_1 = E\left[\left(\frac{X - \mu}{\sigma}\right)^3\right], \quad \gamma_2 = E\left[\left(\frac{X - \mu}{\sigma}\right)^4\right]$$

ile tanımlanır. $\gamma_1 > 0$ sağa çarpık, $\gamma_1 < 0$ sola çarpık dağılımı; γ_2 ise normal dağılımın değeri olan 3 ile karşılaştırılarak kuyruk ağırlığını gösterir.

Örnek 3.10 (Poisson Çarpıklığı). $X \sim \text{Poisson}(\lambda)$ için $\gamma_1 = 1/\sqrt{\lambda}$ olur. λ büyüdükçe çarpıklık azalır ve dağılım giderek simetrik (normale yakın) hâle gelir; bu, Poisson'un büyük λ 'da normal yaklaşımının nicel göstergesidir.



Şekil 3.2 Poisson PMF'leri: küçük λ 'da dağılım sağa çarpıkken, büyük λ 'da simetrikleşir ve normale yaklaşır.

Teorem 3.10 (Markov Eşitsizliği (Ön İzleme)) $X \geq 0$ ve $a > 0$ ise

$$P(X \geq a) \leq \frac{E[X]}{a}.$$

Kanıt. $a \mathbf{1}\{X \geq a\} \leq X$ olduğundan beklenti alınınca $aP(X \geq a) \leq E[X]$ olur; a 'ya bölünür. Bu eşitsizlik, yalnız ortalamayı kullanarak kuyruk olasılığına sınır verir ve Chebyshev eşitsizliğinin temelidir. $\square$

3.8 Gösterge Değişkenleri ve Sayma

Gösterge değişkenleri sayma problemlerini beklenti hesabına dönüştürür. Bu teknik olasılık ve istatistikte çok kullanılır (Bertsekas & Tsitsiklis, 2008; Ross, 2019).

Tanım 3.13 (Gösterge Rassal Değişkeni) Bir A olayı için

$$I_A(\omega) = \begin{cases} 1, & \omega \in A, \\ 0, & \omega \notin A \end{cases}$$

biçiminde tanımlanan değişkene A olayının gösterge değişkeni denir.

Teorem 3.11 (Gösterge Beklenen Değeri) Her A olayı için

$$E[I_A] = P(A)$$

olur.

Kanıt. I_A yalnız 0 ve 1 değerlerini alır. Bu nedenle

$$E[I_A] = 1 \cdot P(I_A = 1) + 0 \cdot P(I_A = 0) = P(A)$$

olur. □

Örnek 3.11 (Kusurlu Sayısı). n üründe I_i değişkeni i 'inci ürün kusurluysa 1, değilse 0 olsun. Toplam kusurlu sayısı

$$X = I_1 + \cdots + I_n$$

şeklinde yazılır. Beklentinin doğrusallığı ile $E[X] = \sum_i P(I_i = 1)$ bulunur.

Örnek 3.12 (Kupon Toplayıcı Problemi). N farklı kupon türü olsun; her satın almada eş olasılıkla bir tür gelir. Tüm türleri toplamak için gereken alış sayısı T 'nin beklenen değerini gösterge-bekleme tekniğiyle bulalım. $T = T_1 + T_2 + \cdots + T_N$ yazılır; burada T_i , $i - 1$ tür toplanmışken yeni bir türü yakalamak için gereken ek alış sayısıdır. Yeni tür gelme olasılığı $(N - i + 1)/N$ olduğundan T_i geometriktir ve $E[T_i] = N/(N - i + 1)$ olur. Doğrusallıkla

$$E[T] = \sum_{i=1}^N \frac{N}{N - i + 1} = N \sum_{j=1}^N \frac{1}{j} = N H_N \approx N(\ln N + \gamma),$$

burada H_N harmonik sayı, $\gamma \approx 0.5772$ Euler sabitidir. Örneğin $N = 50$ için $E[T] \approx 225$ alış gerekir (Ross, 2019).

3.9 Olasılık Üreten Fonksiyon

Negatif olmayan tam sayı değerli rassal değişkenlerde olasılık üreten fonksiyon, dağılımı tek bir fonksiyonda toplar. Özellikle toplamalar ve momentler için yararlıdır (Mukhopadhyay, 2000; Ross, 2019).

Tanım 3.14 (Olasılık Üreten Fonksiyon) $X = 0, 1, 2, \dots$ değerlerini alan bir rassal değişken için

$$G_X(s) = E[s^X] = \sum_{k=0}^{\infty} P(X = k)s^k$$

fonksiyonuna olasılık üreten fonksiyon denir.

Teorem 3.12 (PGF'den Olasılıkların Elde Edilmesi) $G_X(s) = \sum_{k=0}^{\infty} p_k s^k$ ise

$$p_k = \frac{G_X^{(k)}(0)}{k!}$$

olur.

Kanıt. Kuvvet serisinin katsayıları tekildir. G_X fonksiyonunun k kez türevi alınıp $s = 0$ yazıldığında yalnız k 'inci katsayı kalır ve $G_X^{(k)}(0) = k!p_k$ elde edilir. □

Örnek 3.13 (Poisson PGF). $X \sim \text{Poisson}(\lambda)$ ise

$$G_X(s) = \exp\{\lambda(s-1)\}$$

olur. Bu fonksiyon, bağımsız Poisson toplamlarının yine Poisson dağılımlı olduğunu hızlıca göstermekte kullanılır.

Teorem 3.13 (PGF ile Moment ve Toplam) G_X türevlenebilir olsun. O hâlde $E[X] = G'_X(1)$ ve $E[X(X-1)] = G''_X(1)$ olur. Ayrıca X ile Y bağımsızsa

$$G_{X+Y}(s) = G_X(s)G_Y(s).$$

Kanıt. Seri terim terim türevlenip $s = 1$ konulduğunda $G'_X(1) = \sum_k k p_k = E[X]$ ve $G''_X(1) = \sum_k k(k-1)p_k = E[X(X-1)]$ çıkar. Bağımsızlık altında $E[s^{X+Y}] = E[s^X]E[s^Y]$ olduğundan çarpım kuralı elde edilir. $\square$

Örnek 3.14 (Bağımsız Poisson Toplamı). $X \sim \text{Poisson}(\lambda)$ ve $Y \sim \text{Poisson}(\mu)$ bağımsızsa

$$G_{X+Y}(s) = e^{\lambda(s-1)}e^{\mu(s-1)} = e^{(\lambda+\mu)(s-1)},$$

yani $X + Y \sim \text{Poisson}(\lambda + \mu)$ olur. PGF'nin çarpım özelliği toplamın dağılımını anında verir.

3.10 Moment Üreten Fonksiyon

Olasılık üreten fonksiyon yalnız tam sayı değerli değişkenler için tanımlıdır. Sürekli değişkenler dâhil daha genel bir araç, moment üreten fonksiyondur (MGF). Momentleri tek bir fonksiyondan üretir ve dağılımları tek biçimde belirler (Casella & Berger, 2002; Hogg et al., 2019).

Tanım 3.15 (Moment Üreten Fonksiyon) Bir rassal değişken X için 0 etrafında bir aralıkta sonlu olmak üzere

$$M_X(t) = E[e^{tX}]$$

fonksiyonuna moment üreten fonksiyon denir.

Teorem 3.14 (MGF'den Moment Üretimi) $M_X(t)$ sıfır komşuluğunda sonlu ise her k için $E[X^k]$ vardır ve

$$E[X^k] = M_X^{(k)}(0)$$

olur.

Kanıt. $e^{tX} = \sum_{k \geq 0} t^k X^k / k!$ açılımının beklentisi alınırsa $M_X(t) = \sum_{k \geq 0} E[X^k] t^k / k!$ elde edilir. Bu kuvvet serisinde t^k katsayısı $E[X^k] / k!$ olduğundan k 'inci türev sıfırda $E[X^k]$ verir. $\square$

Teorem 3.15 (MGF'nin Tekliği ve Toplam Özelliği) *İki değişkenin MGF'leri bir sıfır komşuluğunda çakışıyorsa dağılımları aynıdır. Ayrıca X ile Y bağımsızsa $M_{X+Y}(t) = M_X(t)M_Y(t)$ olur.*

Kanıt. Teklik, MGF'nin Laplace dönüşümüyle ilişkisinden ve dönüşümün teriminin tek olmasından gelir (ileri düzey ispat ölçü kuramı gerektirir). Toplam özelliği için bağımsızlık altında $E[e^{t(X+Y)}] = E[e^{tX}]E[e^{tY}]$ yazılır. $\square$

Örnek 3.15 (Üstel Dağılımın MGF'si). X üstel(λ) ise $t < \lambda$ için

$$M_X(t) = \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx = \frac{\lambda}{\lambda - t}.$$

Türevelerden $E[X] = M'(0) = 1/\lambda$ ve $E[X^2] = M''(0) = 2/\lambda^2$, dolayısıyla $\text{Var}(X) = 1/\lambda^2$ bulunur.

3.11 Temel Kesikli Dağılımlar

Rassal değişkenleri modellemede bazı dağılımlar tekrar tekrar karşımıza çıkar. Dağılım seçimi, verinin deney yapısına bağlıdır.

Tanım 3.16 (Bernoulli Dağılımı) X yalnız 0 ve 1 değerlerini alıyor ve $P(X = 1) = p$ ise $X \sim \text{Bernoulli}(p)$ denir.

Tanım 3.17 (Binom Dağılımı) n bağımsız Bernoulli(p) denemesindeki başarı sayısı

$$X \sim \text{Bin}(n, p), \quad P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

dağılımına sahiptir.

Teorem 3.16 (Binom Dağılımının Ortalaması ve Varyansı) $X \sim \text{Bin}(n, p)$ ise

$$E[X] = np, \quad \text{Var}(X) = np(1 - p)$$

olur.

Kanıt. $X = I_1 + \dots + I_n$ biçiminde yazılır; burada I_i bağımsız Bernoulli(p) değişkenleridir. Beklentinin doğrusallığı $E[X] = np$ verir. Bağımsızlık nedeniyle varyanslar toplanır ve $\text{Var}(X) = np(1 - p)$ olur. $\square$

Tanım 3.18 (Geometrik Dağılım) Bağımsız Bernoulli denemelerinde ilk başarının geldiği deneme sayısı X ise

$$P(X = k) = (1 - p)^{k-1} p, \quad k = 1, 2, \dots$$

biçimindeki dağılıma geometrik dağılım denir.

Teorem 3.17 (Geometrik Dağılımın Beklenen Değeri) $X \sim \text{Geom}(p)$ ise

$$E[X] = \frac{1}{p}$$

olur.

Kanıt. $|r| < 1$ için $\sum_{k=1}^{\infty} kr^{k-1} = 1/(1-r)^2$ eşitliği kullanılır. $r = 1-p$ alınırsa

$$E[X] = \sum_{k=1}^{\infty} k(1-p)^{k-1}p = p \frac{1}{p^2} = \frac{1}{p}$$

olur. □

Tanım 3.19 (Hipergeometrik Dağılım) N elemanlı bir kitlenin M tanesi başarı türünde olsun. Yerine koymadan n eleman seçildiğinde başarı sayısı X için

$$P(X = x) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}$$

dağılımı hipergeometrik dağılımdır.

Teorem 3.18 (Hipergeometrik Dağılımın Ortalaması) $X \sim \text{Hipergeom}(N, M, n)$ ise

$$E[X] = n \frac{M}{N}$$

olur.

Kanıt. Seçilen n konum için I_i göstergesi, i 'inci seçimin başarı türünden olup olmadığını gösterebilir. $X = I_1 + \dots + I_n$ yazılır. Simetri nedeniyle $E[I_i] = M/N$ olur. Beklentinin doğrusallığından $E[X] = nM/N$ elde edilir. □

Tanım 3.20 (Poisson Dağılımı) $\lambda > 0$ için

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \dots$$

ise $X \sim \text{Poisson}(\lambda)$ denir.

Teorem 3.19 (Geometrik Dağılımın Hafızasızlığı) $X \sim \text{Geom}(p)$ ise her $m, n \geq 0$ için

$$P(X > m+n \mid X > m) = P(X > n).$$

Kesikli dağılımlar içinde bu özelliği yalnız geometrik dağılım taşır.

Kanıt. $P(X > k) = (1-p)^k$ olur. O hâlde

$$P(X > m+n \mid X > m) = \frac{(1-p)^{m+n}}{(1-p)^m} = (1-p)^n = P(X > n).$$

Geçmiş başarısızlıklar gelecekteki bekleyişi etkilemez. □

Teorem 3.20 (Poisson Dağılımının Ortalaması ve Varyansı) $X \sim \text{Poisson}(\lambda)$ ise $E[X] = \text{Var}(X) = \lambda$ olur.

Kanıt. PGF $G_X(s) = e^{\lambda(s-1)}$ kullanılır. $G'_X(1) = \lambda$ olduğundan $E[X] = \lambda$; $G''_X(1) = \lambda^2$ olduğundan $E[X(X-1)] = \lambda^2$ ve $\text{Var}(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$ bulunur. $\square$

Teorem 3.21 (Poisson, Binomun Limiti Olarak) $n \rightarrow \infty$, $p \rightarrow 0$ ve $np \rightarrow \lambda$ iken

$$\binom{n}{k} p^k (1-p)^{n-k} \rightarrow e^{-\lambda} \frac{\lambda^k}{k!}.$$

Kanıt. $p = \lambda/n$ konulur. $\binom{n}{k} p^k = \frac{n(n-1)\cdots(n-k+1)}{k!} \frac{\lambda^k}{n^k} \rightarrow \frac{\lambda^k}{k!}$ ve $(1-p)^{n-k} \rightarrow e^{-\lambda}$ olur. Çarpım istenen limiti verir. Bu, Poisson'un “çok sayıda nadir olay” modeli olmasının nedenidir (Feller, 1968). $\square$

Örnek 3.16 (Nadir Olay Yaklaşımı). $n = 1000$ ürün ve kusur olasılığı $p = 0.002$ ise kusurlu sayısı yaklaşık Poisson(2)'dir. Hiç kusur olmama olasılığı $e^{-2} \approx 0.135$ olarak hızlıca hesaplanır.

Tanım 3.21 (Negatif Binom (Pascal) Dağılımı) Bağımsız Bernoulli(p) denemelerinde r 'inci başarımın geldiği deneme sayısı X ise

$$P(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}, \quad k = r, r+1, \dots$$

dağılımına negatif binom denir. $E[X] = r/p$ ve $\text{Var}(X) = r(1-p)/p^2$ olur.

Tanım 3.22 (Kesikli Düzgün Dağılım) X değerlerini $\{1, 2, \dots, N\}$ kümesinde eş olasılıkla alıyorsa

$$P(X = k) = \frac{1}{N}, \quad E[X] = \frac{N+1}{2}, \quad \text{Var}(X) = \frac{N^2-1}{12}$$

olur. Adil zar ($N = 6$) bunun en bilinen örneğidir.

Örnek 3.17 (Model Seçimi). Bir partiden yerine koymadan 10 ürün seçiliyorsa hipergeometrik model uygundur. Her ürün bağımsız biçimde kusurlu/kusursuz üretiliyorsa binom model; sabit zaman aralığında nadir olay sayısı sayılıyorsa Poisson model doğal seçimdir.

3.12 Kantiller ve Olasılık İntegral Dönüşümü

Dağılım fonksiyonu yalnız olasılık hesaplamak için değil, simülasyon ve kantil hesapları için de kullanılır (Hogg et al., 2019; Ramachandran & Tsokos, 2009).

Tanım 3.23 (Kantil Fonksiyonu) Bir dağılım fonksiyonu F için

$$F^{-1}(u) = \inf\{x : F(x) \geq u\}, \quad 0 < u < 1$$

fonksiyonuna kantil fonksiyonu denir.

Teorem 3.22 (Olasılık İntegral Dönüşümü) X sürekli dağılım fonksiyonu F 'ye sahipse

$$U = F(X)$$

değişkeni $\text{Uniform}(0, 1)$ dağılımına sahiptir.

Kanıt. $0 < u < 1$ için $x_u = F^{-1}(u)$ olsun. Süreklilik ve monotonluk altında

$$P(F(X) \leq u) = P(X \leq x_u) = F(x_u) = u$$

olur. Bu, U 'nun $(0, 1)$ üzerinde düzgün dağılımlı olduğunu gösterir. $\square$

Örnek 3.18 (Ters Dönüşümle Simülasyon). $U \sim \text{Uniform}(0, 1)$ ve $X = F^{-1}(U)$ alınırsa X dağılım fonksiyonu F olan bir değişken üretir. Üstel dağılımda $F^{-1}(u) = -\log(1 - u)/\lambda$ kullanılır.

3.13 Dağılımın Bilgiye Bağlı Yorumu

Aynı fiziksel büyüklük, eldeki bilgi değiştikçe farklı dağılımlarla temsil edilebilir. Bu bakış özellikle Bayesçi yorumda önemlidir (Jaynes, 2003). Örneğin torbadaki kırmızı top oranı hakkında az bilgi varsa dağılım daha yaygın, çok bilgi varsa daha yoğun seçilir.

Tanım 3.24 (Kanıt Altında Dağılım) Eldeki bilgi I açıkça belirtilmek istendiğinde bir değişkenin dağılımı $p(x | I)$ biçiminde yazılır. Bu gösterim, dağılımın yalnız fiziksel sürece değil, bilgi durumuna da bağlı olduğunu vurgular.

Teorem 3.23 (Beklentinin Bilgiye Koşullu Yorumu) Kesikli durumda bilgi I altında beklenen değer

$$E[X | I] = \sum_x xp(x | I)$$

biçimindedir.

Kanıt. Beklenti, dağılım ağırlıklarıyla alınan ağırlıklı ortalamadır. Bilgi I değiştiğinde ağırlıklar $p(x | I)$ olarak güncellenir; formül aynı kalır. $\square$

Örnek 3.19 (Eksik Bilgiyle Torba Modeli). Bir torbadaki kırmızı top oranı bilinmiyorsa çekiliş sonucu için dağılım, eldeki ön bilgiye bağlıdır. Daha fazla gözlem yapıldıkça dağılım daha dar ve bilgilendirici hâle gelir.

3.14 Örnek Veri Tablosu ve Yazılım Çözümleri

Kavram	Formül	Yorum
CDF	$F(x) = P(X \leq x)$	Genel dağılım temsili
PMF	$p(x) = P(X = x)$	Kesikli model
PDF	$P(a < X \leq b) = \int_a^b f(x)dx$	Sürekli model
Beklenti	$E[X]$	Dağılım merkezi
Varyans	$E[(X - \mu)^2]$	Yayılım
PGF	$G(s) = E[s^X]$	Tam sayı değerli dağılım özeti

Python Çözümü

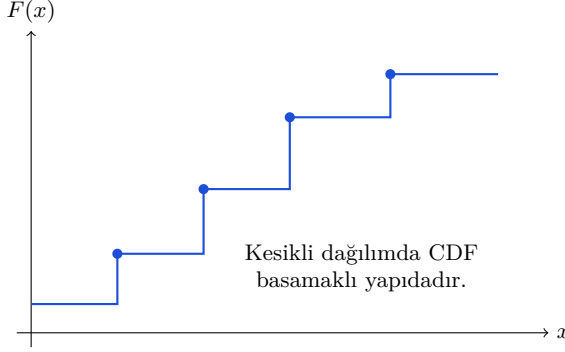
```
import math

p = 0.3
n = 10
prob = math.comb(n, 2) * p**2 * (1 - p)**8
print(prob)
```

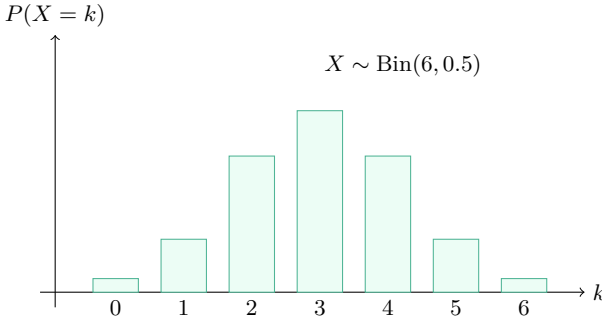
R ile Hesaplama:

```
p <- 0.3
n <- 10
prob <- choose(n, 2) * p * p * (1 - p) * (1 - p) * (1 - p) * (1 - p)
* (1 - p) * (1 - p) * (1 - p) * (1 - p)
prob
```

3.15 Görselleştirme ve Grafikler



Şekil 3.3 Kesikli rassal değişkende dağılım fonksiyonu sıçramalarla artar.



Şekil 3.4 Binom olasılık kütle fonksiyonu çubuk grafiği: $p = 0.5$ için dağılım ortalama $np = 3$ etrafında simetriktir.

Dikkat: Kavram Yanılgısı

Dağılım fonksiyonu ile yoğunluk fonksiyonu karıştırılmamalıdır. $F(x)$ bir olasılık birikimidir; sürekli durumda $f(x)$ tek başına olasılık değildir, olasılık ancak bir aralık üzerinde alan alınarak hesaplanır.

Kendini Dene:

Ara Kontrol: $P(X = 2) = 0$ bilgisi sürekli bir değişkende neden $X = 2$ değerinin imkânsız olduğu anlamına gelmez? İpucu: Nokta olasılığı ile aralık olasılığını ayırın.

3.16 Bölüm Sonu Çözümlü Problemler

1. $X \sim \text{Bernoulli}(0.3)$ ise $E[X]$ ve $\text{Var}(X)$ nedir?

Çözüm. $E[X] = 0.3$ ve $\text{Var}(X) = 0.3(0.7) = 0.21$.

2. $F(2) = 0.60$ ve $F(5) = 0.90$ ise $P(2 < X \leq 5)$ kaçtır?

Çözüm. $P(2 < X \leq 5) = F(5) - F(2) = 0.30$.

3. $X \sim \text{Bin}(8, 0.25)$ ise $P(X = 2)$ değerini yazınız.

Çözüm. $P(X = 2) = \binom{8}{2}(0.25)^2(0.75)^6$.

4. $X \sim \text{Geom}(0.2)$ ise $E[X]$ kaçtır?

Çözüm. $E[X] = 1/0.2 = 5$.

5. 100 parçalık partide 8 kusurlu varsa, yerine koymadan seçilen 10 parçada beklenen kusurlu sayısı nedir?

Çözüm. Hipergeometrik ortalama $10(8/100) = 0.8$ olur.

6. $X \sim \text{Poisson}(3)$ ise $P(X \geq 1)$ değerini bulunuz.

Çözüm. $P(X \geq 1) = 1 - P(X = 0) = 1 - e^{-3} \approx 0.950$.

7. $X \geq 0$ tam sayı değerli ve $P(X > k) = (0.8)^{k+1}$ ise $E[X]$ kaçtır?

Çözüm. Kuyruk toplamıyla $E[X] = \sum_{k \geq 1} P(X \geq k) = \sum_{k \geq 1} (0.8)^k = 0.8/0.2 = 4$.

8. X üstel(λ) ise MGF'sini kullanarak $\text{Var}(X)$ değerini bulunuz.

Çözüm. $M(t) = \lambda/(\lambda - t)$, $M'(0) = 1/\lambda$, $M''(0) = 2/\lambda^2$, dolayısıyla $\text{Var}(X) = 2/\lambda^2 - 1/\lambda^2 = 1/\lambda^2$.

9. $E[X] = 5$, $\text{Var}(X) = 4$ ise $Y = 3X - 2$ için $E[Y]$ ve $\text{Var}(Y)$ nedir?

Çözüm. $E[Y] = 3(5) - 2 = 13$, $\text{Var}(Y) = 9(4) = 36$.

10. Bağımsız $X \sim \text{Poisson}(2)$ ve $Y \sim \text{Poisson}(3)$ için $P(X + Y = 4)$ kaçtır?

Çözüm. $X + Y \sim \text{Poisson}(5)$, dolayısıyla $P = e^{-5}5^4/4! \approx 0.175$.

11. 6 farklı kupon türü varsa hepsini toplamak için beklenen alış sayısını bulunuz.

Çözüm. $E[T] = 6H_6 = 6(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6}) = 6 \cdot 2.45 = 14.7$.

12. $X \sim \text{Geom}(0.25)$ ise $P(X > 4 | X > 2)$ değerini hafızasızlıkla bulunuz.

Çözüm. Hafızasızlık gereği $P(X > 4 | X > 2) = P(X > 2) = (0.75)^2 = 0.5625$.

3.17 Bölüm Sonu Alıştırmaları

Bu alıştırmalar, rassal değişken kavramını, dağılım fonksiyonunu, beklenen değer-varyans hesaplarını ve temel kesikli dağılımları pekiştirmek için üç düzeyde düzenlenmiştir. Cevaplar ve kısa çözüm ipuçları Ekler bölümündeki *Cevap Anahtarı*'nda verilmiştir.

Temel Düzey

- B1. $X \sim \text{Bernoulli}(0.7)$ ise $P(X = 0)$, $E[X]$ ve $\text{Var}(X)$ değerlerini bulunuz.
- B2. X için $P(X = 0) = 0.2$, $P(X = 1) = 0.5$, $P(X = 2) = 0.3$ ise $E[X]$ değerini hesaplayınız.
- B3. $F(1) = 0.25$ ve $F(4) = 0.80$ ise $P(1 < X \leq 4)$ kaçtır?
- B4. $X \sim \text{Poisson}(2)$ ise $P(X = 0)$ ve $P(X = 1)$ değerlerini yazınız.
- B5. $E[X] = 4$ ve $\text{Var}(X) = 9$ ise $Y = 2X + 1$ için $E[Y]$ ve $\text{Var}(Y)$ değerlerini bulunuz.

Orta Düzey

- I1. $X \sim \text{Bin}(10, 0.4)$ ise $P(X = 3)$ değerini formül olarak yazınız ve yaklaşık hesaplayınız.
- I2. $X \sim \text{Geom}(0.25)$ ise $P(X > 3)$ ve $E[X]$ değerlerini bulunuz. Geometrik değişken ilk başarı deneme sayısını gösterebilir.
- I3. Bir partide 20 ürünün 5'i kusurludur. İadesiz seçilen 4 üründe tam 1 kusurlu gelme olasılığını yazınız.
- I4. X kesikli ve $P(X = 1) = 0.2$, $P(X = 3) = 0.3$, $P(X = 5) = 0.5$ ise $\text{Var}(X)$ değerini hesaplayınız.
- I5. Bağımsız $X \sim \text{Poisson}(1.5)$ ve $Y \sim \text{Poisson}(2.5)$ için $P(X + Y = 2)$ değerini bulunuz.

İleri Düzey

- A1. $X \geq 0$ tam sayı değerli ve $P(X \geq k) = (0.6)^k$ ($k = 1, 2, \dots$) ise $E[X]$ değerini kuyruk toplamıyla bulunuz.
- A2. $X \sim \text{Bin}(n, p)$ için gösterge değişkenleri kullanarak $E[X] = np$ olduğunu açıklayınız.
- A3. X için olasılık üreten fonksiyon $G(s) = 0.2 + 0.5s + 0.3s^2$ ise $E[X]$ ve $P(X = 2)$ değerlerini bulunuz.
- A4. $X \sim \text{Poisson}(\lambda)$ için $P(X = k + 1)/P(X = k)$ oranını bulunuz ve k arttıkça olasılıkların ne zaman azalmaya başladığını yorumlayınız.
- A5. $X \sim \text{Geom}(p)$ için $P(X > m + n \mid X > m) = P(X > n)$ hafızasızlık özelliğini doğrudan gösteriniz.

3.18 Kaynak Notu

Bu bölümde rassal değişken, ölçülebilirlik ve dağılım fonksiyonu yapısı Kolmogorov (1956), Capiński ve Kopp (2004) ve Shao (2003) çizgisinde verilmiştir. Kesikli ve sürekli dağılımlar, beklenen değer ve temel modeller Hogg et al. (2019), Hsu (1997), Ross (2019), Grinstead ve Snell (1997) ile uyumlu biçimde düzenlenmiştir. Mühendislik yorumu Soong (2004) ve Montgomery ve Runger (2018), bilgiye bağlı dağılım yorumu Jaynes (2003), olasılık üreten fonksiyon ve ileri istatistiksel dil Mukhopadhyay (2000) kaynaklarıyla ilişkilendirilmiştir. Moment üreten fonksiyon, LOTUS ve dağılım tekliği Casella ve Berger (2002); Poisson'un binom limiti, hafızasızlık ve kuyruk-toplam formülü Feller (1968) ve Ross (2019) çizgisinde sunulmuştur.

Özet:

Rassal değişken, deney sonucunu sayısal bir değere dönüştürür. Dağılım fonksiyonu bütün rassal değişkenler için genel temsildir; kesikli değişkenlerde PMF, sürekli değişkenlerde PDF kullanılır. Beklenen değer ve varyans dağılımın temel özetleridir. Gösterge değişkenleri, PGF ve temel kesikli dağılımlar sayma ve modelleme problemlerini sistematik hâle getirir.

Bölüm 4

Süreklı Dağılımlar, Dönüşümler ve Momentler

4.1 Gerçek Dünya Problemi (Giriş)

Bir parçanın ömrü, bir hastanın bekleme süresi, bir sensörün ölçüm hatası veya bir finansal getirinin büyüklüğü çoğu zaman aralık üzerinde değer alan sürekli değişkenlerle modellenir. Bu tür değişkenlerde tek bir noktanın olasılığı yerine aralık olasılıkları, yoğunluk fonksiyonları, dönüşümler ve momentler temel rol oynar.

Bu Bölüm Ne Sağlar?

Bu bölüm sürekli rassal değişkenleri daha ileri düzeyde ele alır. Önce yoğunluk ve beklenen değer yapısı pekiştirilir; sonra kantiller, dönüşüm kuralları, moment üreten fonksiyonlar, karakteristik fonksiyonlar ve temel sürekli dağılım aileleri incelenir. Amaç, sürekli modelleri yalnız formül listesi olarak değil, dönüşüm ve moment araçlarıyla çalışan bir modelleme dili olarak kurmaktır.

Amaç ve Öğrenme Çıktıları

Bu bölümün sonunda okuyucu

- sürekli yoğunluk ve dağılım fonksiyonu ilişkisini kullanabilir,
- kantil ve ters dönüşüm yöntemini açıklayabilir,
- tek değişkenli ve çok değişkenli dönüşümlerde Jacobian kuralını uygulayabilir,
- moment, merkezi moment, çarpıklık ve basıklığı yorumlayabilir,
- moment üreten ve karakteristik fonksiyonları toplam ve dağılım belirleme problemlerinde kullanabilir,
- üstel, gamma, normal, beta, Weibull ve Laplace gibi temel sürekli dağılımları uygun bağlama yerleştirebilir.

4.2 Sürekli Yoğunluk ve Dağılım Fonksiyonu

Sürekli değişkenlerde olasılık, yoğunluğun alanı olarak hesaplanır. Yoğunluk değeri tek başına olasılık değildir; bir aralık üzerindeki integral olasılıktır. Bu çerçevede modern olasılık ve matematiksel istatistik metinlerinde temel başlangıçtır (Hogg et al., 2019; Ross, 2019; Soong, 2004).

Tanım 4.1 (Sürekli Rassal Değişken) X rassal değişkeninin dağılım fonksiyonu bir f_X yoğunluğu ile

$$F_X(x) = \int_{-\infty}^x f_X(u) du$$

biçiminde yazılabiliyorsa X sürekli rassal değişkendir.

Tanım 4.2 (Yoğunluk Fonksiyonu) $f_X(x) \geq 0$ ve

$$\int_{-\infty}^{\infty} f_X(x) dx = 1$$

koşullarını sağlayan fonksiyon, bir sürekli dağılım için yoğunluk fonksiyonu olabilir.

Teorem 4.1 (Aralık Olasılığı) *Yoğunluğu f_X olan sürekli bir X için*

$$P(a < X \leq b) = \int_a^b f_X(x) dx$$

olur.

Kanıt. Dağılım fonksiyonu yardımıyla

$$P(a < X \leq b) = F_X(b) - F_X(a)$$

yazılır. Yoğunluk temsili yerine konursa

$$F_X(b) - F_X(a) = \int_{-\infty}^b f_X(x) dx - \int_{-\infty}^a f_X(x) dx = \int_a^b f_X(x) dx$$

elde edilir. □

Örnek 4.1 (Düzgün Yoğunluk). $X \sim U(a, b)$ ise $f_X(x) = 1/(b - a)$, $a < x < b$ olur. Bu durumda $[c, d] \subseteq [a, b]$ için

$$P(c < X < d) = \frac{d - c}{b - a}$$

bulunur.

Yoğunluk ve dağılım fonksiyonu yanında, negatif olmayan değişkenlerde beklenen değeri doğrudan sağkalım (kuyruk) olasılığı üzerinden veren bir gösterim de sıkça kullanılır. Bu gösterim hem teorik türetmelerde hem de güvenilirlik analizinde pratiktir (Ross, 2019; Billingsley, 1995).

Teorem 4.2 (Kuyruk İntegrali ile Beklenti) $X \geq 0$ sürekli bir rassal değişken ve $E[X] < \infty$ olsun. Bu durumda

$$E[X] = \int_0^{\infty} P(X > x) dx = \int_0^{\infty} (1 - F_X(x)) dx$$

olur.

Kanıt. Sağkalım olasılığını yoğunluk integraliyle yazıp integrasyon sırasını değiştirelim:

$$\int_0^{\infty} P(X > x) dx = \int_0^{\infty} \int_x^{\infty} f_X(u) du dx = \int_0^{\infty} \int_0^u f_X(u) dx du.$$

İç integral u verir; dolayısıyla ifade $\int_0^{\infty} u f_X(u) du = E[X]$ olur. Sıra değiştirme Tonelli teoremiyle, integrand negatif olmadığından geçerlidir. $\square$

Örnek 4.2 (Üstelde Kuyruk Formülü). $X \sim \text{Exp}(\lambda)$ için $P(X > x) = e^{-\lambda x}$ olduğundan

$$E[X] = \int_0^{\infty} e^{-\lambda x} dx = \frac{1}{\lambda}$$

bulunur. Klasik $\int_0^{\infty} x \lambda e^{-\lambda x} dx$ integraline gerek kalmadan aynı sonuca ulaşılır.

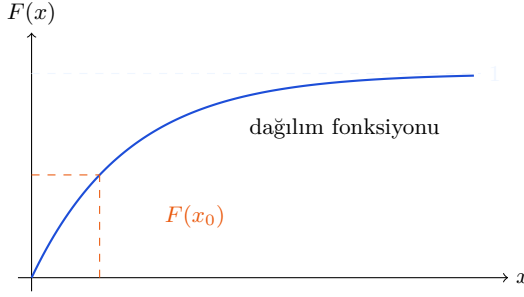
Yaşam süresi ve güvenilirlik modellemesinde, bir parçanın belirli bir yaşa kadar dayandığı bilindiğinde anlık arıza eğilimini ölçen tehlike (hazard) fonksiyonu merkezî bir kavramdır (Soong, 2004; Ramachandran & Tsokos, 2009).

Tanım 4.3 (Tehlike (Hazard) Fonksiyonu) Sağkalım fonksiyonu $S(x) = 1 - F_X(x) > 0$ olan sürekli bir X için tehlike fonksiyonu

$$h(x) = \frac{f_X(x)}{S(x)} = -\frac{d}{dx} \log S(x)$$

olarak tanımlanır. $h(x) dx$ değeri, x yaşına ulaşmış bir birimin $[x, x + dx)$ aralığında arızalanma koşullu olasılığını yaklaşık verir.

Örnek 4.3 (Sabit Tehlike ve Üstel Dağılım). Üstel dağılımda $h(x) = \lambda e^{-\lambda x} / e^{-\lambda x} = \lambda$ sabittir. Sabit tehlike “yaşlanmama” anlamına gelir ve üstel dağılımın hafızasızlık özelliğinin bir başka görünümüdür.



Şekil 4.1 Sürekli bir dağılım fonksiyonu sürekli, azalmayan ve 0'dan 1'e yükselen bir eğridir.

4.3 Kantiller ve Ters Dönüşüm

Kantil fonksiyonu, dağılım fonksiyonunun ters yönde kullanılmasını sağlar. Hem medyan ve yüzdelik hesaplarında hem de simülasyonda temel araçtır (Hogg et al., 2019; Ramachandran & Tsokos, 2009).

Tanım 4.4 (Kantil Fonksiyonu) Dağılım fonksiyonu F için

$$F^{-1}(u) = \inf\{x : F(x) \geq u\}, \quad 0 < u < 1$$

fonksiyonuna kantil fonksiyonu denir.

Teorem 4.3 (Ters Dönüşüm Yöntemi) $U \sim U(0, 1)$ ve $X = F^{-1}(U)$ olsun. F sürekli ve artan ise X 'in dağılım fonksiyonu F olur.

Kanıt. Her x için

$$P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x)$$

yazılır. Son eşitlik U 'nun $(0, 1)$ aralığında düzgün dağılımlı olmasından gelir. $\square$

Örnek 4.4 (Üstel Dağılım Üretimi). Üstel dağılımda $F(x) = 1 - e^{-\lambda x}$, $x \geq 0$ olur. $U \sim U(0, 1)$ için

$$X = -\frac{1}{\lambda} \log(1 - U)$$

değişkeni üstel dağılıma sahiptir.

Kantil fonksiyonu özel değerlerinde dağılımın konum ve yayılım özetlerini verir. Medyan ve çeyrekler, ortalama ve varyansa göre aykırı değerlere daha dayanıklıdır (Hogg et al., 2019; Wasserman, 2004).

Tanım 4.5 (Medyan ve Çeyrekler) $m = F^{-1}(1/2)$ değerine medyan denir. $Q_1 = F^{-1}(1/4)$ ve $Q_3 = F^{-1}(3/4)$ değerleri alt ve üst çeyrekler; $IQR = Q_3 - Q_1$ çeyrekler arası açıklıktır.

Örnek 4.5 (Üstel Dağılımın Medyanı). $X \sim \text{Exp}(\lambda)$ için $1 - e^{-\lambda m} = 1/2$ denklemi çözülür ve

$$m = \frac{\log 2}{\lambda}$$

bulunur. Bu değer ortalama $1/\lambda$ 'dan küçüktür; bu da üstel dağılımın sağa çarpık olduğunu gösterir.

Ters dönüşümün ikiz kardeşi, herhangi bir sürekli değişkeni kendi dağılım fonksiyonundan geçirerek düzgün dağılıma indirgeyen olasılık integral dönüşümüdür. Bu sonuç sözde rassal üretimden hipotez testlerine kadar geniş bir uygulama yelpazesinin temelidir (Casella & Berger, 2002).

Teorem 4.4 (Olasılık İntegral Dönüşümü) X sürekli dağılım fonksiyonu F olan bir rassal değişken olsun. Bu durumda $U = F(X)$ değişkeni $U(0, 1)$ dağılımına sahiptir.

Kanıt. F sürekli ve artan olsun. $0 < u < 1$ için

$$P(U \leq u) = P(F(X) \leq u) = P(X \leq F^{-1}(u)) = F(F^{-1}(u)) = u$$

yazılır. $[0, 1]$ üzerinde dağılım fonksiyonu u olan dağılım $U(0, 1)$ olduğundan iddia kanıtlanır. $\square$

Örnek 4.6 (Düzgünlük Sınaması). Bir model F doğruysa, gözlemlere F uygulan-
duğunda elde edilen değerler yaklaşık $U(0, 1)$ davranmalıdır. Bu fikir, kalıntıların
düzgünlüğünü inceleyen uyum iyiliği grafiklerinin (PP grafiği) temelini oluşturur.

4.4 Dönüşümle Yeni Dağılım Elde Etme

Bir rassal değişken üzerinde yapılan dönüşüm yeni bir dağılım üretir. Ölçek değiştirme, logaritma alma, kare alma veya koordinat değiştirme bu kapsamda düşünülür. Dönüşüm kuralları yoğunluğun doğru taşınmasını sağlar (Jaynes, 2003; Tarantola, 2001).

Tanım 4.6 (Türetilmiş Dağılım) $Y = g(X)$ biçiminde tanımlanan yeni rassal değişkenin dağılımına X 'ten türetilmiş dağılım denir.

Teorem 4.5 (Tek Değişkenli Yoğunluk Dönüşümü) X yoğunluğu f_X olan sürekli bir değişken, g türevlenebilir ve bire bir, $h = g^{-1}$ olsun. Bu durumda

$$f_Y(y) = f_X(h(y)) |h'(y)|$$

olur.

Kanıt. Artan g için $F_Y(y) = P(Y \leq y) = P(X \leq h(y)) = F_X(h(y))$ yazılır. Türev alınırsa $f_Y(y) = f_X(h(y))h'(y)$ elde edilir. Azalan durumda türev işaret değiştirdiğinden mutlak değer gerekir. $\square$

Örnek 4.7 (Ölçek Dönüşümü). $Y = cX$, $c > 0$ ise $h(y) = y/c$ ve $h'(y) = 1/c$ olur. Dolayısıyla

$$f_Y(y) = \frac{1}{c} f_X(y/c)$$

yazılır.

Teorem 4.6 (Çok Değişkenli Jacobian Kuralı) $Y = g(X)$ bire bir ve türevlenebilir bir dönüşüm, $h = g^{-1}$ tersi olsun. Bu durumda

$$f_Y(y) = f_X(h(y)) |\det J_h(y)|$$

olur.

Kanıt. Küçük hacim elemanları dönüşüm altında Jacobian determinantının mutlak değeriyle ölçeklenir. Olasılık kütlesi korunması gerektiğinden $f_Y(y)dy = f_X(x)dx$ yazılır ve $dx = |\det J_h(y)|dy$ kullanılır. $\square$

Örnek 4.8 (Kartezyenden Kutupsala). Düzlemde (X, Y) değişkeninden (R, Θ) değişkenine geçerken alan elemanı $dx dy = r dr d\theta$ olur. Bu nedenle kutupsal yoğunlukta Jacobian çarpanı r olarak görünür.

Dönüşüm bire bir değilse, ters görüntüdeki tüm kollar toplanarak yoğunluk elde edilir. Bu genel kural özellikle simetrik dönüşümlerde gereklidir (Hogg et al., 2019; Ross, 2019).

Teorem 4.7 (Monoton Olmayan Dönüşüm) g türevlenebilir olsun ve her y için $g(x) = y$ denkleminin çözümleri $x_1, \dots, x_k$ olsun. Bu durumda

$$f_Y(y) = \sum_{i=1}^k \frac{f_X(x_i)}{|g'(x_i)|}$$

olur.

Kanıt. g 'nin tek tek monoton olduğu aralıklara ayrılan tanım kümesinde her kol için tek değişkenli dönüşüm formülü uygulanır. $\{Y \leq y\}$ olayının ters görüntüsü bu aralıkların birleşimidir; ayrık katkılar toplandığında verilen ifade elde edilir. $\square$

Örnek 4.9 (Kare Dönüşümü). $Y = X^2$ için $y > 0$ olduğunda iki kol $x = \pm\sqrt{y}$ vardır ve $|g'(x)| = 2|x| = 2\sqrt{y}$ olur. Dolayısıyla

$$f_Y(y) = \frac{f_X(\sqrt{y}) + f_X(-\sqrt{y})}{2\sqrt{y}}.$$

$X \sim N(0, 1)$ alınırsa bu ifade χ_1^2 yoğunluğuna sadeleşir.

İki bağımsız sürekli değişkenin toplamının yoğunluğu, yoğunlukların evrişimi (konvolüsyonu) ile bulunur. Bu kural toplam dağılımlarının analizinde temel araçtır (Feller, 1971; Grimmett & Stirzaker, 2020).

Teorem 4.8 (Toplamanın Yoğunluğu: Evrişim) X ve Y bağımsız sürekli değişkenler, yoğunlukları f_X ve f_Y olsun. $Z = X + Y$ değişkeninin yoğunluğu

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x) dx$$

olur.

Kanıt. $F_Z(z) = P(X+Y \leq z) = \iint_{x+y \leq z} f_X(x) f_Y(y) dx dy$ yazılır. İç integralde $y \leq z - x$ sınırı kullanılıp z 'ye göre türev alındığında $f_Y(z - x)$ belirir ve sonuç elde edilir. $\square$

Örnek 4.10 (İki Düzgün Değişkenin Toplamı). $X, Y \sim U(0, 1)$ bağımsız ise evrişim integrali $Z = X + Y$ için üçgen yoğunluk verir: $f_Z(z) = z$ ($0 \leq z \leq 1$) ve $f_Z(z) = 2 - z$ ($1 \leq z \leq 2$). Bu, iki zarın sürekli karşılığıdır.

4.5 Momentler ve Dağılımın Şekli

Momentler dağılımın merkezini, yayılımını, asimetrisini ve kuyruk davranışını özetler. İlk iki moment çoğu uygulamada yeterli görünse de çarpıklık ve basıklık risk analizi ve kalite kontrol gibi alanlarda önemli bilgi taşır (Hogg et al., 2019; Montgomery & Runger, 2018).

Tanım 4.7 (Moment ve Merkezi Moment) k pozitif tam sayı olmak üzere $E[X^k]$ değerine k 'inci ham moment denir. $\mu = E[X]$ ise

$$E[(X - \mu)^k]$$

k 'inci merkezi momenttir.

Tanım 4.8 (Çarpıklık ve Basıklık) $\sigma^2 = \text{Var}(X)$ olmak üzere üçüncü standartlaştırılmış moment

$$\gamma_1 = E \left[\left(\frac{X - \mu}{\sigma} \right)^3 \right]$$

çarpıklık, dördüncü standartlaştırılmış moment

$$\gamma_2 = E \left[\left(\frac{X - \mu}{\sigma} \right)^4 \right]$$

basıklık ölçüsüdür.

Teorem 4.9 (Karesel Kayıp Altında Ortalama) $E[X^2] < \infty$ ise $E[(X - a)^2]$ ifadesini en küçük yapan a değeri $E[X]$ olur.

Kanıt. $\mu = E[X]$ olsun. Kare tamamlanır:

$$E[(X - a)^2] = E[(X - \mu + \mu - a)^2] = E[(X - \mu)^2] + (\mu - a)^2.$$

İlk terim a 'ya bağlı değildir; ikinci terim en küçük değerini $a = \mu$ iken alır. $\square$

Örnek 4.11 (Ölçüm Değerini Raporlama). *Bir ölçüm birçok kez tekrarlandığında tek sayı ile raporlama yapılacaksa karesel hata ölçütü altında en iyi rapor değeri ölçümlerin ortalamasıdır.*

Konveks fonksiyonlar ile beklenti arasındaki ilişki, momentlerin yorumlanmasında ve birçok eşitsizliğin kanıtında merkezî yer tutar. Jensen eşitsizliği bu ilişkinin temelidir (Casella & Berger, 2002; Williams, 1991).

Teorem 4.10 (Jensen Eşitsizliği) φ konveks bir fonksiyon ve $E[X]$ ile $E[\varphi(X)]$ sonlu olsun. Bu durumda

$$\varphi(E[X]) \leq E[\varphi(X)]$$

olur. φ konkavsa eşitsizlik yön değiştirir.

Kanıt. $\mu = E[X]$ olsun. Konveks φ için μ noktasında bir destek doğrusu vardır: her x için $\varphi(x) \geq \varphi(\mu) + c(x - \mu)$. Her iki tarafın beklentisi alınır; sağ taraftaki doğrusal terimin beklentisi sıfır olduğundan $E[\varphi(X)] \geq \varphi(\mu)$ elde edilir. $\square$

Örnek 4.12 (Ortalama Kareler ile Kare Ortalama). $\varphi(x) = x^2$ konveks olduğundan $E[X]^2 \leq E[X^2]$ olur. Aradaki fark tam olarak $\text{Var}(X)$ 'tir; varyansın negatif olamayacağı bu eşitsizliğin doğrudan bir sonucudur.

Tanım 4.9 (Değişim Katsayısı) $\mu = E[X] > 0$ ve $\sigma = \sqrt{\text{Var}(X)}$ için

$$\text{CV} = \frac{\sigma}{\mu}$$

oranına değişim katsayısı denir. Birimsiz olduğundan farklı ölçekteki dağılımların görelî yayılımını karşılaştırmada kullanılır.

Örnek 4.13 (Üstel Dağılımda Değişim Katsayısı). $X \sim \text{Exp}(\lambda)$ için $\mu = 1/\lambda$ ve $\sigma = 1/\lambda$ olduğundan $\text{CV} = 1$ olur. Değişim katsayısının 1 olması üstel modelin bir imzasıdır; veride $\text{CV} \approx 1$ gözlenmesi üstel varsayımı destekler.

4.6 Moment Üreten Fonksiyon

Moment üreten fonksiyon, dağılımın momentlerini tek bir fonksiyonda toplar. Var olduğu aralıktaki dağılımı belirleyebilir ve bağımsız toplamaları analiz etmeyi kolaylaştırır (Hogg et al., 2019; Mukhopadhyay, 2000).

Tanım 4.10 (Moment Üreten Fonksiyon) X rassal değişkeninin moment üreten fonksiyonu

$$M_X(t) = E[e^{tX}]$$

olarak tanımlanır. Bu fonksiyon $t = 0$ çevresinde sonluysa moment üretmek için kullanılabilir.

Teorem 4.11 (Moment Üretim Teoremi) $M_X(t)$ *sıfırın bir komşuluğunda sonluysa*

$$M_X^{(k)}(0) = E[X^k]$$

olur.

Kanıt. Üstel fonksiyonun seri açılımı

$$e^{tX} = \sum_{k=0}^{\infty} \frac{t^k X^k}{k!}$$

biçimindedir. Uygun sonluluk koşulları altında beklenti ve türev işlemleri yer değiştirir. k kez türev alınıp $t = 0$ yazıldığında $E[X^k]$ elde edilir. $\square$

Teorem 4.12 (Bağımsız Toplamlarda MGF Çarpımı) X ve Y *bağımsız ise*

$$M_{X+Y}(t) = M_X(t)M_Y(t)$$

olur.

Kanıt. Bağımsızlık nedeniyle

$$M_{X+Y}(t) = E[e^{t(X+Y)}] = E[e^{tX}e^{tY}] = E[e^{tX}]E[e^{tY}]$$

yazılır. $\square$

Örnek 4.14 (Gamma Toplamı). *Aynı hız parametresine sahip bağımsız gamma değişkenlerinin MGF'leri çarpıldığında şekil parametreleri toplanır. Bu nedenle ortak hız parametrelili gamma değişkenlerinin toplamı yine gamma dağılımlıdır.*

Örnek 4.15 (Üstel Dağılımın MGF'si). $X \sim \text{Exp}(\lambda)$ *için*

$$M_X(t) = \int_0^{\infty} e^{tx} \lambda e^{-\lambda x} dx = \frac{\lambda}{\lambda - t}, \quad t < \lambda$$

olur. Türevlerden $M_X'(0) = 1/\lambda$ ve $M_X''(0) = 2/\lambda^2$ bulunur; buradan $\text{Var}(X) = 2/\lambda^2 - 1/\lambda^2 = 1/\lambda^2$ elde edilir.

Moment üreten fonksiyonun en güçlü yanı yalnız moment üretmek değil, var olduğu yerde dağılımı tek biçimde belirlemesidir. Bu özellik toplam dağılımlarını tanımda doğrudan kullanılır (Casella & Berger, 2002; Mukhopadhyay, 2000).

Teorem 4.13 (MGF Tekillliği) *X ve Y değişkenlerinin moment üreten fonksiyonları sıfırın bir açık komşuluğunda var ve eşitse, X ile Y aynı dağılıma sahiptir.*

Kanıt. Sıfır komşuluğunda var olan moment üreten fonksiyon, iki yanlı Laplace dönüşümü üzerinden olasılık dağılımını tek biçimde belirler. Dönüşümün tersinin tekliği nedeniyle aynı MGF aynı dağılımı verir. $\square$

Örnek 4.16 (Toplamı Tanıma). *Bağımsız $X_i \sim \text{Exp}(\lambda)$ değişkenlerinin toplamının MGF'si $(\lambda/(\lambda - t))^n$ olur. Bu ifade $\text{Gamma}(n, \lambda)$ dağılımının MGF'sidir; tekillik gereği toplam gamma dağılımlıdır.*

4.7 Karakteristik Fonksiyon

Moment üreten fonksiyon her dağılım için var olmayabilir; karakteristik fonksiyon ise her rassal değişken için tanımlıdır. Bu nedenle limit teoremlerinde ve dağılım belirleme problemlerinde temel araçtır (Hsu, 1997; Kolmogorov, 1956; Shao, 2003).

Tanım 4.11 (Karakteristik Fonksiyon) *X rassal değişkeninin karakteristik fonksiyonu*

$$\varphi_X(t) = E[e^{itX}], \quad t \in \mathbb{R}$$

olarak tanımlanır.

Teorem 4.14 (Bağımsız Toplamların Karakteristik Fonksiyonu) *X ve Y bağımsız ise*

$$\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t)$$

olur.

Kanıt. MGF çarpımındaki aynı fikir kullanılır:

$$\varphi_{X+Y}(t) = E[e^{it(X+Y)}] = E[e^{itX}e^{itY}] = E[e^{itX}]E[e^{itY}].$$

$\square$

Teorem 4.15 (Karakteristik Fonksiyonların Tekillliği) *İki rassal değişkenin karakteristik fonksiyonları her $t \in \mathbb{R}$ için aynıysa dağılımları aynıdır.*

Kanıt. Karakteristik fonksiyon dağılımın Fourier dönüşümüdür. Fourier dönüşümünün tekillik özelliği, aynı karakteristik fonksiyona sahip iki olasılık ölçüsünün aynı dağılımı vermesini sağlar. $\square$

Örnek 4.17 (Normal Toplamı). *Bağımsız normal değişkenlerin karakteristik fonksiyonları çarpıldığında yine normal biçimli bir karakteristik fonksiyon elde edilir. Bu, bağımsız normal toplamalarının normal olduğunu gösterir.*

Karakteristik fonksiyonun her zaman var olması yanında, sınırlı ve sürekli olması da onu limit teoremlerinde elverişli kılar. Aşağıdaki temel özellikler bu kullanımları destekler (Billingsley, 1995; Shao, 2003).

Teorem 4.16 (Karakteristik Fonksiyonun Temel Özellikleri) *Her rassal değişken X için karakteristik fonksiyon φ_X şu özellikleri sağlar:*

$$\varphi_X(0) = 1, \quad |\varphi_X(t)| \leq 1, \quad \varphi_{aX+b}(t) = e^{itb} \varphi_X(at).$$

Kanıt. $\varphi_X(0) = E[e^0] = 1$ olur. İkinci özellik $|E[e^{itX}]| \leq E|e^{itX}| = E[1] = 1$ üçgen eşitsizliğinden gelir. Üçüncüsü için $\varphi_{aX+b}(t) = E[e^{it(aX+b)}] = e^{itb} E[e^{i(at)X}] = e^{itb} \varphi_X(at)$ yazılır. $\square$

Örnek 4.18 (Konum-Ölçek Dönüşümü). *X standart bir dağılıma sahipse, $Y = \mu + \sigma X$ değişkeninin karakteristik fonksiyonu $\varphi_Y(t) = e^{i\mu t} \varphi_X(\sigma t)$ olur. Bu kural, standart formdan konum-ölçek ailesinin tamamına geçmeyi sağlar.*

4.8 Temel Sürekli Dağılım Aileleri

Sürekli dağılım aileleri modelleme bağlamına göre seçilir: bekleme süreleri için üstel/gamma, simetrik hata için normal, oranlar için beta, ömür ve güvenilirlik için Weibull sık kullanılır.

Tanım 4.12 (Üstel Dağılım) $\lambda > 0$ için

$$f(x) = \lambda e^{-\lambda x}, \quad x \geq 0$$

yoğunluğuna sahip dağılıma üstel dağılım denir.

Tanım 4.13 (Gamma Dağılımı) $\alpha > 0, \lambda > 0$ için

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, \quad x > 0$$

yoğunluğuna sahip dağılıma gamma dağılımı denir.

Tanım 4.14 (Normal Dağılım) $\mu \in \mathbb{R}, \sigma > 0$ için

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\}$$

yoğunluğuna sahip dağılıma normal dağılım denir.

Teorem 4.17 (Normalden Ki-Kareye Geçiş) $Z \sim N(0, 1)$ ise $Y = Z^2$ değişkeni χ_1^2 dağılımına sahiptir.

Kanıt. $y > 0$ için

$$P(Y \leq y) = P(-\sqrt{y} \leq Z \leq \sqrt{y}) = \Phi(\sqrt{y}) - \Phi(-\sqrt{y})$$

yazılır. Türev alınırsa

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} y^{-1/2} e^{-y/2}$$

elde edilir. Bu yoğunluk χ_1^2 yoğunluğudur. $\square$

Tanım 4.15 (Beta Dağılımı) $\alpha, \beta > 0$ için

$$f(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}, \quad 0 < x < 1$$

yoğunluğuna sahip dağılıma beta dağılımı denir.

Teorem 4.18 (Beta Dağılımının Ortalaması) $X \sim \text{Beta}(\alpha, \beta)$ ise

$$E[X] = \frac{\alpha}{\alpha + \beta}$$

olur.

Kanıt. Beklenti integrali beta fonksiyonu ile

$$E[X] = \frac{B(\alpha + 1, \beta)}{B(\alpha, \beta)}$$

biçiminde yazılır. Gamma fonksiyonunun $\Gamma(u + 1) = u\Gamma(u)$ özelliği kullanıldığında sonuç $\alpha/(\alpha + \beta)$ olur. $\square$

Tanım 4.16 (Weibull Dağılımı) $\alpha, \lambda > 0$ için sağkalım fonksiyonu

$$P(X > x) = \exp\{-(\lambda x)^\alpha\}, \quad x \geq 0$$

olan dağılıma Weibull dağılımı denir.

Örnek 4.19 (Güvenilirlik Modeli). $\alpha > 1$ olan Weibull modeli, yaşlandıkça arıza riskinin arttığı sistemlerde kullanılır. $\alpha = 1$ özel durumu üstel dağılıma karşılık gelir.

Üstel dağılımı sürekli dağılımlar arasında ayrıcalıklı kılan özellik hafızasızlıktır: geçmiş bekleme, gelecekteki bekleme olasılığını değiştirmez (Ross, 2019).

Teorem 4.19 (Üstel Dağılımın Hafızasızlığı) $X \sim \text{Exp}(\lambda)$ ise her $s, t \geq 0$ için

$$P(X > s + t \mid X > s) = P(X > t)$$

olur. Üstel dağılım, bu özelliği taşıyan tek sürekli dağılımdır.

Kanıt. $P(X > x) = e^{-\lambda x}$ olduğundan

$$P(X > s + t \mid X > s) = \frac{P(X > s + t)}{P(X > s)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = P(X > t).$$

Tersine, sağkalım fonksiyonu S için $S(s+t) = S(s)S(t)$ koşulunu sağlayan ve azalan tek sürekli çözüm $S(x) = e^{-\lambda x}$ biçimindedir. $\square$

Tanım 4.17 (Laplace Dağılımı) $\mu \in \mathbb{R}$, $b > 0$ için

$$f(x) = \frac{1}{2b} \exp \left\{ -\frac{|x - \mu|}{b} \right\}, \quad x \in \mathbb{R}$$

yoğunluğuna sahip dağılıma Laplace (çift üstel) dağılımı denir. Simetrik fakat normalden daha sivri ve ağır kuyrukludur.

Tanım 4.18 (Lognormal Dağılım) $\log X \sim N(\mu, \sigma^2)$ ise X lognormal dağılımlıdır ve yoğunluğu

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp \left\{ -\frac{(\log x - \mu)^2}{2\sigma^2} \right\}, \quad x > 0$$

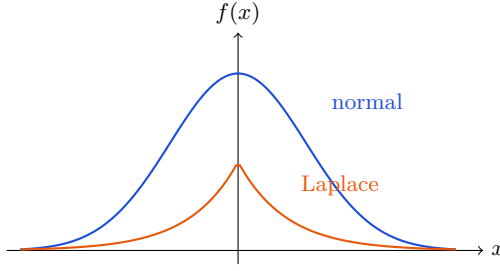
olur. Çarpımsal etkilerin biriktiği gelir, fiyat ve parçacık boyutu gibi pozitif büyüklüklerde kullanılır.

Tanım 4.19 (Cauchy Dağılımı) Konum x_0 ve ölçek $\gamma > 0$ için

$$f(x) = \frac{1}{\pi\gamma \left[1 + \left(\frac{x-x_0}{\gamma} \right)^2 \right]}, \quad x \in \mathbb{R}$$

yoğunluğuna sahip dağılıma Cauchy dağılımı denir. Beklenen değeri ve varyansı tanımsızdır; ağır kuyruk örneği olarak öne çıkar.

Örnek 4.20 (Tanımsız Ortalama). *Cauchy dağılımında $\int |x|f(x) dx$ ıraksar; bu nedenle örneklem ortalaması büyük örneklemde bile bir değere yerleşmez. Bu, büyük sayılar yasasının integrallenebilirlik koşulunun neden gerekli olduğunu gösteren klasik karşı örnektir.*



Şekil 4.2 Normal ve Laplace yoğunlukları: Laplace tepe noktasında daha sivri, kuyruklarda daha ağırdır.

4.9 Örnek Veri Tablosu ve Yazılım Çözümleri

Araç	Tanım	Kullanım
Yoğunluk	$f(x)$	Aralık olasılığı
Kantil	$F^{-1}(u)$	Yüzdelik ve simülasyon
Jacobian	$ h'(y) $ veya $ \det J $	Dönüşüm
MGF	$E[e^{tX}]$	Moment ve toplam
CF	$E[e^{itX}]$	Dağılım belirleme
Weibull	Sağkalım modeli	Güvenilirlik

Python ile Hesaplama:

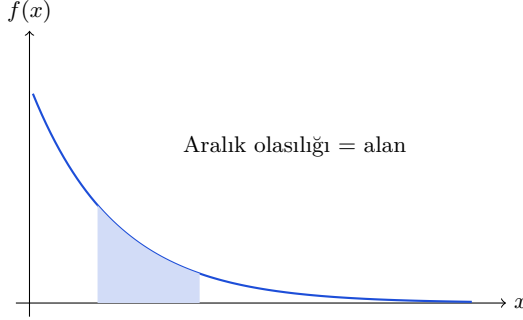
```
import math

lam = 2
prob = math.exp(-lam) - math.exp(-2 * lam)
print(prob)
```

R ile Hesaplama:

```
lambdaValue <- 2
prob <- exp(-lambdaValue * 1) - exp(-lambdaValue * 2)
prob
```

4.10 Görselleştirme ve Grafikler



Şekil 4.3 Sürekli dağılımda olasılık yoğunluğun altında kalan alanla hesaplanır.

Dikkat: Kavram Yanılgısı

Dönüşüm yaparken yalnız $y = g(x)$ yazmak yeterli değildir. Yoğunluk, değişken dönüşümünde Jacobian çarpımı ile ölçek değiştirir; monoton olmayan dönüşümlerde ise tüm ters görüntü kolları hesaba katılmalıdır.

Kendini Dene:

Ara Kontrol: $Y = X^2$ ve X sürekli ise $y > 0$ için neden iki ters kol vardır? Bu iki kol yoğunluk formülüne nasıl girer?

4.11 Bölüm Sonu Çözümlü Problemler

1. $X \sim U(0, 10)$ ise $P(2 < X < 5)$ nedir?
Çözüm. Düzgün dağılımda olasılık uzunluk oranıdır: $(5 - 2)/(10 - 0) = 0.3$.
2. $X \sim \text{Exp}(2)$ ise $P(1 < X \leq 2)$ nedir?
Çözüm. $P(1 < X \leq 2) = e^{-2} - e^{-4}$.
3. $Y = 3X$ ve X yoğunluğu f_X olan sürekli bir değişkense $f_Y(y)$ nedir?
Çözüm. $h(y) = y/3$ olduğundan $f_Y(y) = f_X(y/3)/3$.
4. $M_X(t) = \exp(\mu t + \sigma^2 t^2/2)$ ise $E[X]$ nedir?
Çözüm. $E[X] = M'_X(0) = \mu$.
5. $Z \sim N(0, 1)$ ise Z^2 hangi dağılıma sahiptir?
Çözüm. $Z^2 \sim \chi_1^2$ olur.

6. $X \geq 0$ ve $P(X > x) = e^{-x^2}$ ise $E[X]$ nedir?

Çözüm. Kuyruk integrali kullanılır: $E[X] = \int_0^\infty e^{-x^2} dx = \sqrt{\pi}/2$.

7. $X \sim \text{Exp}(\lambda)$ için medyan nedir ve ortalama ile nasıl karşılaştırılır?

Çözüm. Medyan $\log 2/\lambda \approx 0.693/\lambda$, ortalama $1/\lambda$ 'dır. Medyan ortalamadan küçüktür; dağılım sağa çarpıktır.

8. $X, Y \sim U(0, 1)$ bağımsız ise $Z = X + Y$ için $f_Z(1)$ nedir?

Çözüm. Üçgen yoğunlukta tepe noktasıdır: $f_Z(1) = 1$.

9. $X \sim N(0, 1)$ için $Y = X^2$ yoğunluğunu monoton olmayan dönüşüm kuralıyla yazınız.

Çözüm. İki kol $\pm\sqrt{y}$ ve $|g'| = 2\sqrt{y}$ olduğundan $f_Y(y) = \frac{1}{\sqrt{2\pi}} y^{-1/2} e^{-y/2}$, $y > 0$.

10. $X \sim \text{Exp}(\lambda)$ ise $P(X > 3 \mid X > 1)$ nedir?

Çözüm. Hafızasızlık gereği $P(X > 3 \mid X > 1) = P(X > 2) = e^{-2\lambda}$.

11. $\varphi(x) = x^2$ konveks olduğuna göre Jensen eşitsizliği $E[X]$ ve $E[X^2]$ arasında ne söyler?

Çözüm. $E[X]^2 \leq E[X^2]$ olur; fark $\text{Var}(X) \geq 0$ 'a eşittir.

12. Değişim katsayısı 1 olan pozitif sürekli bir model için doğal aday hangisidir?

Çözüm. Üstel dağılım; çünkü $\text{Exp}(\lambda)$ için $\sigma/\mu = 1$ 'dir.

4.12 Bölüm Sonu Alıştırmaları

Bu alıştırmalar, sürekli yoğunluk yorumunu, kantil ve dönüşüm hesaplarını, üstel-normal modelleri ve moment araçlarını pekiştirmek için üç düzeyde düzenlenmiştir. Cevaplar ve kısa çözüm ipuçları Ekler bölümündeki *Cevap Anahtarı*'nda verilmiştir.

Temel Düzey

B1. $X \sim U(2, 8)$ ise $P(3 < X < 6)$ değerini bulunuz.

B2. $X \sim \text{Exp}(0.5)$ ise $P(X > 4)$ değerini hesaplayınız.

B3. X yoğunluğu $f(x) = 2x$, $0 < x < 1$ olan bir değişken olsun. $P(X > 0.5)$ değerini bulunuz.

B4. $Z \sim N(0, 1)$ ise $P(Z \leq 1.96)$ değerini standart normal tablodan yaklaşık yazınız.

B5. $Y = 2X + 1$ ve X yoğunluğu f_X olan sürekli bir değişkense $f_Y(y)$ dönüşüm formülünü yazınız.

Orta Düzey

- I1. $X \sim \text{Exp}(\lambda)$ için 0.90 kantilini bulunuz.
- I2. X yoğunluğu $f(x) = 3x^2$, $0 < x < 1$ olsun. $Y = X^3$ için Y 'nin dağılımını bulunuz.
- I3. $X \sim U(0, 1)$ ve $Y = -\log X/\lambda$ olsun. Y 'nin dağılımını gösteriniz.
- I4. $X, Y \sim U(0, 1)$ bağımsız ise $S = X + Y$ için $0 < s < 1$ aralığında $f_S(s)$ değerini bulunuz.
- I5. $X \sim N(\mu, \sigma^2)$ ise $Z = (X - \mu)/\sigma$ değişkeninin dağılımını yazınız.

İleri Düzey

- A1. X yoğunluğu $f_X(x) = 1/2$, $-1 < x < 1$ olsun. $Y = X^2$ yoğunluğunu bulunuz.
- A2. $X \geq 0$ ve sağkalım fonksiyonu $S(x) = \exp(-x^3)$ ise tehlike fonksiyonu $h(x)$ değerini bulunuz.
- A3. $X \sim \text{Exp}(\lambda)$ için MGF'yi türeterek $E[X]$ ve $\text{Var}(X)$ değerlerini elde ediniz.
- A4. Bağımsız $X, Y \sim \text{Exp}(\lambda)$ için $T = X + Y$ yoğunluğunu evrişimle bulunuz.
- A5. $X > 0$ ve $Y = \log X$ dönüşümü için $f_Y(y)$ ile $f_X(x)$ arasındaki ilişkiyi yazınız. Bu ilişki lognormal dağılımın yoğunluğunu nasıl açıklar?

4.13 Kaynak Notu

Bu bölümde sürekli yoğunluk, dağılım dönüşümleri ve temel dağılım aileleri Hogg et al. (2019), Ross (2019), Hsu (1997), Grinstead ve Snell (1997) ve Soong (2004) çizgisinde düzenlenmiştir. Olasılık integral dönüşümü, MGF tekilliği ve Jensen eşitsizliği için Casella ve Berger (2002) ile Mukhopadhyay (2000); evrişim ve toplam dağılımları için Feller (1971) ve Grimmer ve Stirzaker (2020); karakteristik fonksiyon özellikleri ve ölçü-teorik temeller için Billingsley (1995), Kolmogorov (1956), Capiński ve Kopp (2004) ve Shao (2003) izlenmiştir. Jacobian ve koordinat dönüşümü vurguları Jaynes (2003) ve Tarantola (2001); tehlike fonksiyonu, güvenilirlik ve mühendislik örnekleri Montgomery ve Runger (2018) ile Ramachandran ve Tsokos (2009); kantil ve dayanıklı özetler Wasserman (2004) kaynaklarıyla ilişkilendirilmiştir.

Özet:

Sürekli değişkenlerde olasılık yoğunluk altında kalan alanla hesaplanır. Kantil fonksiyonu simülasyon ve yüzdelikler için kullanılır. Dönüşümlerde yoğunluk Jacobian ile taşınır. Momentler dağılımın şeklini özetler; MGF ve karakteristik fonksiyonlar moment, toplam ve dağılım belirleme problemlerinde temel araçlardır.

Bölüm 5

Ortak Dağılımlar ve Bağımlılık

5.1 Gerçek Dünya Problemi (Giriş)

Bir sigorta şirketi müşterinin yaşı ile hasar tutarını, bir enerji dağıtım şirketi günlük sıcaklık ile elektrik tüketimini, bir yatırımcı iki hisse senedinin getirilerini, bir hekim boy ile kiloyu birlikte gözler. Bu problemlerin ortak yanı, tek bir rassal değişkenin değil, iki veya daha fazla değişkenin *birlikte* davranışının önemli olmasıdır. İki hisse aynı yönde hareket ediyorsa portföy riski artar; sıcaklık ile tüketim ters yönlü ise kapasite planlaması buna göre yapılır.

Tek değişkenli dağılımlar her değişkeni ayrı ayrı betimler; fakat değişkenler arasındaki ilişkiyi göremez. Ortak dağılımlar bu ilişkiyi modellemenin matematiksel dilidir. Kovaryans ve korelasyon bu ilişkinin yönünü ve gücünü tek sayıya indirger; toplamların varyansı, örnekleme teorisinin ve portföy analizinin temel formülünü verir.

Bu Bölüm Ne Sağlar?

Bu bölüm, tek değişkenli dağılım dilini iki ve daha çok değişkene taşır. Önce ortak dağılım fonksiyonu, ortak olasılık kütle ve yoğunluk fonksiyonları ile marjinal dağılımlar tanımlanır; ardından rassal değişkenlerin bağımsızlığı, iki değişkenli beklenti, kovaryans, korelasyon, toplamların varyansı, iki değişkenli normal dağılım ve multinom dağılım ele alınır. Buradaki kavramlar, örnekleme teorisinin, koşullu dağılımların ve regresyon düşüncesinin temelini oluşturur.

Amaç ve Öğrenme Çıktıları

Bu bölümün sonunda okuyucu

- ortak dağılım, ortak PMF ve ortak PDF kavramlarını yazabilir ve kullanabilir,
- ortak dağılımdan marjinal dağılımları elde edebilir,
- rassal değişkenlerin bağımsızlığını ortak dağılımın çarpanlara ayrılmasıyla ifade edebilir,
- kovaryans ve korelasyonu hesaplayıp yorumlayabilir,
- doğrusal birleşimlerin ve toplamların varyansını kovaryans terimleriyle yazabilir,
- iki değişkenli normal dağılımın temel özelliklerini ve multinom dağılımı uygun bağlama yerleştirebilir.

5.2 Ortak Dağılım Fonksiyonu

İki rassal değişkenin birlikte davranışını betimlemenin en genel yolu, her ikisinin de belirli eşiklerin altında kalma olasılığını veren ortak dağılım fonksiyonudur (Ross, 2019; Hogg et al., 2019).

Tanım 5.1 (Ortak Dağılım Fonksiyonu) X ve Y aynı olasılık uzayı üzerinde tanımlı rassal değişkenler olsun. Her $(x, y) \in \mathbb{R}^2$ için

$$F_{X,Y}(x, y) = P(X \leq x, Y \leq y)$$

fonksiyonuna X ile Y 'nin *ortak (birleşik) dağılım fonksiyonu* denir.

Teorem 5.1 (Dikdörtgen Olasılığı). Her $a_1 < b_1$ ve $a_2 < b_2$ için

$$P(a_1 < X \leq b_1, a_2 < Y \leq b_2) = F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) + F(a_1, a_2)$$

olur.

Kanıt. $\{X \leq b_1, Y \leq b_2\}$ olayı, dikdörtgen bölge ile üç “L biçimli” parçanın ayrık birleşimi olarak yazılır. İçerme-dışlama uygulanır: $F(a_1, b_2)$ ve $F(b_1, a_2)$ çıkarıldığında sol alt köşe bölgesi iki kez çıkarılmış olur; $F(a_1, a_2)$ geri eklenerek dikdörtgenin olasılığı elde edilir. $\square$

Örnek 5.1 (Tek Değişkenli Fonksiyonlara İndirgeme). $y \rightarrow \infty$ alındığında $\{Y \leq y\}$ olayı tüm örnek uzaya yaklaşır; dolayısıyla

$$\lim_{y \rightarrow \infty} F_{X,Y}(x, y) = P(X \leq x) = F_X(x)$$

olur. Ortak dağılım fonksiyonu, tek değişkenli dağılım fonksiyonlarını sınır davranışında içerir.

5.3 Kesikli Ortak Dağılımlar ve Marjinaller

Kesikli değişken çiftlerinde ortak davranış, her değer çiftine olasılık atayan ortak olasılık kütle fonksiyonuyla verilir. Ortak tablodan tek değişkenin dağılımına geçme işlemine *marjinalleştirme* denir; adını, tablo kenarlarına (marjlara) yazılan satır ve sütun toplamlarından alır (Blitzstein & Hwang, 2019; Ross, 2019).

Tanım 5.2 (Ortak Olasılık Kütle Fonksiyonu) Kesikli X ve Y için

$$p_{X,Y}(x, y) = P(X = x, Y = y)$$

fonksiyonuna *ortak olasılık kütle fonksiyonu* denir. Bu fonksiyon negatif değildir ve tüm değer çiftleri üzerinde toplamı 1'dir.

Tanım 5.3 (Marjinal PMF) Ortak PMF'si $p_{X,Y}$ olan X değişkeninin *marjinal olasılık kütle fonksiyonu*

$$p_X(x) = \sum_y p_{X,Y}(x, y)$$

olarak tanımlanır; p_Y benzer biçimde x üzerinden toplanarak elde edilir.

Teorem 5.2 (Marjinalin Geçerliliği). p_X fonksiyonu tek başına geçerli bir olasılık kütle fonksiyonudur: negatif değildir ve toplamı 1'dir.

Kanıt. Negatif olmama, negatif olmayan terimlerin toplamından gelir. Ayrıca

$$\sum_x p_X(x) = \sum_x \sum_y p_{X,Y}(x, y) = 1$$

olur; çünkü çifte toplam, ortak PMF'nin tüm kütlelerini toplar. $\square$

Örnek 5.2 (İki Zarın Minimumu ve Maksimumu). İki adil zar atılsın; $X = \min$, $Y = \max$ olsun. $x < y$ için (x, y) çifti iki sıralı sonuçtan, $x = y$ için tek sonuçtan gelir:

$$p_{X,Y}(x, y) = \begin{cases} 2/36, & x < y, \\ 1/36, & x = y, \\ 0, & x > y. \end{cases}$$

Marjinalleştirme ile örneğin $p_X(1) = \sum_{y=1}^6 p_{X,Y}(1, y) = 1/36 + 5 \cdot (2/36) = 11/36$ bulunur. Ortak tablo, minimum ile maksimum arasındaki güçlü bağımlı-

lığı da gösterir: $p_{X,Y}(3,1) = 0$ olması, minimumun maksimumdan büyük olamamasının doğrudan ifadesidir.

Marjinal dağılımlar ortak dağılımdan her zaman elde edilir; fakat ters yönde geçiş genel olarak mümkün değildir. Aynı marjinal çiftine sahip birbirinden çok farklı ortak dağılımlar vardır. Bağımlılık bilgisi yalnız ortak dağılımda saklıdır.

5.4 Sürekli Ortak Dağılımlar ve Marjinaller

Sürekli değişken çiftlerinde olasılık, düzlemdeki bölgeler üzerinde ortak yoğunluğun integraliyle hesaplanır.

Tanım 5.4 (Ortak Yoğunluk Fonksiyonu) Negatif olmayan bir $f_{X,Y}$ fonksiyonu her (x, y) için

$$F_{X,Y}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(u, v) dv du$$

koşulunu sağlıyorsa $f_{X,Y}$ fonksiyonuna (X, Y) çiftinin *ortak olasılık yoğunluk fonksiyonu* denir. Bu durumda düzlemdeki uygun her B bölgesi için

$$P((X, Y) \in B) = \iint_B f_{X,Y}(x, y) dx dy$$

olur.

Tanım 5.5 (Marjinal Yoğunluk) Ortak yoğunluğu $f_{X,Y}$ olan X değişkeninin *marjinal yoğunluğu*

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy$$

olarak tanımlanır; f_Y benzer biçimde x üzerinden integrale bulunur.

Örnek 5.3 (Üçgen Bölgede Düzgün Dağılım). (X, Y) çifti $0 < x < y < 1$ üçgeninde düzgün dağılsın. Üçgenin alanı $1/2$ olduğundan ortak yoğunluk bu bölgede $f_{X,Y}(x, y) = 2$ 'dir. Marjinaller

$$f_X(x) = \int_x^1 2 dy = 2(1-x), \quad 0 < x < 1, \quad f_Y(y) = \int_0^y 2 dx = 2y, \quad 0 < y < 1$$

olur. Ortak dağılım düzgün olduğu halde marjinaller düzgün değildir; kısıt bölgesinin biçimi marjinallere eğim kazandırır. Bu örnek, koşullu dağılımlar bölümünde koşullu yoğunluk hesabıyla yeniden ele alınacaktır.

Örnek 5.4 (Bölge Olasılığı). *Yukarıdaki üçgen modelde $P(Y > 2X)$ olasılığı, $y > 2x$ kısıtını sağlayan alt bölge üzerinde integraldir:*

$$P(Y > 2X) = \int_0^{1/2} \int_{2x}^1 2 \, dy \, dx = \int_0^{1/2} 2(1 - 2x) \, dx = [2x - 2x^2]_0^{1/2} = \frac{1}{2}.$$

5.5 Rassal Değişkenlerin Bağımsızlığı

Olayların bağımsızlığı, rassal değişkenlere ortak dağılımın çarpanlara ayrılması biçiminde taşınır. Bu tanım, rassal örneklem modelinin ve çarpım biçimli olabirliklerin temelidir (Casella & Berger, 2002; Ross, 2019).

Tanım 5.6 (Bağımsız Rassal Değişkenler) X ve Y rassal değişkenleri her x, y için

$$F_{X,Y}(x, y) = F_X(x)F_Y(y)$$

koşulunu sağlıyorsa *bağımsız* adını alır. Kesikli durumda bu koşul $p_{X,Y}(x, y) = p_X(x)p_Y(y)$, sürekli durumda $f_{X,Y}(x, y) = f_X(x)f_Y(y)$ eşitliğine denktir.

Teorem 5.3 (Bağımsızlıkta Çarpım Beklentisi). X ile Y *bağımsız ve ilgili beklentiler sonlu ise her uygun g ve h fonksiyonları için*

$$\mathbb{E}[g(X)h(Y)] = \mathbb{E}[g(X)] \mathbb{E}[h(Y)]$$

olur. Özel olarak $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ sağlanır.

Kanıt. Kesikli durumda ortak PMF çarpanlara ayrılır:

$$\mathbb{E}[g(X)h(Y)] = \sum_x \sum_y g(x)h(y) p_X(x)p_Y(y) = \left(\sum_x g(x)p_X(x) \right) \left(\sum_y h(y)p_Y(y) \right).$$

Sürekli durumda toplamaların yerini integraller alır ve aynı çarpanlara ayırma uygulanır. $\square$

Örnek 5.5 (Bağımsızlığın Yoğunluktan Okunması). $f_{X,Y}(x, y) = 4xy$, $0 < x < 1$, $0 < y < 1$ *yoğunluğu $f_X(x) = 2x$ ile $f_Y(y) = 2y$ çarpımına ayrılır; X ve Y bağımsızdır. Buna karşılık üçgen bölgedeki düzgün model çarpanlara ayıramaz: destek bölgesi dikdörtgen olmadığı için değişkenler bağımlıdır. Destek bölgesinin dikdörtgen olması bağımsızlığın gerekli (ama tek başına yeterli olmayan) koşuludur.*

5.6 İki Değişkenli Beklenti

Tek değişkendeki bilinçsiz istatistikçi yasası (LOTUS), iki değişkenli fonksiyonlara doğrudan genişler; beklentinin doğrusallığının genel ispatı da bu genişlemeden gelir.

Theorem 5.4 (İki Değişkenli LOTUS). *g uygun bir fonksiyon olsun. Kesikli durumda*

$$\mathbb{E}[g(X, Y)] = \sum_x \sum_y g(x, y) p_{X,Y}(x, y),$$

süreklili durumda

$$\mathbb{E}[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y) dx dy$$

olur (ilgili toplam/integral mutlak yakınsadığında).

Kanıt. Tek değişkenli LOTUS kanıtındaki gruplama argümanı çifte toplam ve çifte integral üzerinde aynen tekrarlanır: $g(X, Y)$ 'nin aldığı her değer olasılığı, o değeri üreten (x, y) çiftlerinin ortak olasılıklarının toplamıdır. $\square$

Theorem 5.5 (Beklentinin Doğrusallığı (Genel Kanıt)). *Beklentileri sonlu X, Y ve sabitler a, b için*

$$\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$$

olur; bağımsızlık gerekmez.

Kanıt. $g(x, y) = ax + by$ alınıp iki değişkenli LOTUS uygulanır. Çifte toplam iki parçaya ayrılır; birinci parçada y , ikinci parçada x üzerinden önce toplanarak marjinaler belirir:

$$\mathbb{E}[aX + bY] = a \sum_x x p_X(x) + b \sum_y y p_Y(y) = a\mathbb{E}[X] + b\mathbb{E}[Y].$$

Süreklili durumu aynı biçimde integrallerle yürütülür. $\square$

5.7 Kovaryans

Kovaryans, iki değişkenin ortalamaları etrafında aynı yönde mi yoksa ters yönde mi dalgalandığını ölçer. Pozitif kovaryans birlikte artışı, negatif kovaryans ters yönlü hareketi gösterir (Ross, 2019; Bertsekas & Tsitsiklis, 2008).

Tanım 5.7 (Kovaryans) $\mathbb{E}[X] = \mu_X$ ve $\mathbb{E}[Y] = \mu_Y$ sonlu olsun. X ile Y 'nin kovaryansı

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mu_X)(Y - \mu_Y)]$$

olarak tanımlanır. Özel olarak $\text{Cov}(X, X) = \text{Var}(X)$ olur.

Teorem 5.6 (Kovaryansın Hesaplama Formülü). *İkinci momentler sonlu ise*

$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

olur.

Kanıt. Çarpım açılır ve beklentinin doğrusallığı uygulanır:

$$\mathbb{E}[(X - \mu_X)(Y - \mu_Y)] = \mathbb{E}[XY] - \mu_Y \mathbb{E}[X] - \mu_X \mathbb{E}[Y] + \mu_X \mu_Y = \mathbb{E}[XY] - \mu_X \mu_Y.$$

□

Teorem 5.7 (Kovaryansın Temel Özellikleri). *Sonlu ikinci momentler altında ve sabitler a, b, c için*

$$\text{Cov}(X, Y) = \text{Cov}(Y, X), \quad \text{Cov}(aX + b, cY + d) = ac \text{Cov}(X, Y),$$

$$\text{Cov}(X + Z, Y) = \text{Cov}(X, Y) + \text{Cov}(Z, Y)$$

olur; yani kovaryans simetrik ve her iki değişkeninde doğrusaldır. Sabit eklemek kovaryansı değiştirmez.

Kanıt. Her özellik, tanımın açılıp beklentinin doğrusallığının uygulanmasıyla doğrudan elde edilir. Sabitler merkezleme sırasında düşer; ölçek çarpanları dışarı çıkar; toplam, merkezlenmiş terimlerin toplamına dağılır. □

Teorem 5.8 (Bağımsızlık Kovaryansı Sıfırlar). *X ile Y bağımsız ve ikinci momentler sonlu ise $\text{Cov}(X, Y) = 0$ olur.*

Kanıt. Bağımsızlıkta $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ olduğundan hesaplama formülü $\text{Cov}(X, Y) = 0$ verir. □

Örnek 5.6 (Kovaryansı Sıfır Ama Bağımlı). *X değişkeni $-1, 0, 1$ değerlerini eşit olasılıkla alsın ve $Y = X^2$ olsun. $\mathbb{E}[X] = 0$ ve $\mathbb{E}[XY] = \mathbb{E}[X^3] = 0$ olduğundan $\text{Cov}(X, Y) = 0$ 'dır. Oysa Y , X 'in fonksiyonudur; değişkenler güçlü biçimde bağımlıdır. Kovaryans yalnız doğrusal ilişkiyi ölçer; sıfır kovaryans bağımsızlık anlamına gelmez.*

“Bağımsızlık $\Rightarrow$ kovaryans sıfır” çıkarımı tek yönlüdür. Ters yön yalnız özel ailelerde (örneğin iki değişkenli normal dağılımda) geçerlidir. Sıfır kovaryansı bağımsızlık kanıtı olarak kullanmak yaygın bir modelleme hatasıdır.

5.8 Korelasyon Katsayısı

Kovaryans ölçek birimine bağlıdır: uzunluk metreden santimetreye çevrildiğinde kovaryans 100 katına çıkar. Korelasyon katsayısı, kovaryansı standart sapmalara bölerek birimsiz ve $[-1, 1]$ aralığına sıkıştırılmış bir ilişki ölçüsü verir.

Tanım 5.8 (Korelasyon Katsayısı) $\sigma_X, \sigma_Y > 0$ olmak üzere

$$\rho(X, Y) = \text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

oranına *korelasyon katsayısı* denir.

Teorem 5.9 (Korelasyonun Sınırlılığı). *Sonlu ikinci momentler altında*

$$-1 \leq \rho(X, Y) \leq 1$$

olur. Ayrıca $|\rho| = 1$ olması, olasılık 1 ile $Y = aX + b$ biçiminde tam doğrusal bir ilişkinin varlığına denktir ($\rho = 1$ için $a > 0$, $\rho = -1$ için $a < 0$).

Kanıt. $U = X - \mu_X$, $V = Y - \mu_Y$ olsun. Her gerçekte t için $\mathbb{E}[(U - tV)^2] \geq 0$ yazılır; bu ifade t 'ye göre

$$\sigma_Y^2 t^2 - 2 \text{Cov}(X, Y) t + \sigma_X^2$$

polinomudur. Hiçbir t 'de negatif olmaması diskriminantın pozitif olamamasını gerektirir: $\text{Cov}(X, Y)^2 \leq \sigma_X^2 \sigma_Y^2$, yani $|\rho| \leq 1$. (Bu, Cauchy–Schwarz eşitsizliğinin merkezlenmiş biçimidir; eşitsizlik olasılık eşitsizlikleri bölümünde genel haliyle yeniden ele alınacaktır.) $|\rho| = 1$ ise diskriminant sıfırdır ve polinomun çift katlı kökü t^* için $\mathbb{E}[(U - t^*V)^2] = 0$ olur; beklenen değeri sıfır olan negatif olmayan değişken olasılık 1 ile sıfırdır, dolayısıyla $U = t^*V$ ve Y, X 'in doğrusal fonksiyonudur. $\square$

Yorum:

Korelasyon doğrusal ilişkinin *gücünü ve yönünü* ölçer; nedensellik söylemez. Ayrıca ρ^2 değeri, en iyi doğrusal tahminin açıkladığı değişkenlik oranı olarak yorumlanır; bu yorum koşullu beklenti bölümündeki en iyi doğrusal tahmin teoremiyle kesinleşir.

Örnek 5.7 (Ölçek Değişmezliği). $\rho(aX + b, cY + d) = \text{sgn}(ac) \rho(X, Y)$ olur; çünkü kovaryans ac ile, standart sapmalar $|a|$ ve $|c|$ ile çarpılır. Boy santimetreye ya da metreyle ölçülsün, boy-kilo korelasyonu değişmez.

5.9 Doğrusal Birleşimlerin ve Toplamların Varyansı

Tek değişkende $\text{Var}(aX+b) = a^2 \text{Var}(X)$ kuralı vardı. İki ve daha çok değişkende varyans, kovaryans terimleriyle genişler; bu formül örnekleme teorisinin ve portföy riskinin temel hesabıdır.

Teorem 5.10 (Doğrusal Birleşimin Varyansı). *Sonlu ikinci momentler altında*

$$\text{Var}(aX + bY) = a^2 \text{Var}(X) + b^2 \text{Var}(Y) + 2ab \text{Cov}(X, Y)$$

olur. Daha genel olarak

$$\text{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \text{Var}(X_i) + 2 \sum_{i < j} \text{Cov}(X_i, X_j).$$

Kanıt. $\text{Var}(aX + bY) = \text{Cov}(aX + bY, aX + bY)$ yazılır ve kovaryansın iki değişkendeki doğrusallığı uygulanır; dört terim $a^2 \text{Var}(X)$, $b^2 \text{Var}(Y)$ ve iki kez $ab \text{Cov}(X, Y)$ verir. Genel toplam için aynı açılım tüm çiftlere uygulanır: köşegen terimler varyansları, köşegen dışı terimler kovaryansları üretir. $\square$

Sonuç 5.1 (Bağımsız Toplamda Varyansların Toplanması). $X_1, \dots, X_n$ *ikişerli bağımsız (dolayısıyla kovaryansları sıfır) ise*

$$\text{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \text{Var}(X_i)$$

olur. Örnekleme teorisindeki $\text{Var}(\bar{X}) = \sigma^2/n$ sonucu bu kuralın doğrudan uygulamasıdır.

Örnek 5.8 (İki Hisseli Portföy). *İki hissenin getirileri X ve Y ; $\sigma_X = \sigma_Y = 0.20$ ve ağırlıklar eşit ($a = b = 0.5$) olsun. Portföy getirisi $R = 0.5X + 0.5Y$ için*

$$\text{Var}(R) = 0.25(0.04) + 0.25(0.04) + 2(0.25) \text{Cov}(X, Y) = 0.02 + 0.5 \text{Cov}(X, Y).$$

$\rho = 1$ ise $\text{Var}(R) = 0.04$ (çeşitlendirme yarar sağlamaz); $\rho = 0$ ise $\text{Var}(R) = 0.02$; $\rho = -1$ ise $\text{Var}(R) = 0$ olur. Çeşitlendirmenin riski azaltması, matematiksel olarak kovaryans teriminin küçülmesidir.

Yorum:

İadesiz örneklemede gözlemler arasında negatif kovaryans oluşur; toplamın varyans formülündeki kovaryans terimleri bu nedenle örneklem ortalamasının varyansını küçültür. Örnekleme teorisi bölümündeki sonlu anakütle düzeltme çarpanı $(N - n)/(N - 1)$, tam olarak bu negatif kovaryans teriminin toplanmasından doğar.

5.10 İki Değişkenli Normal Dağılım

Sürekli ortak modellerin en önemlisi iki değişkenli (bivariat) normal dağılımdır. Ölçüm hatası çiftleri, boy-kilo türü antropometrik veriler ve varlık getirileri için klasik modeldir (Casella & Berger, 2002; Hogg et al., 2019).

Tanım 5.9 (İki Değişkenli Normal Dağılım) $\mu_X, \mu_Y \in \mathbb{R}$, $\sigma_X, \sigma_Y > 0$ ve $-1 < \rho < 1$ için ortak yoğunluk

$$f(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\{-Q(x, y)\},$$

$$Q(x, y) = \frac{1}{2(1-\rho^2)} \left[\frac{(x-\mu_X)^2}{\sigma_X^2} - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2} \right].$$

olan (X, Y) çiftine *iki değişkenli normal dağılımlı* denir. Buradaki ρ parametresi tam olarak $\text{Corr}(X, Y)$ değeridir.

Teorem 5.11 (Temel Özellikler). (X, Y) *iki değişkenli normal ise:*

1. *Marjinaler normaldir: $X \sim N(\mu_X, \sigma_X^2)$ ve $Y \sim N(\mu_Y, \sigma_Y^2)$.*
2. *Her doğrusal birleşim $aX + bY$ normal dağılımlıdır.*
3. *$\rho = 0$ ise X ile Y bağımsızdır; yani bu ailede sıfır korelasyon bağımsızlığa denktir.*

Kanıt. (1) ve (2), üstel ifadedeki karesel formun kare tamamlamayla ayrıştırılması ve iç değişken üzerinden integralin standart normal integrale indirgenmesiyle elde edilir. (3) için $\rho = 0$ konduğunda çapraz terim düşer ve yoğunluk

$$f(x, y) = \frac{1}{\sigma_X\sqrt{2\pi}} e^{-(x-\mu_X)^2/(2\sigma_X^2)} \cdot \frac{1}{\sigma_Y\sqrt{2\pi}} e^{-(y-\mu_Y)^2/(2\sigma_Y^2)}$$

biçiminde marjinalerin çarpımına ayrılır; bu, bağımsızlığın tanımıdır. $\square$

Teorem 5.12 (Koşullu Dağılım ve Regresyon Doğrusu). (X, Y) *iki değişkenli normal ise $X = x$ verildiğinde Y 'nin koşullu dağılımı yine normaldir:*

$$Y \mid X = x \sim N\left(\mu_Y + \rho \frac{\sigma_Y}{\sigma_X}(x - \mu_X), \sigma_Y^2(1 - \rho^2)\right).$$

Koşullu ortalama x 'in doğrusal fonksiyonudur; bu doğru, klasik regresyon doğrusudur.

Kanıt. Koşullu yoğunluk, ortak yoğunluğun marjinaline bölünmesiyle elde edilir. Üstel ifadede y 'ye göre kare tamamlandığında, ortalaması $\mu_Y + \rho(\sigma_Y/\sigma_X)(x - \mu_X)$ ve varyansı $\sigma_Y^2(1 - \rho^2)$ olan normal yoğunluk çıkar. Ayrıntılı cebir koşullu dağılımlar bölümündeki koşullu yoğunluk tanımının doğrudan uygulamasıdır. $\square$

Örnek 5.9 (Boy ve Kilo). *Bir yetişkin grubunda boy $X \sim N(170, 10^2)$ (cm), kilo $Y \sim N(75, 12^2)$ (kg) ve $\rho = 0.6$ olsun. Boyu 185 cm olan bir kişide beklenen kilo*

$$\mathbb{E}[Y \mid X = 185] = 75 + 0.6 \cdot \frac{12}{10}(185 - 170) = 75 + 10.8 = 85.8 \text{ kg},$$

koşullu standart sapma $12\sqrt{1 - 0.36} = 9.6$ kg olur. Bilgi (boy) geldikçe belirsizlik 12'den 9.6'ya düşer; azalma oranı $1 - \rho^2$ ile belirlenir.

Marjinallerin her ikisinin normal olması, ortak dağılımın iki değişkenli normal olduğunu *garanti etmez*. Marjinalleri normal olan fakat ortak yapısı normal olmayan çiftler kurulabilir. İki değişkenli normallik, ortak yoğunluğun kendisine ilişkin bir varsayımdır.

5.11 Multinom Dağılım

Binom dağılımı iki kategorili denemeleri sayar. Sonuçları üç veya daha çok kategoriye düşen bağımsız denemelerde kategori sayılarının ortak dağılımı multinom dağılımdır; kombinatorik temeli multinom katsayılarıdır (Feller, 1968).

Tanım 5.10 (Multinom Dağılım) Her denemenin k kategoriden birine $p_1, \dots, p_k$ olasılıklarıyla ($\sum_i p_i = 1$) düştüğü n bağımsız deneme yapılsın. N_i , i 'inci kategorideki deneme sayısı olsun. $n_1 + \dots + n_k = n$ olan her değer ataması için

$$P(N_1 = n_1, \dots, N_k = n_k) = \frac{n!}{n_1! n_2! \dots n_k!} p_1^{n_1} p_2^{n_2} \dots p_k^{n_k}$$

olur ve $(N_1, \dots, N_k) \sim \text{Multinom}(n; p_1, \dots, p_k)$ yazılır.

Teorem 5.13 (Marjinaler ve Kovaryans Yapısı). $(N_1, \dots, N_k) \sim \text{Multinom}(n; p_1, \dots, p_k)$ ise her $i \neq j$ için

$$N_i \sim \text{Bin}(n, p_i), \quad \text{Cov}(N_i, N_j) = -n p_i p_j.$$

Kanıt. Yalnız i 'inci kategoriye odaklanıldığında her deneme “kategori i / diğerleri” ikilisine iner; dolayısıyla $N_i \sim \text{Bin}(n, p_i)$. Kovaryans için gösterge değişkenleri kullanılır: $I_m^{(i)}$, m 'inci denemenin i kategorisine düştüğünün göstergesi olsun. Aynı deneme iki farklı kategoriye birden düşemeyeceğinden $\mathbb{E}[I_m^{(i)} I_m^{(j)}] = 0$; farklı denemeler bağımsız olduğundan $\mathbb{E}[I_m^{(i)} I_{m'}^{(j)}] = p_i p_j$ ($m \neq m'$). Buradan

$$\mathbb{E}[N_i N_j] = n(n-1)p_i p_j, \quad \text{Cov}(N_i, N_j) = n(n-1)p_i p_j - (np_i)(np_j) = -np_i p_j$$

elde edilir. Negatif işaret doğaldır: toplam n sabitken bir kategorinin fazla dolması diğerlerinin payını azaltır. $\square$

Örnek 5.10 (Üç Kategorili Anket). *Bir soruya evet, hayır, kararsız cevapları 0.5, 0.3, 0.2 olasılıklarıyla verilsin ve $n = 100$ kişiye sorulsun. Tam olarak 55 evet, 28 hayır ve 17 kararsız gözlenme olasılığı*

$$\frac{100!}{55! 28! 17!} (0.5)^{55} (0.3)^{28} (0.2)^{17}$$

olur. Ayrıca $\text{Cov}(N_{\text{evet}}, N_{\text{hayır}}) = -100(0.5)(0.3) = -15$: evet sayısı arttıkça hayır sayısı sistematik olarak azalır.

5.12 Örnek Veri Tablosu ve Yazılım Çözümleri

Aşağıdaki küçük veri seti, beş üründe reklam harcaması (x , bin TL) ile satış tutarını (y , bin TL) temsil etsin.

Ürün	1	2	3	4	5
x	2	4	6	8	10
y	30	45	50	70	80

Örneklem ortalamaları $\bar{x} = 6$ ve $\bar{y} = 55$ 'tir. Örneklem kovaryansı ve korelasyonu, sapma çarpımlarının toplamından hesaplanır:

$$s_{xy} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \frac{250}{4} = 62.5, \quad r = \frac{s_{xy}}{s_x s_y} \approx 0.988.$$

Korelasyonun 1'e çok yakın olması, reklam ile satış arasında güçlü ve pozitif doğrusal ilişkiyi gösterir.

```
x <- c(2, 4, 6, 8, 10)
y <- c(30, 45, 50, 70, 80)
cov(x, y)
cor(x, y)
```

Python Çözümü

```
import numpy as np

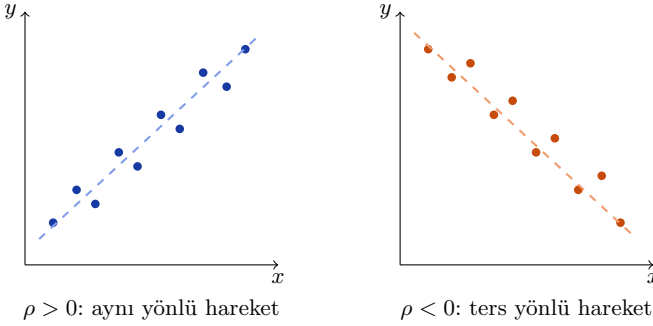
x = np.array([2, 4, 6, 8, 10])
y = np.array([30, 45, 50, 70, 80])

kovaryans = np.cov(x, y, ddof=1)[0, 1]
korelasyon = np.corrcoef(x, y)[0, 1]
```

```
print(kovaryans, korelasyon)
```

5.13 Görselleştirme ve Grafikler

Korelasyonun işareti ve gücü, saçılım grafiğinde noktaların bir doğru etrafında ne kadar sıkı toplandığıyla görülür.



Şekil 5.1 Saçılım grafikleri: pozitif korelasyonda noktalar artan, negatif korelasyonda azalan bir doğru etrafında toplanır. $|\rho|$ büyüdükçe noktalar doğruya yaklaşır.

Dikkat: Kavram Yanılgısı

Sıfır kovaryans bağımsızlık demek değildir. Kovaryans yalnız doğrusal ilişkiyi ölçer; $Y = X^2$ gibi doğrusal olmayan tam bağımlılık durumlarında kovaryans sıfır olabilir.

Kendini Dene:

Ara Kontrol: Bir ortak dağılımdan marjinaler bulunabilir; fakat yalnız marjinallerden ortak dağılım genellikle neden belirlenemez?

5.14 Bölüm Sonu Çözümlü Problemler

1. $p_{X,Y}(x,y)$ ortak PMF tablosu $p(0,0) = 0.3$, $p(0,1) = 0.2$, $p(1,0) = 0.1$, $p(1,1) = 0.4$ olsun. $p_X(1)$ ve $p_Y(1)$ değerlerini bulun.

Çözüm. $p_X(1) = p(1,0) + p(1,1) = 0.1 + 0.4 = 0.5$ ve $p_Y(1) = p(0,1) + p(1,1) = 0.2 + 0.4 = 0.6$.

2. Yukarıdaki tabloda X ile Y bağımsız mıdır?

Çözüm. $p(1, 1) = 0.4$ fakat $p_X(1)p_Y(1) = 0.5 \cdot 0.6 = 0.3 \neq 0.4$. Çarpanlara ayrılma tek bir hücrede bile bozulduğundan değişkenler bağımlıdır.

3. Aynı tablo için $\text{Cov}(X, Y)$ değerini hesaplayın.

Çözüm. $\mathbb{E}[X] = 0.5$, $\mathbb{E}[Y] = 0.6$ ve $\mathbb{E}[XY] = 1 \cdot 1 \cdot p(1, 1) = 0.4$. Dolayısıyla

$$\text{Cov}(X, Y) = 0.4 - (0.5)(0.6) = 0.1.$$

4. $f_{X,Y}(x, y) = cxy$, $0 < x < 1$, $0 < y < 1$ yoğunluğunda c sabitini bulun ve bağımsızlığı inceleyin.

Çözüm. $\int_0^1 \int_0^1 xy \, dx \, dy = 1/4$ olduğundan $c = 4$. Yoğunluk $f_X(x)f_Y(y) = (2x)(2y)$ biçiminde çarpanlara ayrıldığı için X ile Y bağımsızdır.

5. (X, Y) , $0 < x < y < 1$ üçgeninde düzgün dağılsın ($f = 2$). $P(X < 1/4)$ olasılığını marjinal yoğunlukla bulun.

Çözüm. $f_X(x) = 2(1 - x)$ olduğundan

$$P(X < 1/4) = \int_0^{1/4} 2(1 - x) \, dx = [2x - x^2]_0^{1/4} = \frac{1}{2} - \frac{1}{16} = \frac{7}{16}.$$

6. $\text{Var}(X) = 9$, $\text{Var}(Y) = 16$ ve $\text{Cov}(X, Y) = 6$ ise $\text{Var}(X + Y)$ ve $\text{Var}(X - Y)$ nedir?

Çözüm. $\text{Var}(X + Y) = 9 + 16 + 2(6) = 37$ ve $\text{Var}(X - Y) = 9 + 16 - 2(6) = 13$.

7. $\sigma_X = 3$, $\sigma_Y = 4$ ve $\text{Cov}(X, Y) = 6$ ise korelasyon katsayısı nedir?

Çözüm. $\rho = 6/(3 \cdot 4) = 0.5$.

8. $Y = 3X + 2$ ise $\rho(X, Y)$ nedir?

Çözüm. $\text{Cov}(X, Y) = 3 \text{Var}(X)$ ve $\sigma_Y = 3\sigma_X$ olduğundan $\rho = 3\sigma_X^2/(3\sigma_X^2) = 1$. Tam artan doğrusal ilişki $\rho = 1$ verir.

9. X değişkeni $-1, 0, 1$ değerlerini eşit olasılıkla alsın ve $Y = X^2$ olsun. $\text{Cov}(X, Y)$ değerini bulun ve sonucu yorumlayın.

Çözüm. $\mathbb{E}[X] = 0$ ve $\mathbb{E}[XY] = \mathbb{E}[X^3] = (-1 + 0 + 1)/3 = 0$ olduğundan $\text{Cov}(X, Y) = 0$. Buna rağmen Y , X tarafından tamamen belirlenir: sıfır kovaryans bağımsızlık demek değildir.

10. İki hissenin getirileri için $\sigma_X = \sigma_Y = 0.2$ ve $\rho = -0.5$ olsun. Eşit ağırlıklı portföy $R = 0.5X + 0.5Y$ için $\text{Var}(R)$ nedir?

Çözüm. $\text{Cov}(X, Y) = \rho\sigma_X\sigma_Y = -0.02$ olur.

$$\text{Var}(R) = 0.25(0.04) + 0.25(0.04) + 2(0.25)(-0.02) = 0.02 - 0.01 = 0.01.$$

Negatif korelasyon, tek hissenin varyansı 0.04'ü portföyde 0.01'e indirir.

11. (X, Y) iki değişkenli normal; $\mu_X = 50$, $\mu_Y = 100$, $\sigma_X = 5$, $\sigma_Y = 20$ ve $\rho = 0.8$ olsun. $\mathbb{E}[Y | X = 55]$ ve $\text{Var}(Y | X = 55)$ değerlerini bulun.

Çözüm.

$$\mathbb{E}[Y | X = 55] = 100 + 0.8 \cdot \frac{20}{5}(55 - 50) = 100 + 16 = 116,$$

$$\text{Var}(Y \mid X = 55) = 20^2(1 - 0.64) = 144.$$

12. $(N_1, N_2, N_3) \sim \text{Multinom}(60; 0.5, 0.3, 0.2)$ olsun. $\mathbb{E}[N_2]$, $\text{Var}(N_2)$ ve $\text{Cov}(N_1, N_2)$ değerlerini bulun.

Çözüm. Marjinal $N_2 \sim \text{Bin}(60, 0.3)$ olduğundan $\mathbb{E}[N_2] = 18$ ve $\text{Var}(N_2) = 60(0.3)(0.7) = 12.6$. Kovaryans yapısından $\text{Cov}(N_1, N_2) = -60(0.5)(0.3) = -9$.

5.15 Bölüm Sonu Alıştırmaları

Bu alıştırmalar, ortak dağılım tablolarını, marjinalleştirme ve koşullama işlemlerini, bağımsızlık testini, kovaryans-korelasyon hesaplarını ve çok değişkenli model yorumlarını pekiştirmek için üç düzeyde düzenlenmiştir. Cevaplar ve kısa çözüm ipuçları Ekler bölümündeki *Cevap Anahtarı*'nda verilmiştir.

Temel Düzey

- B1. Ortak PMF $p(0, 0) = 0.2$, $p(0, 1) = 0.3$, $p(1, 0) = 0.1$, $p(1, 1) = 0.4$ olsun. $p_X(0)$ ve $p_Y(0)$ değerlerini bulunuz.
- B2. Aynı ortak PMF için $P(X = 1, Y = 1)$ ile $P(X = 1)P(Y = 1)$ değerlerini karşılaştırarak bağımsızlığı inceleyiniz.
- B3. $E[X] = 2$, $E[Y] = 5$ ve $E[XY] = 12$ ise $\text{Cov}(X, Y)$ değerini hesaplayınız.
- B4. $\text{Var}(X) = 4$, $\text{Var}(Y) = 9$ ve $\text{Cov}(X, Y) = 3$ ise $\text{Var}(X + Y)$ değerini bulunuz.
- B5. $\sigma_X = 2$, $\sigma_Y = 5$ ve $\text{Cov}(X, Y) = -4$ ise korelasyon katsayısını hesaplayınız.

Orta Düzey

- I1. $f_{X,Y}(x, y) = c$, $0 < x < 1$, $0 < y < 2$ bölgesinde ortak yoğunluk olsun. c değerini ve $P(X < 1/2, Y > 1)$ olasılığını bulunuz.
- I2. $f_{X,Y}(x, y) = 6xy^2$, $0 < x < 1$, $0 < y < 1$ olsun. X ve Y bağımsız mıdır?
- I3. Ortak PMF $p(0, 0) = 0.1$, $p(0, 1) = 0.2$, $p(1, 0) = 0.3$, $p(1, 1) = 0.4$ olsun. $P(Y = 1 \mid X = 1)$ değerini bulunuz.
- I4. $\text{Var}(X) = 1$, $\text{Var}(Y) = 4$ ve $\rho = -0.5$ ise $\text{Var}(2X - Y)$ değerini hesaplayınız.
- I5. X ve Y bağımsız, $E[X] = 1$, $E[Y] = 3$, $\text{Var}(X) = 2$, $\text{Var}(Y) = 5$ ise $E[XY]$ ve $\text{Var}(X + Y)$ değerlerini bulunuz.

İleri Düzey

- A1. (X, Y) , $0 < x < y < 1$ üçgeninde düzgün dağılsın. $f_Y(y)$ marjinal yoğunluğunu bulunuz.
- A2. X değişkeni -1 ve 1 değerlerini eşit olasılıkla alsın, $Y = X^2$ olsun. X ve Y bağımsız mıdır? Kovaryansla birlikte yorumlayınız.
- A3. (X, Y) iki değişkenli normal; $\mu_X = 10$, $\mu_Y = 20$, $\sigma_X = 2$, $\sigma_Y = 5$, $\rho = 0.6$ olsun. $E[Y | X = 12]$ ve $\text{Var}(Y | X = 12)$ değerlerini bulunuz.
- A4. $(N_1, N_2, N_3) \sim \text{Multinom}(100; 0.2, 0.5, 0.3)$ için $\text{Cov}(N_1, N_3)$ ve $\text{Corr}(N_1, N_3)$ değerlerini hesaplayınız.
- A5. X ve Y bağımsız standart normal değişkenler olsun. $U = X + Y$ ve $V = X - Y$ için $\text{Cov}(U, V)$ değerini bulunuz ve normal varsayımı altında bağımsızlık hakkında yorum yapınız.

5.16 Kaynak Notu

Bu bölümdeki ortak dağılım, marjinalleştirme, bağımsızlık, kovaryans, korelasyon ve toplamların varyansı sunumu; Ross (2019), Hogg, McKean ve Craig (2019), Blitzstein ve Hwang (2019), Bertsekas ve Tsitsiklis (2008) ve Casella ve Berger (2002) çizgisindeki klasik olasılık ve matematiksel istatistik kaynaklarıyla uyumludur. İki değişkenli normal dağılım ve koşullu regresyon yapısı Casella ve Berger (2002) ile Hogg, McKean ve Craig (2019); multinom dağılımın kombinatorik temeli ve kovaryans yapısı Feller (1968); ortak dağılımların ölçü-teorik çerçevesi Billingsley (1995) ve Grimmett ve Stirzaker (2020) izlenerek verilmiştir.

Özet:

Ortak dağılımlar, iki veya daha çok rassal değişkenin birlikte davranışını betimler; marjinal dağılımlar ortak dağılımdan toplanarak ya da integrale elde edilir, fakat tersi mümkün değildir. Bağımsızlık, ortak dağılımın marjinallerin çarpımına ayrılmasıdır. Kovaryans doğrusal birlikteliğin yönünü, korelasyon ise birimsiz gücünü ölçer; sıfır kovaryans bağımsızlık anlamına gelmez. Toplamların varyansı kovaryans terimleriyle genişler ve örnekleme teorisi ile portföy analizinin temel formülünü verir. İki değişkenli normal dağılımda koşullu ortalama doğrusaldır ve sıfır korelasyon bağımsızlığa denktir; multinom dağılım çok kategorili sayımların ortak modelidir.

Bölüm 6

Örnekleme Teorisi ve Örnekleme Dağılımları

6.1 Gerçek Dünya Problemi (Giriş)

Bir fabrikanın ürettiği tüm ampullerin ortalama ömrünü, bir ülkedeki tüm seçmenlerin oy tercihlerini veya bir hastanenin tüm hastalarında ortalama iyileşme süresini çoğu zaman tek tek gözleyemeyiz. Bunun yerine anakütleden seçilen sınırlı sayıda gözlemle çalışırız. Temel soru şudur: Sadece örneklem verisine bakarak anakütle hakkında ne kadar güvenilir bilgi üretebiliriz?

Bu Bölüm Ne Sağlar?

Bu bölüm, anakütle-örneklem ayırımından başlayarak rassal örneklem, istatistik, örnekleme dağılımı, örneklem ortalaması, örneklem varyansı, standart hata, sonlu anakütle düzeltmesi, sıralı istatistikler, normal anakütlerde kesin dağılımlar ve büyük örneklem yaklaşımlarını sistematik biçimde ele alır. Amaç, çıkarımsal istatistiğin temelindeki “istatistiğin dağılımı” fikrini olasılık kuramı içinde açık hâle getirmektir.

Amaç ve Öğrenme Çıktıları

Bu bölümün sonunda okuyucu

- anakütle, örneklem, parametre ve istatistik kavramlarını ayırt edebilir,
- rassal örneklemün ortak dağılımını yazabilir,
- örneklem ortalaması ve örneklem varyansının beklenen değer ve varyansını hesaplayabilir,
- standart hata ve sonlu anakütle düzeltmesini yorumlayabilir,
- normal anakütlerde $\bar{X}$, S^2 , Z , t ve χ^2 dağılımlarının nasıl ortaya çıktığını açıklayabilir,
- merkezi limit teoremini örnekleme dağılımlarına uygulayabilir.

6.2 Anakütle, Örneklem ve İstatistik

İstatistiksel çıkarımda anakütle ilgilenilen tüm birimleri, örneklem ise anakütleden gözlenen sınırlı alt kümeyi temsil eder. Anakütleyle ait bilinmeyen sayısal özelliklere parametre, örneklemden hesaplanan sayısal değerlere istatistik denir. Bu ayrım, olasılık modelinden istatistiksel çıkarıma geçişin temelidir (Hogg, McKean & Craig, 2019; Ross, 2019).

Tanım 6.1 (Anakütle ve Parametre) İlgilenilen tüm birimlerin oluşturduğu topluluğa anakütle denir. Anakütle dağılımının ortalaması μ , varyansı σ^2 , oranı p gibi sabit fakat çoğu zaman bilinmeyen niceliklere parametre denir.

Tanım 6.2 (Örneklem ve İstatistik) Anakütleden gözlenen $X_1, \dots, X_n$ rasal değişkenlerine örneklem denir. Örneklemden hesaplanan ve bilinmeyen parametre içermeyen her fonksiyona istatistik denir:

$$T = T(X_1, \dots, X_n).$$

Örnek 6.1 (Ortalama Ömür). *Bir ampul üretim hattında anakütle, üretilen tüm ampullerin ömür dağılımıdır. Anakütle ortalaması μ bilinmeyen parametredir. Seçilen n ampulün ömürleri $X_1, \dots, X_n$ ise*

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

örneklem ortalaması adlı bir istatistiktir.

Parametre sabit, istatistik rassaldır. Çünkü parametre anakütle dağılımının özelliğidir; istatistik ise seçilen örneklemeye göre değişir. Örneklemeye dağılımı fikri bu ayırmadan doğar.

Örneklemin anakütle dağılımını ne ölçüde temsil ettiğini gösteren en doğrudan istatistik, gözlemleri basamak basamak biriktiren ampirik dağılım fonksiyonudur. Bu fonksiyon hem grafiksel betimlemenin hem de birçok dayanıklı yöntemin temelidir (Wasserman, 2004; van der Vaart, 1998).

Tanım 6.3 (Ampirik Dağılım Fonksiyonu) $X_1, \dots, X_n$ örneklemini için ampirik dağılım fonksiyonu

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}\{X_i \leq x\}$$

olarak tanımlanır; yani x 'ten küçük veya eşit gözlemlerin oranıdır.

Teorem 6.1 (Ampirik Dağılımın Yansızlığı). *Her sabit x için $nF_n(x) \sim \text{Bin}(n, F(x))$ olur ve*

$$E(F_n(x)) = F(x), \quad \text{Var}(F_n(x)) = \frac{F(x)[1 - F(x)]}{n}$$

sağlanır.

Kanıt. Sabit x için $\mathbf{1}\{X_i \leq x\}$ değişkenleri bağımsız Bernoulli($F(x)$)'dir; çünkü $P(X_i \leq x) = F(x)$. Bağımsız Bernoulli göstergelerinin toplamı binom dağılır. Binom ortalaması $nF(x)$ ve varyansı $nF(x)[1 - F(x)]$ olduğundan n 'e bölündüğünde verilen değerler elde edilir. $\square$

Örnek 6.2 (Ampirik Oranın Tahmini). *Bir aralıktaki gözlem oranı, o aralığın gerçek olasılığını yansıtarak tahmin eder. Örneğin $n = 200$ gözlemden 150'si bir eşik altında ise $F_n(\text{eşik}) = 0.75$ olur ve bu, $F(\text{eşik})$ için doğal bir tahmindir.*

6.3 Rassal Örneklem

Olasılık kuramında en temel örnekleme modeli bağımsız ve aynı dağılımlı gözlemlerden oluşan rassal örneklemdir. Bu varsayım, her gözlemin aynı anakütle dağılımından geldiğini ve gözlemler arasında olasılıksal bağımlılık bulunmadığını söyler.

Tanım 6.4 (Rassal Örneklem) $X_1, \dots, X_n$ rassal değişkenleri bağımsız ve aynı dağılıma sahipse bu diziye büyüklüğü n olan rassal örneklem denir. Ortak dağılım kesikli durumda

$$P(X_1 = x_1, \dots, X_n = x_n) = \prod_{i=1}^n p(x_i),$$

sürekli durumda ise

$$f_{X_1, \dots, X_n}(x_1, \dots, x_n) = \prod_{i=1}^n f(x_i)$$

şeklinde dir.

Örnek 6.3 (Bernoulli Örneklemi). *Bir ürünün kusurlu olup olmadığı X_i ile gösterilsin. Kusurlu ürün için $X_i = 1$, sağlam ürün için $X_i = 0$ alınır ve $X_i \sim \text{Bernoulli}(p)$ olur. Bağımsız n ürün incelendiğinde $X_1, \dots, X_n$ bir Bernoulli örneklemdir.*

Teorem 6.2 (Rassal Örneklemenin Olabilirlik Çekirdeği). $X_1, \dots, X_n$ yoğunluğu veya olasılık fonksiyonu $f(x; \theta)$ olan bir dağılımdan rassal örneklem olsun. Gözlenen $x_1, \dots, x_n$ için ortak model

$$L(\theta; x_1, \dots, x_n) = \prod_{i=1}^n f(x_i; \theta)$$

biçimindedir.

Kanıt. Rassal örneklem tanımı gereği gözlemler bağımsızdır ve her biri aynı $f(x; \theta)$ modeline sahiptir. Bağımsız rassal değişkenlerin ortak dağılımı marjinal dağılımların çarpımı olduğundan sonuç elde edilir. $\square$

Rassal örneklemelerin doğrusal birleşimlerinin dağılımını bulmada moment üreten fonksiyon güçlü bir araçtır. Özellikle normal örnekleme toplamının yine normal olması, örneklem teorisinin pek çok kesin sonucunun kaynağıdır (Casella & Berger, 2002).

Teorem 6.3 (Bağımsız Normallerin Doğrusal Birleşimi). $X_i \sim N(\mu_i, \sigma_i^2)$ bağımsız ve $a_1, \dots, a_n$ sabitler olsun. Bu durumda

$$\sum_{i=1}^n a_i X_i \sim N\left(\sum_{i=1}^n a_i \mu_i, \sum_{i=1}^n a_i^2 \sigma_i^2\right)$$

olur.

Kanıt. $N(\mu_i, \sigma_i^2)$ değişkeninin moment üreten fonksiyonu $\exp(\mu_i t + \sigma_i^2 t^2 / 2)$ 'dir. Bağımsızlık nedeniyle $\sum a_i X_i$ 'nin MGF'si bireysel MGF'lerin çarpımıdır:

$$\prod_{i=1}^n \exp\left(a_i \mu_i t + \frac{1}{2} a_i^2 \sigma_i^2 t^2\right) = \exp\left(t \sum a_i \mu_i + \frac{1}{2} t^2 \sum a_i^2 \sigma_i^2\right).$$

Bu, belirtilen parametrelili normal dağılımın MGF'sidir; MGF tekillliğiyle sonuç kanıtlanır. $\square$

Örnek 6.4 (İki Bağımsız Ölçümün Farkı). $X \sim N(10, 4)$ ve $Y \sim N(8, 9)$ bağımsız ise $X - Y \sim N(10 - 8, 4 + 9) = N(2, 13)$ olur. Farkın standart sapması $\sqrt{13} \approx 3.61$ 'dir.

6.4 Örneklem Dağılımı

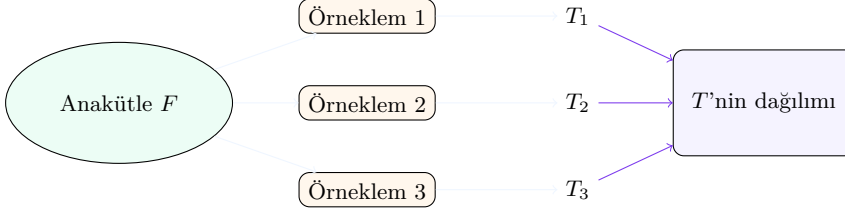
Bir istatistik örneklemden örneklem değiştiği için kendisinin de bir olasılık dağılımı vardır. Bu dağılıma örneklem dağılımı denir. Güven aralıkları, hipotez testleri ve tahmin edici değerlendirmeleri bu dağılımlar üzerine kurulur.

Tanım 6.5 (Örneklem Dağılımı) $T = T(X_1, \dots, X_n)$ bir istatistik olsun. T 'nin olasılık dağılımına örneklem dağılımı denir.

Örnek 6.5 (Yazı-Tura Oranı). n kez atılan adil olmayan bir paranın yazı gelme olasılığı p olsun. $X_i \sim \text{Bernoulli}(p)$ ve

$$\hat{p} = \frac{1}{n} \sum_{i=1}^n X_i$$

örneklem oranıdır. $\sum X_i \sim \text{Bin}(n, p)$ olduğundan $\hat{p}$ yalnızca $0, 1/n, 2/n, \dots, 1$ değerlerini alır ve örnekleme dağılımı binom dağılımından türetilir.



Şekil 6.1 Örnekleme dağılımı: aynı anakütleden alınan her örneklem bir istatistik değeri üretir; bu değerlerin dağılımı örnekleme dağılımıdır.

6.5 Örneklem Ortalamasının Dağılımı

Örneklem ortalaması, anakütle ortalamasını tahmin etmek için en temel istatistiktir. Dağılımı tam olarak her zaman kolay bulunmasa da beklenen değeri ve varyansı oldukça genel koşullarda hesaplanabilir.

Teorem 6.4 (Örneklem Ortalamasının Beklenen Değeri ve Varyansı). $X_1, \dots, X_n$ ortalaması μ ve varyansı $\sigma^2 < \infty$ olan bir anakütleden rassal örneklem olsun. O halde

$$E(\bar{X}) = \mu, \quad \text{Var}(\bar{X}) = \frac{\sigma^2}{n}.$$

Kanıt. Beklenen değer in doğrusallığıyla

$$E(\bar{X}) = E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \mu.$$

Bağımsızlık nedeniyle toplamın varyansı varyansların toplamıdır:

$$\text{Var}(\bar{X}) = \text{Var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n \text{Var}(X_i) = \frac{\sigma^2}{n}.$$

□

Tanım 6.6 (Standart Hata) Bir istatistiğin örnekleme dağılımının standart sapmasına standart hata denir. Örneklem ortalaması için

$$SE(\bar{X}) = \frac{\sigma}{\sqrt{n}}$$

olur.

Yorum:

Örneklem büyüklüğü dört katına çıkarıldığında ortalamanın standart hatası yarıya iner. Bu nedenle örneklem büyüklüğünü artırmanın getirisi doğrusal değil, karekök hızındadır.

Örnek 6.6 (Standart Hatanın Hesabı). *Anakütle standart sapması $\sigma = 20$ olsun. $n = 25$ için $SE(\bar{X}) = 20/\sqrt{25} = 4$; $n = 100$ için $SE(\bar{X}) = 20/\sqrt{100} = 2$ olur. Örneklem dört katına çıktığında standart hata yarıya inmiştir. Aynı kesinliği iki katına çıkarmak için örnekleme dört katına çıkarmak gerekir.*

6.6 Örneklem Varyansı

Anakütle varyansı bilinmediğinde değişkenliği örneklemden tahmin ederiz. Bu amaçla paydada n yerine $n - 1$ kullanan örneklem varyansı temel rol oynar.

Tanım 6.7 (Örneklem Varyansı) $X_1, \dots, X_n$ örnekleme için örneklem varyansı

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

olarak tanımlanır.

Teorem 6.5 (Örneklem Varyansının Yansızlığı). $X_1, \dots, X_n$ ortalaması μ ve varyansı $\sigma^2 < \infty$ olan bir anakütleden rassal örneklem olsun. O halde

$$E(S^2) = \sigma^2.$$

Kanıt. Temel ayrışım

$$\sum_{i=1}^n (X_i - \bar{X})^2 = \sum_{i=1}^n (X_i - \mu)^2 - n(\bar{X} - \mu)^2$$

şeklindedir. Beklenen değer alınırsa

$$E \left[\sum_{i=1}^n (X_i - \bar{X})^2 \right] = n\sigma^2 - n \frac{\sigma^2}{n} = (n-1)\sigma^2.$$

Her iki taraf $n - 1$ 'e bölündüğünde $E(S^2) = \sigma^2$ bulunur. $\square$

Hesaplama da örneklem varyansı çoğu zaman sapmaların karesini tek tek toplamak yerine kareler toplamı üzerinden bulunur. Bu özdeşlik hem elle hem de yazılımla hesabı kolaylaştırır (Mood, Graybill & Boes, 1974).

Teorem 6.6 (Örneklem Varyansının Hesaplama Formülü).

$$\sum_{i=1}^n (X_i - \bar{X})^2 = \sum_{i=1}^n X_i^2 - n\bar{X}^2$$

olur; dolayısıyla

$$S^2 = \frac{1}{n-1} \left(\sum_{i=1}^n X_i^2 - n\bar{X}^2 \right).$$

Kanıt. Kareyi açalım:

$$\sum_i (X_i - \bar{X})^2 = \sum_i X_i^2 - 2\bar{X} \sum_i X_i + n\bar{X}^2.$$

$\sum_i X_i = n\bar{X}$ yerine konursa orta terim $-2n\bar{X}^2$ olur ve ifade $\sum_i X_i^2 - n\bar{X}^2$ 'ye sadeleşir. $\square$

Örnek 6.7 (Sayısal Örneklem Varyansı). 5, 7, 9 gözlemleri için $\bar{X} = 7$, $\sum X_i^2 = 25 + 49 + 81 = 155$ ve $n\bar{X}^2 = 3 \cdot 49 = 147$ olur. Buradan kareler toplamı $155 - 147 = 8$ ve

$$S^2 = \frac{8}{3-1} = 4, \quad S = 2$$

bulunur.

6.7 Sonlu Anakütleden Örneklem

Bazı uygulamalarda anakütle sonludur ve örneklem iadesiz yapılır. Bu durumda gözlemler bağımsız değildir; örneklem ortalamasının varyansı bağımsız örneklem durumuna göre daha küçüktür.

Teorem 6.7 (Sonlu Anakütle Düzeltmesi). *Büyüklüğü N olan sonlu bir anakütleden iadesiz basit rassal örneklem ile n birim seçilsin. Anakütle varyansı σ^2 ile gösterilirse örneklem ortalamasının varyansı*

$$\text{Var}(\bar{X}) = \frac{\sigma^2}{n} \left(\frac{N-n}{N-1} \right)$$

biçimindedir.

Kanıt. İadesiz örneklemde farklı gözlemler arasında negatif kovaryans vardır. Toplamın varyansı

$$\text{Var} \left(\sum_{i=1}^n X_i \right) = \sum_{i=1}^n \text{Var}(X_i) + 2 \sum_{i < j} \text{Cov}(X_i, X_j)$$

şeklinde açılır. Sonlu anakütlerde bu kovaryans terimleri hesaba katıldığında çarpan $(N - n)/(N - 1)$ elde edilir. $n = 1$ için düzeltme 1'dir; $n = N$ için varyans 0 olur. $\square$

Sonlu anakütle düzeltmesi yalnızca anakütlenin anlamlı bir kısmı iadesiz seçildiğinde önemlidir. N çok büyük ve n/N küçükse düzeltme yaklaşık 1 olur.

Örnek 6.8 (Düzeltilmenin Etkisi). $N = 500$ anakütleden iadesiz $n = 50$ birim seçilsin ve $\sigma^2 = 100$ olsun. Bağımsız örneklem varyansı $100/50 = 2$ 'dir. Sonlu anakütle çarpanı $(500 - 50)/(499) \approx 0.902$ olduğundan düzeltilmiş varyans $2 \times 0.902 \approx 1.80$, standart hata $\sqrt{1.80} \approx 1.34$ olur. $N = 50000$ olsaydı çarpan ≈ 0.999 ile ihmal edilebilirdi.

6.8 Normal Anakütlerde Kesin Örneklem Dağılımları

Normal dağılım, örneklem teorisinde özel bir yere sahiptir. Eğer anakütle normal ise örneklem ortalamasının dağılımı tam olarak normaldir; ayrıca örneklem ortalaması ile örneklem varyansı bağımsızdır.

Teorem 6.8 (Normal Anakütlerde Örneklem Ortalaması). $X_1, \dots, X_n \sim N(\mu, \sigma^2)$ bağımsız olsun. O halde

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

ve

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$$

olur.

Kanıt. Bağımsız normal rassal değişkenlerin doğrusal birleşimi normaldir. $\bar{X}$ bu değişkenlerin doğrusal birleşimidir. Önceki teoremden ortalaması μ , varyansı σ^2/n olur. Standartlaştırma $N(0, 1)$ dağılımını verir. $\square$

Teorem 6.9 (Ki-Kare ve Student t Dağılımları). $X_1, \dots, X_n \sim N(\mu, \sigma^2)$ bağımsız olsun. Bu durumda

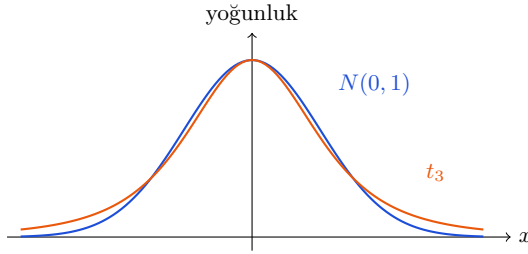
$$\frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$$

ve $\bar{X}$ ile S^2 bağımsızdır. Ayrıca

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}.$$

Kanıt. Normal örnekleme toplam kareler ortalamaya ait bir bileşen ve artık bileşenine ayrılır. Artık bileşen $n - 1$ serbestlik dereceli ki-kare dağılımına sahiptir. Normal örnekleme ortalama ve artık kareler bağımsızdır. Standart normal değişkenin bağımsız ki-kare değişkeninin serbestlik derecesine bölünmüş kareköküne oranı Student t dağılımını verir. $\square$

Örnek 6.9 (Bilinmeyen Varyansla Standartlaştırma). *Anakütle standart sapması σ bilinmediğinde, standartlaştırmada σ yerine S kullanılır. Bu durumda Z yerine $T = (\bar{X} - \mu)/(S/\sqrt{n})$ değişkeni elde edilir ve dağılımı normal değil t_{n-1} olur. Küçük örneklerde bu fark belirgindir; t dağılımının kuyrukları normalden ağırdır.*



Şekil 6.2 Student t dağılımı normale göre daha basık tepeli ve daha ağır kuyrukludur; serbestlik derecesi büyüdükçe normale yaklaşır.

6.9 Örneklem Oranı ve Binom Modeli

İkili sonuçlarda örneklem oranı, anakütle oranı p için doğal tahmin edicidir. Kusurlu-sağlam, başarılı-başarısız, evet-hayır gibi değişkenler bu çerçevede modellenir.

Teorem 6.10 (Örneklem Oranının Ortalama ve Varyansı). $X_1, \dots, X_n \sim \text{Bernoulli}(p)$ bağımsız olsun ve

$$\hat{p} = \frac{1}{n} \sum_{i=1}^n X_i$$

tanımlansın. O halde

$$E(\hat{p}) = p, \quad \text{Var}(\hat{p}) = \frac{p(1-p)}{n}.$$

Ayrıca $Y = \sum_{i=1}^n X_i \sim \text{Bin}(n, p)$ olduğundan $\hat{p} = Y/n$ olur.

Kant. Bernoulli değişkeni için $E(X_i) = p$ ve $\text{Var}(X_i) = p(1 - p)$ değerleri kullanılır. Örneklem ortalaması teoremi doğrudan sonucu verir. Toplamın binom dağılımına sahip olması bağımsız Bernoulli denemelerinin tanımından gelir. $\square$

6.10 Sıralı İstatistikler

Bazı istatistikler gözlemlerin büyüklük sırasına dayanır. Minimum, maksimum, medyan ve kantiller sıralı istatistiklerle ifade edilir.

Tanım 6.8 (Sıralı İstatistik) $X_1, \dots, X_n$ örneklemini küçükten büyüğe sıralandığında elde edilen değişkenler

$$X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$$

şeklinde gösterilir. $X_{(k)}$ değişkenine k 'nci sıralı istatistik denir.

Teorem 6.11 (Minimum ve Maksimum Dağılımları). *Dağılım fonksiyonu F olan sürekli bir anakütleden rassal örneklem alınsın. Bu durumda*

$$P(X_{(1)} > x) = [1 - F(x)]^n, \quad P(X_{(n)} \leq x) = [F(x)]^n.$$

Dolayısıyla

$$F_{X_{(1)}}(x) = 1 - [1 - F(x)]^n, \quad F_{X_{(n)}}(x) = [F(x)]^n.$$

Kant. $X_{(1)} > x$ olayı tüm gözlemlerin x 'ten büyük olmasıdır. Bağımsızlık nedeniyle olasılık $[1 - F(x)]^n$ olur. $X_{(n)} \leq x$ olayı ise tüm gözlemlerin x veya altında olmasıdır; olasılığı $[F(x)]^n$ 'dir. $\square$

Tek tek sıralı istatistiklerin tam yoğunluğu, kombinatorik bir argümanla yazılabilir. Bu sonuç medyan, çeyrek ve uç değer dağılımlarının ortak temelidir (David & Nagaraja, 2003; Casella & Berger, 2002).

Teorem 6.12 (k 'nci Sıralı İstatistiğin Yoğunluğu). *Yoğunluğu f , dağılım fonksiyonu F olan sürekli anakütleden n büyüklüğünde örneklem için $X_{(k)}$ yoğunluğu*

$$f_{X_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!} [F(x)]^{k-1} [1 - F(x)]^{n-k} f(x)$$

olur.

Kant. $X_{(k)}$ 'nin $[x, x + dx)$ aralığında olması için gözlemlerden biri bu aralıkta ($\approx f(x)dx$), $k-1$ tanesi x 'in altında ($[F(x)]^{k-1}$) ve $n-k$ tanesi x 'in üstünde ($[1 - F(x)]^{n-k}$) olmalıdır. Bu birimleri seçmenin yol sayısı $\frac{n!}{(k-1)!(n-k)!}$ multinom katsayısıdır; çarpıldığında yoğunluk elde edilir. $\square$

Örnek 6.10 (Düzgün Örneklemde Medyan). $X_i \sim U(0, 1)$ ve $n = 2m + 1$ tek olsun. Ortanca sıralı istatistik $X_{(m+1)}$ bir $\text{Beta}(m + 1, m + 1)$ dağılımına sahiptir; beklenen değeri $1/2$, dağılımı $1/2$ etrafında simetriktir. Bu, örneklem medyanının anakütle medyanını yansız tahmin ettiği sezgisini destekler.

Örnek 6.11 (Açıklık). Örneklem açıklığı $R = X_{(n)} - X_{(1)}$ ile tanımlanır. Maksimum ve minimumun ortak dağılımından açıklığın dağılımı türetilir; kalite kontrolde süreç değişkenliğinin hızlı bir ölçüsü olarak kullanılır.

6.11 Merkezi Limit Teoremi ve Büyük Örneklem Yaklaşımı

Örneklem teorinin en güçlü sonuçlarından biri merkezi limit teoremidir. Anakütle normal olmasa bile, uygun koşullarda örneklem ortalamasının standartlaştırılmış hali büyük n için yaklaşık standart normal dağılır.

Teorem 6.13 (Merkezi Limit Teoremi). $X_1, X_2, \dots$ bağımsız ve aynı dağılımlı, $E(X_i) = \mu$ ve $\text{Var}(X_i) = \sigma^2 < \infty$ olsun. O halde

$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} N(0, 1)$$

yakınsaması geçerlidir.

Kanıt. Bu sonuç bağımsız aynı dağılımlı değişkenler için klasik merkezi limit teoremidir. Moment üreten fonksiyon veya karakteristik fonksiyon yöntemiyle ispatlanabilir. Standartlaştırılmış toplamın karakteristik fonksiyonu limitte $\exp(-t^2/2)$ fonksiyonuna, yani standart normal dağılımın karakteristik fonksiyonuna yakınsar. $\square$

Örnek 6.12 (Ortalama Teslimat Süresi). Bir kargo firmasının teslimat süresi sağa çarpık bir dağılıma sahip olsun. Tek bir teslimat süresi normal olmayabilir; ancak $n = 100$ teslimatın ortalaması, sonlu varyans koşulu altında yaklaşık normal dağılır. Bu nedenle ortalama teslimat süresi için yaklaşık hata payı hesaplanabilir.

Merkezi limit teoreminin tarihsel olarak ilk özel hâli, binom dağılımının normal yaklaşımını veren De Moivre–Laplace teoremidir. Oran ve sayım verilerinde büyük örneklem çıkarımının temelini oluşturur (Feller, 1968; Montgomery & Runger, 2018).

Teorem 6.14 (De Moivre–Laplace). $Y \sim \text{Bin}(n, p)$ olsun. $n \rightarrow \infty$ için

$$\frac{Y - np}{\sqrt{np(1-p)}} \xrightarrow{d} N(0, 1)$$

yakınsaması geçerlidir. Eşdeğer olarak örneklem oranı için

$$\frac{\hat{p} - p}{\sqrt{p(1-p)/n}} \xrightarrow{d} N(0, 1).$$

Kanıt. $Y = \sum_{i=1}^n X_i$ biçiminde bağımsız *Bernoulli*(p) değişkenlerinin toplamı yazılır. Bu değişkenlerin ortalaması p , varyansı $p(1-p)$ sonludur. Merkezi limit teoremi doğrudan uygulanır ve standartlaştırılmış toplam standart normale yakınsar. $\square$

Yorum:

Binom olasılıklarını normalle yaklaşık hesaplarken, kesikli bir değişkeni sürekli bir dağılımla değerlendirdiğimiz için süreklilik düzeltmesi kullanılır: $P(Y \leq k) \approx \Phi((k+0.5-np)/\sqrt{np(1-p)})$. Yaklaşım $np \geq 5$ ve $n(1-p) \geq 5$ olduğunda iyi çalışır.

Örnek 6.13 (Oran için Normal Yaklaşım). $n = 400$ ve $p = 0.5$ için $\hat{p}$ 'nin standart hatası $\sqrt{0.25/400} = 0.025$ olur. De Moivre–Laplace gereği $\hat{p}$ yaklaşık $N(0.5, 0.025^2)$ dağılır; dolayısıyla $\hat{p}$ 'nin yaklaşık %95 olasılıkla 0.5 ± 0.049 aralığında olması beklenir.

Uygulamada bilinmeyen σ yerine tutarlı bir tahmin kullanmak ya da ortalamanın bir fonksiyonunun dağılımını bulmak gerekir. Slutsky teoremi ve Delta yöntemi bu iki ihtiyacı karşılar (van der Vaart, 1998; Casella & Berger, 2002).

Teorem 6.15 (Slutsky Teoremi). $Z_n \xrightarrow{d} Z$ ve $W_n \xrightarrow{p} c$ (sabit) olsun. Bu durumda

$$Z_n + W_n \xrightarrow{d} Z + c, \quad W_n Z_n \xrightarrow{d} cZ$$

olur.

Kanıt. Olasılıkta sabite yakınsama, dağılımda yakınsamayı bozmayacak biçimde toplam ve çarpım işlemlerine taşınır. Ortak dağılımın limit davranışı (Z, c) çiftine indirgendiğinden sürekli dönüşüm teoremi ile sonuç elde edilir. $\square$

Örnek 6.14 (σ Yerine S Kullanmak). Büyük örnekleme $S \xrightarrow{p} \sigma$ olduğundan, $(\bar{X} - \mu)/(\sigma/\sqrt{n}) \xrightarrow{d} N(0, 1)$ ifadesinde σ yerine S konabilir. Slutsky teoremi gereği $(\bar{X} - \mu)/(S/\sqrt{n})$ de standart normale yakınsar; bu, büyük örnekleme t ve z yaklaşımlarının örtüşmesini açıklar.

Teorem 6.16 (Delta Yöntemi). $\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{d} N(0, \sigma^2)$ ve g fonksiyonu μ noktasında türevlenebilir, $g'(\mu) \neq 0$ olsun. Bu durumda

$$\sqrt{n}(g(\bar{X}_n) - g(\mu)) \xrightarrow{d} N(0, [g'(\mu)]^2 \sigma^2)$$

olur.

Kanıt. g fonksiyonu μ çevresinde birinci dereceden Taylor açılımıyla $g(\bar{X}_n) \approx g(\mu) + g'(\mu)(\bar{X}_n - \mu)$ biçiminde yaklaştırılır. Kalan terim olasılıkla ihmal edilebilir mertebededir. $\sqrt{n}(\bar{X}_n - \mu)$ 'nin limit normalliği $g'(\mu)$ çarpanıyla ölçeklenerek aktarılır. $\square$

Örnek 6.15 (Logaritmik Dönüşüm). $\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{d} N(0, \sigma^2)$ ve $\mu > 0$ ise $g(x) = \log x$ için $g'(\mu) = 1/\mu$ olur. Delta yöntemi gereği $\sqrt{n}(\log \bar{X}_n - \log \mu) \xrightarrow{d} N(0, \sigma^2/\mu^2)$. Bu, oransal değişkenlerde logaritmik ölçeğin varyansı kararlı kılma nedenini gösterir.

6.12 Örnek Veri Tablosu ve Yazılım Çözümleri

Aşağıdaki küçük veri seti, üretim hattından seçilen 12 parçanın dayanım değerlerini temsil etsin.

Parça	1	2	3	4	5	6	7	8	9	10	11	12
Dayanım	48	52	51	49	55	50	47	53	54	49	52	50

Bu veriler için örneklem ortalaması, örneklem standart sapması ve ortalamanın standart hatası hesaplanır. Anakütle standart sapması bilinmediği için standart hata $S/\sqrt{n}$ ile tahmin edilir.

```
x <- c(48, 52, 51, 49, 55, 50, 47, 53, 54, 49, 52, 50)
mean(x)
sd(x)
sd(x) / sqrt(length(x))
```

Python Çözümü

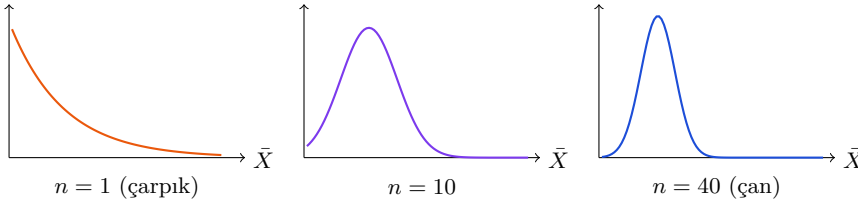
```
import numpy as np

x = np.array([48, 52, 51, 49, 55, 50, 47, 53, 54, 49, 52, 50])
n = len(x)
ortalama = np.mean(x)
s = np.std(x, ddof=1)
standart_hata = s / np.sqrt(n)

print(n, ortalama, s, standart_hata)
```

6.13 Görselleştirme ve Grafikler

Örneklem dağılımını sezgisel olarak görmek için aynı anakütleden çok sayıda örneklem alınır ve her örneklem ortalaması kaydedilir. Tekil gözlemler çarpık dağılmış olsa bile örneklem ortalamalarının histogramı çoğu zaman daha simetrik görünür.



Şekil 6.3 Merkezi limit teoremi: anakütle çarpık olsa bile örneklem büyüklüğü arttıkça örneklem ortalamasının dağılımı çan eğrisine yaklaşır ve daralır.

Aşağıdaki benzetim, çarpık bir anakütleden alınan ortalamaların histogramını üretir.

```
import numpy as np
import matplotlib.pyplot as plt

rng = np.random.default_rng(42)
ortalamalar = []

for _ in range(10000):
    orneklem = rng.exponential(scale=10, size=40)
    ortalamalar.append(np.mean(orneklem))

plt.hist(ortalamalar, bins=40, density=True, edgecolor="black")
plt.xlabel("Örneklem ortalaması")
plt.ylabel("Yoğunluk")
plt.title("Örneklem Dağılımı: n = 40")
plt.show()
```

Dikkat: Kavram Yanılgısı

Örneklem dağılımı, verinin histogramı değildir. Verinin histogramı tek örneklemdeki gözlemleri gösterir; örneklem dağılımı ise aynı örneklem süreci tekrarlandığında bir istatistiğin nasıl değişeceğini anlatır.

Kendini Dene:

Ara Kontrol: Örneklem büyüklüğü iki katına çıkınca $\bar{X}$ değerinin standart hatası neden yarıya değil, $1/\sqrt{2}$ katına iner?

6.14 Bölüm Sonu Çözümlü Problemler

1. Bir anakütlenin ortalaması $\mu = 80$, standart sapması $\sigma = 12$ olsun. $n = 36$ büyüklüğünde rassal örneklem için $E(\bar{X})$ ve $SE(\bar{X})$ değerlerini bulun.

Çözüm. Örneklem ortalamasının beklenen değeri anakütle ortalamasına eşittir:

$$E(\bar{X}) = 80.$$

Standart hata

$$SE(\bar{X}) = \frac{12}{\sqrt{36}} = 2$$

olur.

2. $X_1, \dots, X_{25} \sim N(100, 15^2)$ bağımsız olsun. $\bar{X}$ dağılımını yazın.

Çözüm. Normal anakütleden gelen örneklemde ortalama tam olarak normal dağılır:

$$\bar{X} \sim N\left(100, \frac{15^2}{25}\right) = N(100, 9).$$

Standart sapması 3'tür.

3. Bir üretim sürecinde kusurlu oranı $p = 0.08$ olsun. $n = 200$ ürün için örneklem oranının beklenen değeri ve standart hatası nedir?

Çözüm.

$$E(\hat{p}) = p = 0.08, \quad SE(\hat{p}) = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.08 \cdot 0.92}{200}} \approx 0.0192.$$

4. $N = 1000$ büyüklüğündeki sonlu bir anakütleden iadesiz $n = 100$ birim seçiliyor. Bağımsız örnekleme varyansına göre sonlu anakütle düzeltme çarpanını bulun.

Çözüm. Düzeltme çarpanı

$$\frac{N-n}{N-1} = \frac{1000-100}{999} = \frac{900}{999} \approx 0.901.$$

Standart hata düzeyinde düzeltme bunun kareköküdür.

5. $X_1, \dots, X_n$ dağılım fonksiyonu F olan sürekli bir anakütleden gelsin. Maksimumun x 'ten küçük veya eşit olma olasılığını bulun.

Çözüm. Maksimumun x 'ten küçük veya eşit olması tüm gözlemlerin x 'ten küçük veya eşit olması demektir:

$$P(X_{(n)} \leq x) = P(X_1 \leq x, \dots, X_n \leq x) = [F(x)]^n.$$

6. Gözlemler 4, 8, 9, 11, 13 olsun. Hesaplama formülüyle örneklem varyansını bulun.

Çözüm. $\bar{X} = 9$, $\sum X_i^2 = 16 + 64 + 81 + 121 + 169 = 451$, $n\bar{X}^2 = 5 \cdot 81 = 405$. Kareler toplamı $451 - 405 = 46$ ve $S^2 = 46/4 = 11.5$.

7. $X_i \sim U(0, 1)$, $n = 2m + 1$ tek olsun. Örneklem medyanının dağılımı nedir? **Çözüm.** Ortanca sıralı istatistik $X_{(m+1)} \sim \text{Beta}(m+1, m+1)$ olur; beklenen değeri $1/2$, dağılımı $1/2$ etrafında simetrik.

8. $Y \sim \text{Bin}(100, 0.5)$ için $P(Y \leq 60)$ değerini süreklilik düzeltmeli normal yaklaşımla tahmin edin.

Çözüm. $np = 50$, $\sqrt{np(1-p)} = 5$. $z = (60 + 0.5 - 50)/5 = 2.1$, dolayısıyla $P(Y \leq 60) \approx \Phi(2.1) \approx 0.982$.

9. Sabit bir x için ampirik dağılım fonksiyonu $F_n(x)$ 'in varyansı nedir ve n büyüdükçe nasıl davranır?

Çözüm. $\text{Var}(F_n(x)) = F(x)[1 - F(x)]/n$ olur; $n \rightarrow \infty$ iken sıfıra gider, yani $F_n(x)$ giderek $F(x)$ 'e yoğunlaşır.

6.15 Bölüm Sonu Alıştırmaları

- E1. Örneklem dağılımının ham verinin dağılımı değil, bir istatistiğin dağılımı olduğunu açıklayın.
 E2. Varyansı $\sigma^2 = 25$ olan bir anakütleden $n = 100$ büyüklüğünde örneklem alındığında $\bar{X}$ için standart hatayı hesaplayın.
 E3. Merkezi limit teoreminin iyi bir normal yaklaşım verdiği pratik bir koşul belirtin.
 E4. Bir örneklem yöntemi değerlendirirken yanlılık ve varyansın rollerini karşılaştırın.

6.16 Kaynak Notu

Bu bölümdeki örneklem dağılımı, yansız örneklem varyansı, normal örneklemde t ve χ^2 dağılımları ile merkezi limit teoremi sunumu Hogg, McKean ve Craig (2019), Ross (2019), Montgomery ve Runger (2018), Casella ve Berger (2002) ve Mood, Graybill ve Boes (1974) çizgisindeki klasik olasılık ve matematiksel istatistik yaklaşımıyla uyumludur. Ampirik dağılım fonksiyonu ve dayanıklı özetler için Wasserman (2004) ile van der Vaart (1998); sıralı istatistiklerin tam dağılımları için David ve Nagaraja (2003); De Moivre–Laplace teoremi ve binom normal yaklaşımı için Feller (1968) izlenmiştir.

Özet:

Örnekleme teorisi, örneklemden hesaplanan istatistiklerin rassal değişken olduğunu vurgular. Rassal örnekleme örnekleme ortalaması μ etrafında yoğunlaşır ve varyansı σ^2/n olur. Örneklem varyansı S^2 , σ^2 için yansızdır. Normal anakütlede Z , t ve χ^2 dağılımları kesin olarak elde edilir; normal olmayan anakütlelerde merkezi limit teoremi büyük örneklem yaklaşımını sağlar.

Bölüm 7

Asimptotik Teori ve Limit Teoremleri

7.1 Gerçek Dünya Problemi (Giriş)

Bir e-ticaret platformu her dakika binlerce sipariş, arama ve ödeme isteği alır. Tek bir kullanıcının davranışı düzensiz olabilir; bazı kullanıcılar çok hızlı işlem yapar, bazıları uzun süre bekler, bazıları ise sistemi hiç kullanmaz. Buna rağmen yöneticinin ilgilendiği büyüklük çoğu zaman tek tek kullanıcılar değil, toplam yük, ortalama yanıt süresi veya hata oranıdır. Benzer biçimde bir üretim hattında tek ürünün dayanımı değişken olabilir; fakat çok sayıda üründen elde edilen ortalama dayanım daha kararlı davranır.

Asimptotik teori bu kararlılığı matematiksel olarak açıklar. Örneklem büyüklüğü arttığında rassal ortalamalar, toplamlar ve tahmin ediciler hangi sınırlara yaklaşır? Yaklaşım hangi anlamda gerçekleşir? Büyük örneklemelerde normal dağılım neden bu kadar sık ortaya çıkar? Bu sorular, modern istatistiksel çıkarımın ve simülasyon temelli hesaplamanın temelidir.

Bu Bölüm Ne Sağlar?

Bu bölüm, rassal değişken dizilerinde yakınsama kavramlarını tanıtır; olasılıkta, dağılımda, ortalama karede ve hemen hemen kesin yakınsama arasındaki farkları açıklar. Ardından büyük sayılar yasası, merkezi limit teoremi, süreklilik düzeltmesi, delta yöntemi, Slutsky teoremi, sürekli dönüşüm teoremi ve asimptotik tahmin edici mantığı ele alınır. Amaç, büyük örneklem sonuçlarını yalnız formül olarak değil, çıkarımsal düşünmenin omurgası olarak kurmaktır.

Amaç ve Öğrenme Çıktıları

Bu bölümün sonunda okuyucu

- farklı yakınsama türlerini tanımlayabilir ve aralarındaki ilişkileri açıklayabilir,
- zayıf ve kuvvetli büyük sayılar yasalarının ne söylediğini yorumlayabilir,
- merkezi limit teoremini toplam, ortalama ve oran problemlerine uygulayabilir,
- kesikli dağılımlar için normal yaklaşımda süreklilik düzeltmesini kullanabilir,
- sürekli dönüşüm, Slutsky ve delta yöntemini temel düzeyde uygulayabilir,
- tutarlılık ve asimptotik normallik kavramlarını tahmin ediciler bağlamında açıklayabilir.

7.2 Asimptotik Düşünce

Asimptotik teori, örneklem büyüklüğü n sonsuza giderken rassal niceliklerin davranışını inceler. Burada amaç sonsuz büyüklükte veri elde etmek değildir; amaç, büyük fakat sonlu örneklemelerde hangi yaklaşık kuralların güvenilir olduğunu anlamaktır. Bu yaklaşım Hogg, McKean ve Craig (2019), Casella ve Berger (2002), Ross (2019) ve Billingsley (1995) çizgisindeki modern olasılık ve istatistik kitaplarında temel bir yer tutar.

Tanım 7.1 (Rassal Değişken Dizisi) Aynı olasılık uzayı üzerinde tanımlı $X_1, X_2, \dots$ rassal değişkenlerine rassal değişken dizisi denir. Asimptotik teori bu dizinin $n \rightarrow \infty$ iken bir rassal değişkene veya sabite yaklaşıp yaklaşmadığını inceler.

Örnek 7.1 (Örneklem Ortalaması Dizisi). $X_1, X_2, \dots$ bağımsız ve aynı dağılımlı gözlemler olsun. Her n için

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$$

başka bir rassal değişkendir. Dolayısıyla $\bar{X}_1, \bar{X}_2, \dots$ bir rassal değişken dizisidir. Büyük sayılar yasası bu dizinin anakütle ortalamasına yaklaşmasını açıklar.

Asimptotik sonuçlar “ n sonsuzdur” demek değildir. Bu sonuçlar, büyük n için kullanılabilecek yaklaşık hesap kuralları verir. Yaklaşımın kalitesi dağılımın kuyruğuna, bağımsızlık varsayımına, varyansın sonlu olup olmasına ve örneklem büyüklüğüne bağlıdır.

Asimptotik analizde rassal niceliklerin büyüklük mertebesi kısa biçimde ifade etmek için stokastik mertebe gösterimi kullanılır. Bu gösterim, hata terimlerinin ne kadar hızlı küçüldüğünü izlemeyi kolaylaştırır (van der Vaart, 1998).

Tanım 7.2 (Stokastik Mertebe: O_p ve o_p) $X_n = o_p(1)$ ifadesi $X_n \xrightarrow{p} 0$ anlamına gelir. $X_n = O_p(1)$ ise her $\varepsilon > 0$ için öyle bir M ve N vardır ki $n \geq N$ için $P(|X_n| > M) < \varepsilon$ olur; yani dizi olasılıkta sınırlıdır. Genel olarak $X_n = o_p(a_n)$ ile $X_n/a_n \xrightarrow{p} 0$ ve $X_n = O_p(a_n)$ ile X_n/a_n olasılıkta sınırlı kastedilir.

Örnek 7.2 (Ortalamanın Mertebesi). *Merkezi limit teoreminden $\sqrt{n}(\bar{X}_n - \mu) = O_p(1)$ olur; dolayısıyla $\bar{X}_n - \mu = O_p(1/\sqrt{n})$ yazılır. Bu, ortalamanın anakütle değerine $1/\sqrt{n}$ hızında yaklaştığını kısaca ifade eder.*

7.3 Yakınsama Türleri

Rassal değişkenler sayılar gibi tek bir yoldan yakınsamaz. Bir dizi bazen olasılık bakımından yaklaşır, bazen dağılım bakımından yaklaşır, bazen de neredeyse her örneklem yolu üzerinde yaklaşır. Bu nedenle yakınsama türlerini ayırmak gerekir.

Tanım 7.3 (Olasılıkta Yakınsama) X_n dizisi ve X rassal değişkeni için her $\varepsilon > 0$ değerinde

$$P(|X_n - X| > \varepsilon) \rightarrow 0$$

ise X_n , X 'e olasılıkta yakınsar denir ve

$$X_n \xrightarrow{p} X$$

yazılır.

Tanım 7.4 (Dağılımda Yakınsama) X_n dizisinin dağılım fonksiyonu F_n , X değişkeninin dağılım fonksiyonu F olsun. F 'nin sürekli olduğu her x noktasında

$$F_n(x) \rightarrow F(x)$$

ise X_n , X 'e dağılımda yakınsar denir ve

$$X_n \xrightarrow{d} X$$

yazılır.

Tanım 7.5 (Ortalama Karede Yakınsama)

$$E[(X_n - X)^2] \rightarrow 0$$

ise X_n , X 'e ortalama karede yakınsar denir. Bu yakınsama türü özellikle varyans hesaplarıyla kullanışlıdır.

Tanım 7.6 (Hemen Hemen Kesin Yakınsama)

$$P\left(\lim_{n \rightarrow \infty} X_n = X\right) = 1$$

ise X_n , X 'e hemen hemen kesin yakınsar denir ve $X_n \xrightarrow{a.s.} X$ yazılır.

Teorem 7.1 (Yakınsama İlişkileri). *Aşağıdaki çıkarımlar geçerlidir:*

$$X_n \xrightarrow{a.s.} X \implies X_n \xrightarrow{P} X \implies X_n \xrightarrow{d} X.$$

Ayrıca X_n değişkenleri X 'e ortalama karede yakınsıyorsa $X_n \xrightarrow{P} X$ olur.

Kanıt. Hemen hemen kesin yakınsama, örneklem yollarının olasılık 1 ile yakınsadığını söyler; bu durumda belirli bir ε eşliğini aşan yolların olasılığı sıfıra gider ve olasılıkta yakınsama elde edilir. Olasılıkta yakınsama, dağılım fonksiyonlarının süreklilik noktalarında limit dağılıma yaklaşmasını sağlar. Ortalama karede yakınsama için Markov eşitsizliği uygulanır:

$$P(|X_n - X| > \varepsilon) \leq \frac{E[(X_n - X)^2]}{\varepsilon^2} \rightarrow 0.$$

□

Yorum:

Dağılımda yakınsama en zayıf yakınsama türlerinden biridir. Merkezi limit teoreminde kullanılan yakınsama dağılımda yakınsamadır; çünkü standartlaştırılmış ortalama sabit bir değere değil, standart normal dağılıma yaklaşır.

Yakınsama türleri arasındaki bağlantıların çoğu Markov ve Chebyshev eşitsizlikleri üzerinden kurulur. Bu iki eşitsizlik, olasılık sınırlarının asimptotik analizdeki temel araçlarıdır (Billingsley, 1995; Grimmett & Stirzaker, 2020).

Teorem 7.2 (Markov ve Chebyshev Eşitsizlikleri). $X \geq 0$ ve $a > 0$ için

$$P(X \geq a) \leq \frac{E[X]}{a}$$

olur (Markov). $E[X] = \mu$, $\text{Var}(X) = \sigma^2$ ve $k > 0$ için

$$P(|X - \mu| \geq k) \leq \frac{\sigma^2}{k^2}$$

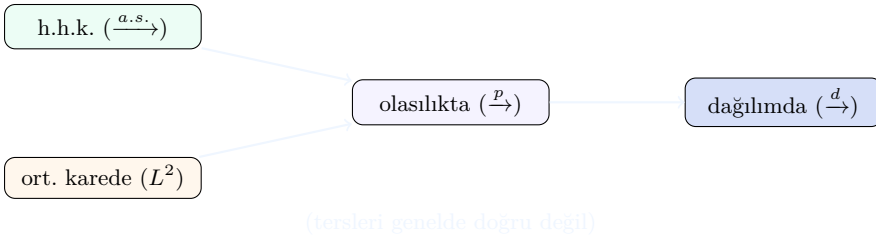
olur (Chebyshev).

Kanıt. Markov için $a \mathbf{1}\{X \geq a\} \leq X$ eşitsizliğinin beklentisi alınır: $aP(X \geq a) \leq E[X]$. Chebyshev için Markov eşitsizliği $(X - \mu)^2 \geq 0$ değişkenine ve $a = k^2$ eşğine uygulanır:

$$P(|X - \mu| \geq k) = P((X - \mu)^2 \geq k^2) \leq \frac{E[(X - \mu)^2]}{k^2} = \frac{\sigma^2}{k^2}.$$

□

Örnek 7.3 (Olasılıkta Yakınsama Ama Hemen Hemen Kesin Değil). $[0, 1]$ üzerinde düzgün seçilen bir noktada, giderek daralan ama art arda tüm aralığı tarayan göstergeler düşünölsün. Bu göstergeler olasılıkta sıfıra yakınsar; çünkü “1” olma olasılığı giderek küçölür. Ancak her nokta sonsuz kez “1” değerini gördüğünden hemen hemen kesin yakınsama yoktur. Bu, olasılıkta yakınsamanın hemen hemen kesin yakınsamadan kesinlikle zayıf olduğunu gösteren klasik örnektir.



Şekil 7.1 Yakınsama türleri arasındaki çıkarım ilişkileri. Hemen hemen kesin ve ortalama karede yakınsama olasılıkta yakınsamayı, o da dağılımda yakınsamayı gerektirir.

7.4 Büyük Sayılar Yasası

Büyük sayılar yasası, örneklem ortalamasının anakütle ortalaması etrafında kararlılaştığını söyler. Bu sonuç, frekans yorumunun, Monte Carlo simölasyonunun ve tutarlı tahmin edici fikrinin temelidir.

Teorem 7.3 (Zayıf Büyük Sayılar Yasası). $X_1, X_2, \dots$ bağımsız ve aynı dağılımlı, $E(X_i) = \mu$ ve $\text{Var}(X_i) = \sigma^2 < \infty$ olsun. O halde

$$\bar{X}_n \xrightarrow{P} \mu.$$

Kanıt. Örneklem ortalaması için $E(\bar{X}_n) = \mu$ ve $\text{Var}(\bar{X}_n) = \sigma^2/n$ olur. Chebyshev eşitsizliğiyle her $\varepsilon > 0$ için

$$P(|\bar{X}_n - \mu| \geq \varepsilon) \leq \frac{\text{Var}(\bar{X}_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \rightarrow 0.$$

Bu, olasılıkta yakınsama tanımıdır. $\square$

Teorem 7.4 (Kuvvetli Büyük Sayılar Yasası). $X_1, X_2, \dots$ bağımsız ve aynı dağılımlı, $E|X_i| < \infty$ ve $E(X_i) = \mu$ olsun. O halde

$$\bar{X}_n \xrightarrow{a.s.} \mu.$$

Kuvvetli yakınsamanın temel aracı, olayların “sonsuz kez” gerçekleşip gerçekleşmeyeceğini olasılıkların toplamına bağlayan Borel–Cantelli lemmasıdır (Billingsley, 1995).

Teorem 7.5 (Borel–Cantelli Lemması). $A_1, A_2, \dots$ olayları için $\sum_{n=1}^{\infty} P(A_n) < \infty$ ise

$$P(A_n \text{ sonsuz kez gerçekleşir}) = 0$$

olur.

Kanıt. $\{A_n \text{ sonsuz kez}\} = \bigcap_m \bigcup_{n \geq m} A_n$ olarak yazılır. Her m için birleşimin olasılığı alt-toplamlanabilirlikle $\sum_{n \geq m} P(A_n)$ ile sınırlıdır. Toplam yakınsadığından bu kuyruk toplamı $m \rightarrow \infty$ iken sifıra gider; dolayısıyla kesişimin olasılığı sıfırdır. $\square$

Kuvvetli BSY için fikir. Sapma olaylarının $A_n = \{|\bar{X}_n - \mu| > \varepsilon\}$ olasılıkları yeterince hızlı azaldığında $\sum_n P(A_n) < \infty$ olur. Borel–Cantelli lemması bu sapmaların yalnız sonlu kez gerçekleştiğini, dolayısıyla $\bar{X}_n$ ’nin μ ’ya hemen hemen kesin yakınsadığını verir. Genel durumda uygun kesme teknikleriyle aynı sonuca ulaşılır. $\square$

Örnek 7.4 (Monte Carlo İntegrali). $U_1, \dots, U_n \sim \text{Uniform}(0, 1)$ bağımsız olsun. g integrallenebilir ise

$$\frac{1}{n} \sum_{i=1}^n g(U_i) \xrightarrow{p} E[g(U)] = \int_0^1 g(u) du.$$

Bu nedenle integral, rassal noktalar üzerinden alınan ortalama ile yaklaşık hesaplanabilir.

Büyük sayılar yasasının tarihsel olarak ilk biçimi, bağıl frekansın olasılığa yakınsamasını söyleyen Bernoulli teoremidir. Olasılığın frekans yorumunun matematiksel temelidir (Feller, 1968).

Teorem 7.6 (Bernoulli Teoremi (Frekans Yorumu)). $Y_n \sim \text{Bin}(n, p)$ olsun. Bağıl frekans $\hat{p}_n = Y_n/n$ için her $\varepsilon > 0$ değerinde

$$P(|\hat{p}_n - p| \geq \varepsilon) \rightarrow 0, \quad \text{yani} \quad \hat{p}_n \xrightarrow{p} p$$

olur.

Kanıt. $\hat{p}_n$ bağımsız *Bernoulli*(p) göstergelerinin ortalamasıdır; ortalaması p , varyansı $p(1 - p)/n$ 'dir. Chebyshev eşitsizliği uygulanır:

$$P(|\hat{p}_n - p| \geq \varepsilon) \leq \frac{p(1 - p)}{n\varepsilon^2} \rightarrow 0.$$

□

Örnek 7.5 (Monte Carlo Tahmininde Hata Mertebesi). $\hat{\theta}_n = n^{-1} \sum g(U_i)$ tahmininin standart hatası $\sigma_g/\sqrt{n}$ mertebesindedir; burada $\sigma_g^2 = \text{Var}(g(U))$. Hatanın $1/\sqrt{n}$ ile küçülmesi, doğruluğu 10 kat artırmak için örneklem büyüklüğünün 100 kat artırılması gerektiği anlamına gelir. Bu, Monte Carlo yönteminin temel maliyet kuralıdır.

7.5 Merkezi Limit Teoremi

Büyük sayılar yasası ortalamanın nereye yaklaştığını söyler; merkezi limit teoremi ise bu yaklaşımın etrafındaki dalgalanmanın biçimini açıklar. Bu nedenle güven aralıkları ve hipotez testleri için merkezi limit teoremi kritik önemdedir.

Teorem 7.7 (Merkezi Limit Teoremi). $X_1, X_2, \dots$ bağımsız ve aynı dağılımlı, $E(X_i) = \mu$ ve $\text{Var}(X_i) = \sigma^2$ olmak üzere $0 < \sigma^2 < \infty$ olsun. O halde

$$\frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} N(0, 1).$$

Eşdeğer olarak

$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} N(0, 1).$$

Kanıt. İspatın klasik yolu karakteristik fonksiyonları kullanır. X_i değişkenleri standartlaştırılır ve toplamın karakteristik fonksiyonu bağımsızlık nedeniyle çarpım biçiminde yazılır. Taylor açılımı, her terimin küçük katkı yaptığını; çarpımın limitte $\exp(-t^2/2)$ fonksiyonuna yakınsadığını gösterir. Bu fonksiyon standart normal dağılımın karakteristik fonksiyonudur. □

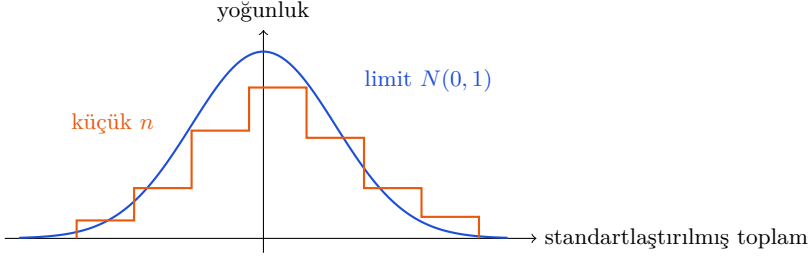
Örnek 7.6 (Ortalama Yanıt Süresi). Bir sunucu isteğinin yanıt süresi sağa çarpık dağılsın; ortalama $\mu = 120$ ms ve standart sapma $\sigma = 60$ ms olsun. $n = 100$ bağımsız isteğin ortalama yanıt süresi için

$$\bar{X}_{100} \approx N\left(120, \frac{60^2}{100}\right)$$

yaklaşımı kullanılır. Buna göre ortalamanın 132 ms'yi aşma olasılığı yaklaşık

$$P\left(Z > \frac{132 - 120}{6}\right) = P(Z > 2)$$

olur.



Şekil 7.2 Merkezi limit teoremi: standartlaştırılmış toplamın dağılımı n büyüdükçe standart normal eğriye yakınsar.

Klasik merkezi limit teoremi aynı dağılımlı değişkenler içindir. Bağımsız ama özdeş olmayan toplamlar için Lindeberg ve Lyapunov koşulları altında aynı sonuç geçerlidir; bu, gerçek veride katkıların eşit olmadığı durumlarda normal yaklaşımı haklı kılar (Billingsley, 1995).

Teorem 7.8 (Lyapunov Merkezi Limit Teoremi). $X_1, X_2, \dots$ bağımsız, $E(X_i) = \mu_i$, $\text{Var}(X_i) = \sigma_i^2$ ve $s_n^2 = \sum_{i=1}^n \sigma_i^2$ olsun. Bir $\delta > 0$ için Lyapunov koşulu

$$\frac{1}{s_n^{2+\delta}} \sum_{i=1}^n E|X_i - \mu_i|^{2+\delta} \rightarrow 0$$

sağlanıyorsa

$$\frac{\sum_{i=1}^n (X_i - \mu_i)}{s_n} \xrightarrow{d} N(0, 1)$$

olur.

Kanıt. Lyapunov koşulu, hiçbir tek terimin toplam değişkenliğe baskın olmadığını güvence altına alır. Bu koşul altında standartlaştırılmış toplamın karakteristik fonksiyonu $\exp(-t^2/2)$ 'ye yakınsar; tekillik gereği limit dağılım standart normaldir. $\square$

Merkezi limit teoremi limitin ne olduğunu söyler ama hatanın büyüklüğünü vermez. Berry–Esseen teoremi, normal yaklaşımın hatasının örneklem büyüklüğüyle nasıl küçüldüğünü açık bir sınırla ifade eder (Shao, 2003).

Teorem 7.9 (Berry–Esseen). $X_1, \dots, X_n$ bağımsız aynı dağılımlı, $E(X_i) = \mu$, $\text{Var}(X_i) = \sigma^2$ ve $\rho = E|X_i - \mu|^3 < \infty$ olsun. F_n standartlaştırılmış toplamın dağılım fonksiyonu, Φ standart normal dağılım fonksiyonu ise mutlak bir C sabiti için

$$\sup_x |F_n(x) - \Phi(x)| \leq \frac{C\rho}{\sigma^3\sqrt{n}}$$

olur.

Kanıt. İspat, F_n ile Φ 'nin karakteristik fonksiyonlarının farkını Esseen düzgünlük eşitsizliğiyle sınırlamaya dayanır. Üçüncü mutlak momentin sonlu olması, fark fonksiyonunun büyüklüğünü $1/\sqrt{n}$ mertebesinde kontrol etmeyi sağlar. $\square$

Yorum:

Berry–Esseen sınırı, normal yaklaşım hatasının $1/\sqrt{n}$ hızında küçüldüğünü söyler. Bu nedenle dağılım çok çarpıksa (büyük ρ/σ^3), aynı doğruluğa ulaşmak için daha büyük örneklem gerekir.

7.6 Normal Yaklaşım ve Süreklilik Düzeltmesi

Merkezi limit teoremi yalnız sürekli değişkenler için değil, toplam biçimindeki kesikli değişkenler için de kullanılır. Ancak kesikli bir dağılım sürekli normal eğriyle yaklaştırılırken sınırlar yarım birim kaydırılır. Buna süreklilik düzeltmesi denir.

Teorem 7.10 (Binom Dağılımı İçin Normal Yaklaşım). $Y \sim \text{Bin}(n, p)$ ve $0 < p < 1$ olsun. n büyük olduğunda

$$\frac{Y - np}{\sqrt{np(1-p)}} \approx N(0, 1)$$

yaklaşımı kullanılır. Aralık olasılıklarında

$$P(a \leq Y \leq b) \approx \Phi\left(\frac{b + 0.5 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a - 0.5 - np}{\sqrt{np(1-p)}}\right)$$

şeklinde süreklilik düzeltmesi yapılır.

Kanıt. Binom değişkeni bağımsız Bernoulli değişkenlerinin toplamıdır. Her Bernoulli değişkeni için ortalama p , varyans $p(1-p)$ olduğundan merkezi limit teoremi doğrudan uygulanır. Süreklilik düzeltmesi, kesikli değerlerin her birini sürekli ekseninde yarım birimlik aralıkla temsil eder. $\square$

Örnek 7.7 (Kusurlu Ürün Sayısı). Bir hatta kusurlu oranı $p = 0.04$ ve örneklem büyüklüğü $n = 500$ olsun. Kusurlu sayısı $Y \sim \text{Bin}(500, 0.04)$ için $np = 20$ ve $np(1-p) = 19.2$ olur. $P(Y \leq 25)$ yaklaşık olarak

$$\Phi\left(\frac{25.5 - 20}{\sqrt{19.2}}\right)$$

ile hesaplanır.

Teorem 7.11 (Poisson Dağılımı İçin Normal Yaklaşım). $Y \sim \text{Poisson}(\lambda)$ olsun. λ büyük olduğunda

$$\frac{Y - \lambda}{\sqrt{\lambda}} \approx N(0, 1)$$

yaklaşımı kullanılabilir. Kesikli sınırlar için süreklilik düzeltmesi önerilir.

Örnek 7.8 (Çağrı Merkezi Yüğü). Bir çağrı merkezine saatte ortalama $\lambda = 100$ çağrı geliyorsa, gelen çağrı sayısı $Y \sim \text{Poisson}(100)$ ile modellenir. $P(Y \leq 110)$ yaklaşık olarak süreklilik düzeltmesiyle

$$\Phi\left(\frac{110.5 - 100}{\sqrt{100}}\right) = \Phi(1.05) \approx 0.853$$

biçiminde hesaplanır. Böylece kapasite planlamasında üst yük tahmini yapılabilir.

7.7 Sürekli Dönüşüm ve Slutsky Teoremi

Asimptotik sonuçlar çoğu zaman doğrudan ilgilendiğimiz istatistik için değil, onun bir parçası için elde edilir. Sürekli dönüşüm ve Slutsky teoremleri, yakınsama sonuçlarını dönüştürülmüş istatistiklere taşır.

Teorem 7.12 (Sürekli Dönüşüm Teoremi). $X_n \xrightarrow{p} X$ ve g sürekli bir fonksiyon ise

$$g(X_n) \xrightarrow{p} g(X)$$

olur. Benzer biçimde $X_n \xrightarrow{d} X$ ise ve g süreklilik koşullarını sağlıyorsa

$$g(X_n) \xrightarrow{d} g(X)$$

sonucu elde edilir.

Kanıt. Süreklilik, X_n değerleri X 'e yakın olduğunda $g(X_n)$ değerlerinin de $g(X)$ 'e yakın olmasını sağlar. Olasılıkta yakınsama için bu fikir doğrudan ε - δ argümanı ile gösterilir. Dağılımda yakınsama için sonuç, dağılım fonksiyonlarının süreklilik noktalarında yakınsamasıyla elde edilir. $\square$

Teorem 7.13 (Slutsky Teoremi). $X_n \xrightarrow{d} X$ ve $Y_n \xrightarrow{p} c$ olsun. Bu durumda

$$X_n + Y_n \xrightarrow{d} X + c, \quad X_n Y_n \xrightarrow{d} cX.$$

Ayrıca $c \neq 0$ ise

$$\frac{X_n}{Y_n} \xrightarrow{d} \frac{X}{c}$$

geçerlidir.

Kanıt. Y_n olasılıkta sabit c 'ye yaklaştığı için büyük örneklemelerde rassal dalgalanması ihmal edilebilir düzeye iner. Ortak yakınsama ve sürekli dönüşüm teoremi toplam, çarpım ve oran fonksiyonlarına uygulandığında sonuçlar elde edilir. $\square$

Örnek 7.9 (Bilinmeyen Varyansla Standartlaştırma). *Merkezi limit teoremi*

$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} N(0, 1)$$

sonucunu verir. Eğer $S_n \xrightarrow{p} \sigma$ ise Slutsky teoremiyle

$$\frac{\bar{X}_n - \mu}{S_n/\sqrt{n}} \xrightarrow{d} N(0, 1)$$

olur. Büyük örneklem güven aralıklarında bilinmeyen σ yerine S_n kullanılmasının nedeni budur.

7.8 Delta Yöntemi

Bazı tahmin ediciler doğrudan ortalama veya oran değil, bunların logaritması, karesi, ters değeri gibi dönüşümleridir. Delta yöntemi, asimptotik normal bir tahmin edicinin düzgün dönüşümünün de yaklaşık normal olduğunu söyler.

Teorem 7.14 (Delta Yöntemi).

$$\sqrt{n}(T_n - \theta) \xrightarrow{d} N(0, \tau^2)$$

olsun. g fonksiyonu θ noktasında türevlenebilir ve $g'(\theta) \neq 0$ ise

$$\sqrt{n}\{g(T_n) - g(\theta)\} \xrightarrow{d} N(0, [g'(\theta)]^2 \tau^2).$$

Kanıt. g fonksiyonu θ çevresinde birinci dereceden Taylor açılımıyla

$$g(T_n) = g(\theta) + g'(\theta)(T_n - \theta) + r_n$$

şeklinde yazılır. Kalan terim r_n , T_n değeri θ 'ya yaklaştığında birinci dereceden terime göre ihmal edilebilir olur. Bu nedenle limit dağılım $g'(\theta)$ katsayısıyla ölçeklenmiş normal dağılımdır. $\square$

Örnek 7.10 (Logaritmik Dönüşüm). T_n pozitif bir parametre θ için

$$\sqrt{n}(T_n - \theta) \xrightarrow{d} N(0, \tau^2)$$

sağlasın. $g(x) = \log x$ için $g'(\theta) = 1/\theta$ olduğundan

$$\sqrt{n}(\log T_n - \log \theta) \xrightarrow{d} N\left(0, \frac{\tau^2}{\theta^2}\right).$$

7.9 Tutarlılık ve Asimptotik Normallik

Asimptotik teori tahmin edicileri değerlendirmek için iki temel kavram sunar: tutarlılık ve asimptotik normallik. Tutarlılık, tahmin edicinin doğru parametreye yaklaşmasını; asimptotik normallik ise yaklaşım hatasının yaklaşık normal biçimde dağılmasını ifade eder.

Tanım 7.7 (Tutarlı Tahmin Edici) T_n bir θ parametresi için tahmin edici olsun. Eğer

$$T_n \xrightarrow{p} \theta$$

ise T_n tutarlı tahmin edicidir.

Tutarlılığı doğrulamanın en pratik yolu, çoğu zaman ortalama karesel hatanın sifıra gitmesini göstermektir. Bu, yanlışlık ve varyansı birlikte ele alan bir ölçüttür (Casella & Berger, 2002; Shao, 2003).

Teorem 7.15 (MSE ile Tutarlılık). T_n tahmin edicisi için ortalama karesel hata

$$\text{MSE}(T_n) = E[(T_n - \theta)^2] = (\text{Yanlışlık}(T_n))^2 + \text{Var}(T_n)$$

biçiminde ayrışır. Eğer $\text{MSE}(T_n) \rightarrow 0$ ise $T_n \xrightarrow{p} \theta$ olur, yani T_n tutarlıdır.

Kanıt. Ayrışım için $E[(T_n - \theta)^2] = E[(T_n - ET_n)^2] + (ET_n - \theta)^2$ yazılır; çapraz terim sıfırdır. $\text{MSE}(T_n) \rightarrow 0$ ise ortalama karede yakınsama geçerlidir; bu da Markov eşitsizliğiyle olasılıkta yakınsamayı verir. $\square$

Örnek 7.11 (Örneklem Ortalamasının Tutarlılığı). $\bar{X}_n$ için yanlışlık sıfır ve $\text{Var}(\bar{X}_n) = \sigma^2/n \rightarrow 0$ olduğundan $\text{MSE}(\bar{X}_n) = \sigma^2/n \rightarrow 0$ olur. Dolayısıyla örneklem ortalaması anakütle ortalamasının tutarlı tahmin edicisidir.

Tanım 7.8 (Asimptotik Normallik) T_n tahmin edicisi için

$$\sqrt{n}(T_n - \theta) \xrightarrow{d} N(0, V)$$

ise T_n asimptotik normaldir. Bu durumda büyük n için

$$T_n \approx N\left(\theta, \frac{V}{n}\right)$$

yaklaşımı kullanılır.

Örnek 7.12 (Örneklem Oranı). $X_i \sim \text{Bernoulli}(p)$ ve $\hat{p}_n = n^{-1} \sum X_i$ olsun. Büyük sayılar yasası $\hat{p}_n \xrightarrow{P} p$ sonucunu verir. Merkezi limit teoremi ise

$$\sqrt{n}(\hat{p}_n - p) \xrightarrow{d} N(0, p(1-p))$$

sonucunu sağlar. Böylece $\hat{p}_n$ hem tutarlı hem de asimptotik normaldir.

7.10 Örnek Veri Tablosu ve Yazılım Çözümleri

Aşağıdaki veri seti, bir sistemde ardışık 15 işlem süresini milisaniye cinsinden temsil etsin.

İşlem	1	2	3	4	5	6	7	8
Süre	118	131	125	144	109	136	121	128
İşlem	9	10	11	12	13	14	15	
Süre	117	152	133	126	140	119	130	

Bu verilerle örneklem ortalaması, örneklem standart sapması ve ortalama için yaklaşık standart hata hesaplanabilir. Büyük örneklem yaklaşımı, n büyüdükçe ortalamanın yaklaşık normal dağılımla modellenmesini sağlar.

```
x <- c(118, 131, 125, 144, 109, 136, 121, 128,
       117, 152, 133, 126, 140, 119, 130)
mean(x)
sd(x)
sd(x) / sqrt(length(x))
```

Python Çözümü

```
import numpy as np

x = np.array([118, 131, 125, 144, 109, 136, 121, 128,
              117, 152, 133, 126, 140, 119, 130])
n = len(x)
ortalama = np.mean(x)
s = np.std(x, ddof=1)
standart_hata = s / np.sqrt(n)

print(n, ortalama, s, standart_hata)
```

7.11 Görselleştirme ve Grafikler

Merkezi limit teoremini görmek için çarpık bir anakütleden çok sayıda örneklem alınabilir. Her örneklem için ortalama hesaplandığında, ortalamaların histogramı tekil gözlemlerin dağılımından daha simetrik ve normal eğriye daha yakın görünür.

```
import numpy as np
import matplotlib.pyplot as plt

rng = np.random.default_rng(123)
ortalamalar = []

for _ in range(10000):
    orneklem = rng.exponential(scale=10, size=50)
    ortalamalar.append(np.mean(orneklem))

plt.hist(ortalamalar, bins=40, density=True, edgecolor="black")
plt.xlabel("Örneklem ortalaması")
plt.ylabel("Yoğunluk")
plt.title("Merkezi Limit Teoremi Simülasyonu")
plt.show()
```

Yorum:

Bu simülasyon, anakütle üstel ve sağa çarpık olsa bile örneklem ortalamalarının daha düzenli bir dağılıma yöneldiğini gösterir. Ancak ağır kuyruklu ve varyansı sonsuz dağılımlarda klasik merkezi limit teoremi doğrudan uygulanmayabilir.

Dikkat: Kavram Yanılgısı

Merkezi limit teoremi, her küçük örneklemde dağılımın normal olacağını söylemez. Teorem, uygun koşullarda ve büyük n için standartlaştırılmış toplamların yaklaşık normal davranacağını belirtir.

Kendini Dene:

Ara Kontrol: Bir simülasyonda örneklem ortalaması n arttıkça dalgalanmayı sürdürüyor ama dalgalanma bandı daralıyorsa bu durum büyük sayılar yasasıyla nasıl uyumludur?

Mini-Proje / Vaka Çalışması:

Rastgele yürüyüş ve büyük sayılar: Bağımsız $X_i \in \{-1, 1\}$ adımları üretilip $S_n = \sum_{i=1}^n X_i$ ve S_n/n eğrilerini aynı grafikte karşılaştırın. S_n tipik olarak büyümeye devam ederken S_n/n değerinin sifıra yaklaşmasını yorumlayın.

7.12 Bölüm Sonu Çözümlü Problemler

1. $X_1, \dots, X_n$ bağımsız ve aynı dağılımlı, $E(X_i) = 50$ ve $\text{Var}(X_i) = 25$ olsun. $\bar{X}_n$ için zayıf büyük sayılar yasasının sonucunu yazın.

Çözüm. Zayıf büyük sayılar yasasına göre

$$\bar{X}_n \xrightarrow{p} 50.$$

Ayrıca Chebyshev sınırı

$$P(|\bar{X}_n - 50| \geq \varepsilon) \leq \frac{25}{n\varepsilon^2}$$

biçimindedir.

2. Aynı koşullarda $n = 100$ için $P(|\bar{X}_n - 50| \geq 5)$ olasılığına Chebyshev üst sınırı bulun.

Çözüm.

$$P(|\bar{X}_{100} - 50| \geq 5) \leq \frac{25}{100 \cdot 5^2} = 0.01.$$

3. X_i değişkenlerinin ortalaması 10, standart sapması 4 olsun. $n = 64$ için $\bar{X}_n$ dağılımını merkezi limit teoremiyle yaklaşık yazın.

Çözüm. Merkezi limit teoremine göre

$$\bar{X}_{64} \approx N\left(10, \frac{4^2}{64}\right) = N(10, 0.25).$$

Bu dağılımın standart sapması 0.5'tir.

4. $Y \sim \text{Bin}(200, 0.30)$ olsun. $P(Y \leq 70)$ olasılığını normal yaklaşım için standartlaştırın.

Çözüm. $np = 60$ ve $np(1 - p) = 42$ olur. Süreklilik düzeltmesiyle

$$P(Y \leq 70) \approx \Phi\left(\frac{70.5 - 60}{\sqrt{42}}\right).$$

5. T_n için $\sqrt{n}(T_n - \theta) \xrightarrow{d} N(0, 9)$ olsun. $g(x) = x^2$ ve $\theta = 2$ için delta yöntemi sonucunu bulun.

Çözüm. $g'(x) = 2x$ olduğundan $g'(2) = 4$ olur. Delta yöntemine göre

$$\sqrt{n}(T_n^2 - 4) \xrightarrow{d} N(0, 4^2 \cdot 9) = N(0, 144).$$

6. $X_n \xrightarrow{d} N(0, 1)$ ve $Y_n \xrightarrow{p} 2$ olsun. X_n/Y_n limit dağılımını bulun.

Çözüm. Slutsky teoremine göre

$$\frac{X_n}{Y_n} \xrightarrow{d} \frac{Z}{2}, \quad Z \sim N(0, 1).$$

Dolayısıyla limit dağılım $N(0, 1/4)$ olur.

7. $X \geq 0$ ve $E[X] = 4$ olsun. Markov eşitsizliğiyle $P(X \geq 10)$ için üst sınır bulun.

Çözüm. $P(X \geq 10) \leq E[X]/10 = 4/10 = 0.4$.

8. $Y_n \sim \text{Bin}(n, p)$ için bağıl frekans $\hat{p}_n = Y_n/n$. $\hat{p}_n$ hangi anlamda p 'ye yakınsar ve neden?

Çözüm. Bernoulli teoremiyle $\hat{p}_n \xrightarrow{p} p$. Çünkü $\text{Var}(\hat{p}_n) = p(1-p)/n \rightarrow 0$ ve Chebyshev sınırı sıfıra gider.

9. Bir Monte Carlo tahmininin standart hatası 0.02 ise, hatayı 0.005'e düşürmek için örneklem kaç katına çıkarılmalı?

Çözüm. Hata $1/\sqrt{n}$ ile ölçeklendiğinden $0.02/0.005 = 4$ kat azalma, $4^2 = 16$ kat örneklem gerektirir.

10. T_n tahmincisi için yanlılık $1/n$ ve varyans $9/n$ olsun. T_n tutarlı mıdır?

Çözüm. $\text{MSE}(T_n) = (1/n)^2 + 9/n \rightarrow 0$ olduğundan T_n tutarlıdır.

11. $\sqrt{n}(T_n - \theta) \xrightarrow{d} N(0, \tau^2)$, $\theta > 0$ ve $g(x) = \sqrt{x}$ olsun. Delta yöntemiyle limit varyansını bulun.

Çözüm. $g'(x) = 1/(2\sqrt{x})$ olduğundan limit varyans $[g'(\theta)]^2 \tau^2 = \tau^2/(4\theta)$ olur.

7.13 Bölüm Sonu Alıştırmaları

- E1. Zayıf büyük sayılar yasasını ifade edin ve pratik anlamını açıklayın.
 E2. Olasılıkta yakınsayan bir dizi örneği verin.
 E3. Merkezi limit teoremini kullanarak standartlaştırılmış bir örneklem ortalamasını yaklaşıkledir.
 E4. Delta yönteminin dönüştürülmüş bir tahmin edicinin dağılımını yaklaşıkledir dırmadaki rolünü açıklayın.

7.14 Kaynak Notu

Bu bölümdeki yakınsama türleri, büyük sayılar yasaları, merkezi limit teoremi, Slutsky teoremi ve delta yöntemi sunumu; Billingsley (1995), Casella ve Berger (2002), Hogg, McKean ve Craig (2019), Ross (2019), Grimmett ve Stirzaker (2020) ve Shao (2003) gibi klasik olasılık ve matematiksel istatistik kaynaklarında yer alan standart çerçeveye uyumludur. Markov ve Chebyshev eşitsizlikleri, Bernoulli teoremi ve Lyapunov merkezi limit teoremi için Feller (1968); MSE ile tutarlılık ve asimptotik tahmin edici çerçevesi için Casella ve Berger (2002) ile Shao (2003) izlenmiştir.

Özet:

Asimptotik teori, örneklem büyüklüğü arttığında rassal değişken dizilerinin nasıl davrandığını inceler. Büyük sayılar yasası örneklem ortalamasının anakütle ortalamasına yaklaştığını, merkezi limit teoremi ise uygun standartlaştırılmış toplamaların normal dağılıma yaklaştığını söyler. Sürekli dönüşüm, Slutsky ve delta yöntemi bu limit sonuçlarını dönüştürülmüş istatistiklere taşır. Tutarlılık doğru parametreye yaklaşmayı, asimptotik normallik ise büyük örneklem hata dağılımını ifade eder.

Bölüm 8

Özel Dağılımlar ve Dağılım Aileleri

8.1 Gerçek Dünya Problemi (Giriş)

Bir çağrı merkezinde dakikada gelen çağrı sayısı, bir üretim hattında kusurlu ürün sayısı, bir cihazın arızalanana kadar çalışma süresi, bir sınav puanının ortalama etrafındaki dalgalanması ve bir ağ bağlantısının ilk kopmaya kadar dayanma süresi birbirinden farklı görünür. Fakat hepsi aynı temel soruya çıkar: Gözlenen rassal büyüklük hangi yapıya sahiptir ve bu yapı hangi olasılık dağılımıyla modellenmelidir?

Dağılım seçimi ezberlenmiş formül listesi değildir. Başarı-başarısızlık deneyleri Bernoulli ve binom modellerine; nadir olay sayıları Poisson modeline; ilk başarıya kadar bekleme geometrik modele; sürekli bekleme süreleri üstel ve gamma ailelerine; toplam etkiler ve ölçüm hataları normal modele doğal biçimde bağlanır. Bu bölümde özel dağılımlar, problem türüne göre seçilen modelleme araçları olarak ele alınır.

Bu Bölüm Ne Sağlar?

Bu bölüm, dağılım ailesi ve parametre fikriyle başlar; ardından kesikli dağılımlar, sürekli dağılımlar, bekleme süresi modelleri, normal aileden türeyen dağılımlar, dağılımlar arası ilişkiler ve dağılım seçimi kontrol listesi sunulur. Amaç, her dağılımı yalnız olasılık fonksiyonu veya yoğunluk formülüyle değil; kullanım bağlamı, parametre yorumu, beklenen değer, varyans ve temel özellikleriyle birlikte öğretmektir.

Amaç ve Öğrenme Çıktıları

Bu bölümün sonunda okuyucu

- dağılım ailesi ve parametre kavramlarını modelleme bağlamında açıklayabilir,
- Bernoulli, binom, geometrik, negatif binom, hipergeometrik ve Poisson dağılımlarını ayırt edebilir,
- düzgün, üstel, gamma, beta ve normal dağılımların temel özelliklerini kullanabilir,
- üstel ve geometrik dağılımların belleksizlik özelliğini yorumlayabilir,
- normal dağılımdan türeyen ki-kare, Student t ve F dağılımlarının nerede kullanıldığını açıklayabilir,
- veri tipine ve süreç yapısına göre uygun dağılım seçimi yapabilir.

8.2 Dağılım Ailesi ve Parametre Mantiğı

Bir dağılım ailesi, aynı matematiksel biçimi paylaşan fakat parametre değerlerine göre değişen dağılımlar kümesidir. Parametreler modelin konumunu, yayılımını, olay hızını veya başarı olasılığını belirler. Bu bakış, özel dağılımları ezber listesi olmaktan çıkarır ve onları modelleme dili haline getirir.

Tanım 8.1 (Dağılım Ailesi) Parametre kümesi Θ olmak üzere, $\theta \in \Theta$ değerlerine bağlı olasılık dağılımları kümesine dağılım ailesi denir. Bu aile kesikli durumda $p(x; \theta)$, sürekli durumda $f(x; \theta)$ ile gösterilebilir.

Örnek 8.1 (Binom Ailesi). $X \sim \text{Bin}(n, p)$ dağılımında n deneme sayısını, p ise her denemedeki başarı olasılığını gösterir. n ve p değiştikçe aynı aile içinde farklı dağılımlar elde edilir.

Aynı ortalama farklı dağılımlardan gelebilir. Örneğin ortalaması 10 olan bir Poisson değişkeni yalnız negatif olmayan tam sayı değerleri alır; ortalaması 10 olan normal değişken ise sürekli ve simetrik. Bu nedenle yalnız ortalamaya bakarak dağılım seçilmez.

8.3 Bernoulli ve Binom Dağılımları

İki sonuçlu deneyler olasılık modellemesinin en temel yapı taşlarından biridir. Başarı–başarısızlık, kusurlu–sağlam, evet–hayır, geçti–kaldı gibi durumlar Bernoulli değişkeniyle temsil edilir. Bağımsız Bernoulli denemelerinin toplamı ise binom dağılımını verir.

Tanım 8.2 (Bernoulli Dağılımı) X yalnız 0 ve 1 değerlerini alıyor ve

$$P(X = 1) = p, \quad P(X = 0) = 1 - p$$

ise $X \sim \text{Bernoulli}(p)$ denir. Bu durumda

$$E(X) = p, \quad \text{Var}(X) = p(1 - p).$$

Tanım 8.3 (Binom Dağılımı) $X_1, \dots, X_n$ bağımsız $\text{Bernoulli}(p)$ değişkenleri ve $S_n = \sum_{i=1}^n X_i$ olsun. Bu durumda $S_n \sim \text{Bin}(n, p)$ ve

$$P(S_n = k) = \binom{n}{k} p^k (1 - p)^{n-k}, \quad k = 0, 1, \dots, n.$$

Ayrıca

$$E(S_n) = np, \quad \text{Var}(S_n) = np(1 - p).$$

Kanıt. $S_n = k$ olayı tam k başarı ve $n - k$ başarısızlık içeren dizilerden oluşur. Her belirli dizinin olasılığı bağımsızlık nedeniyle $p^k(1-p)^{n-k}$ olur. Böyle dizilerin sayısı $\binom{n}{k}$ olduğundan olasılık fonksiyonu elde edilir. Beklenen değer ve varyans, bağımsız Bernoulli değişkenlerinin toplamından gelir. $\square$

Örnek 8.2 (Kalite Kontrol). *Bir üretim hattında bir ürünün kusurlu olma olasılığı 0.03 olsun. Bağımsız seçilen 100 ürünündeki kusurlu sayısı*

$$S \sim \text{Bin}(100, 0.03)$$

ile modellenir. Beklenen kusurlu sayısı $100 \cdot 0.03 = 3$ olur.

Binom dağılımının moment üreten fonksiyonu hem moment hesabını hem de bağımsız binom değişkenlerinin toplamını analiz etmeyi kolaylaştırır (Ross, 2019).

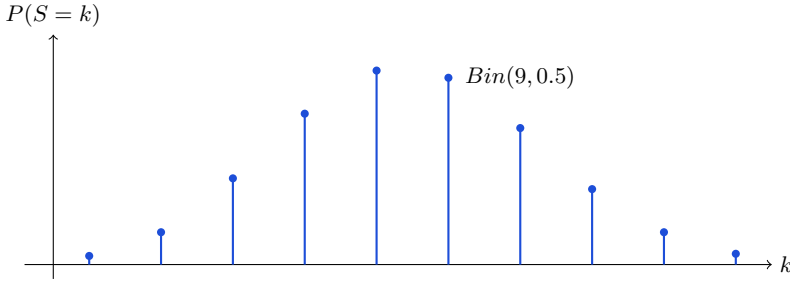
Teorem 8.1 (Binom MGF ve Toplamı). $S_n \sim \text{Bin}(n, p)$ için

$$M_{S_n}(t) = (1 - p + pe^t)^n$$

olur. Ayrıca $X \sim \text{Bin}(m, p)$ ve $Y \sim \text{Bin}(n, p)$ bağımsız ve ortak p ise

$$X + Y \sim \text{Bin}(m + n, p).$$

Kanıt. Bernoulli için $M(t) = 1 - p + pe^t$ 'dir. Binom bağımsız Bernoulli toplamı olduğundan MGF'ler çarpılır ve $(1 - p + pe^t)^n$ elde edilir. Toplam için $M_{X+Y}(t) = (1 - p + pe^t)^m (1 - p + pe^t)^n = (1 - p + pe^t)^{m+n}$ olur; bu $\text{Bin}(m + n, p)$ 'nin MGF'sidir ve tekillik gereği sonuç çıkar. $\square$



Şekil 8.1 Binom olasılık fonksiyonu: $p = 0.5$ için dağılım simetrik; tepe ortalama np civarındadır.

8.4 Geometrik ve Negatif Binom Dağılımları

Binom dağılımında deneme sayısı sabit, başarı sayısı rassaldır. Geometrik ve negatif binom dağılımlarında ise başarı sayısı hedeflenir; bu başarıya ulaşmak için gereken deneme sayısı rassaldır.

Tanım 8.4 (Geometrik Dağılım) Bağımsız Bernoulli denemelerinde ilk başarının gerçekleştiği deneme sayısı T olsun. Bu durumda

$$P(T = k) = (1 - p)^{k-1}p, \quad k = 1, 2, \dots$$

ve $T \sim \text{Geom}(p)$ denir. Bu tanım altında

$$E(T) = \frac{1}{p}, \quad \text{Var}(T) = \frac{1-p}{p^2}.$$

Teorem 8.2 (Geometrik Dağılımın Belleksizlik Özelliği). $T \sim \text{Geom}(p)$ ise pozitif tam sayılar s, t için

$$P(T > s + t \mid T > s) = P(T > t)$$

geçerlidir.

Kanıt. Geometrik dağılım için $P(T > m) = (1 - p)^m$ olur. Bu nedenle

$$P(T > s + t \mid T > s) = \frac{P(T > s + t)}{P(T > s)} = \frac{(1 - p)^{s+t}}{(1 - p)^s} = (1 - p)^t = P(T > t).$$

□

Tanım 8.5 (Negatif Binom Dağılımı) Bağımsız Bernoulli denemelerinde r 'inci başarının gerçekleştiği deneme sayısı T olsun. O halde

$$P(T = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}, \quad k = r, r+1, \dots$$

olur. $r = 1$ özel durumunda geometrik dağılım elde edilir.

Örnek 8.3 (Üçüncü Satışa Kadar Arama). *Bir satış temsilcisinin her aramada satış yapma olasılığı $p = 0.2$ olsun. Üçüncü satışın 10'uncu aramada gerçekleşme olasılığı*

$$\binom{9}{2} (0.2)^3 (0.8)^7$$

ile hesaplanır.

Teorem 8.3 (Negatif Binomun Ortalama ve Varyansı). $T \sim \text{NegBin}(r, p)$, yani r 'inci başarıya kadar deneme sayısı ise

$$E(T) = \frac{r}{p}, \quad \text{Var}(T) = \frac{r(1-p)}{p^2}$$

olur.

Kanıt. T , r bağımsız geometrik bekleme süresinin toplamıdır: $T = \sum_{i=1}^r G_i$, $G_i \sim \text{Geom}(p)$. Her geometrik için $E(G_i) = 1/p$ ve $\text{Var}(G_i) = (1-p)/p^2$ dir. Bağımsızlık nedeniyle beklenenler ve varyanslar toplanır; sonuç elde edilir. $\square$

Yorum:

Geometrik dağılım, kesikli dağılımlar içinde belleksizlik özelliğini taşıyan tek dağılımdır. Sürekli karşılığı üstel dağılımdır. Bu nedenle “geçmiş başarısızlıklar gelecekteki bekleme süresini etkilemez” yorumu yalnız bu iki dağılım için geçerlidir.

8.5 Hipergeometrik Dağılım

Binom dağılımı bağımsız denemelere dayanır. Ancak sonlu bir anakütleden iadesiz örnekleme yapıldığında denemeler bağımsız değildir. Bu durumda hipergeometrik dağılım kullanılır.

Tanım 8.6 (Hipergeometrik Dağılım) N birimlik bir anakütleden K başarı ve $N - K$ başarısızlık bulunsun. İadesiz n birim seçildiğinde örneklemden başarı sayısı X için

$$P(X = x) = \frac{\binom{K}{x} \binom{N-K}{n-x}}{\binom{N}{n}}$$

olur. Uygun x değerleri için bu dağılıma hipergeometrik dağılım denir.

Örnek 8.4 (Partiden Ürün Seçimi). *Bir partide $N = 50$ ürünün $K = 6$ tanesi kusurlu olsun. İadesiz $n = 5$ ürün seçildiğinde tam 1 kusurlu ürün gelme olasılığı*

$$\frac{\binom{6}{1}\binom{44}{4}}{\binom{50}{5}}$$

şeklindedir.

Teorem 8.4 (Hipergeometrik Ortalama ve Binom Yaklaşımı). *Hipergeometrik dağılımda $p = K/N$ olmak üzere*

$$E(X) = np, \quad \text{Var}(X) = np(1-p) \frac{N-n}{N-1}$$

olur. N çok büyük ve n/N küçükse hipergeometrik dağılım $\text{Bin}(n, p)$ ile yaklaşılabılır.

Kanıt. $X = \sum_{i=1}^n I_i$ yazılır; burada I_i , i 'inci çekilişin başarı göstergesidir. Simetri nedeniyle $E(I_i) = K/N = p$ olduğundan $E(X) = np$. Varyansta göstergeler arası negatif kovaryans sonlu anakütle çarpanı $(N-n)/(N-1)$ 'i üretir. $N \rightarrow \infty$ iken bu çarpan 1'e gider ve iadesiz seçim iadeli seçime, yani binoma yaklaşır. $\square$

8.6 Poisson Dağılımı

Poisson dağılımı, sabit bir zaman, alan veya hacim aralığında meydana gelen olay sayısını modellemek için kullanılır. Çağrı sayısı, trafik kazası sayısı, radyoaktif bozunma sayısı veya bir metindeki yazım hatası sayısı bu modele örnek olabilir.

Tanım 8.7 (Poisson Dağılımı) $\lambda > 0$ için X değişkeni

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \dots$$

olasılık fonksiyonuna sahipse $X \sim \text{Poisson}(\lambda)$ denir. Bu durumda

$$E(X) = \lambda, \quad \text{Var}(X) = \lambda.$$

Teorem 8.5 (Binom-Poisson Yaklaşımı). *n büyük, p küçük ve $np = \lambda$ yaklaşık sabit ise*

$$\text{Bin}(n, p) \approx \text{Poisson}(\lambda)$$

yaklaşımı kullanılabilir.

Kanıt. $X \sim \text{Bin}(n, p)$ ve $p = \lambda/n$ alınırsa

$$P(X = k) = \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}.$$

$n \rightarrow \infty$ iken $\binom{n}{k}/n^k \rightarrow 1/k!$ ve $(1 - \lambda/n)^n \rightarrow e^{-\lambda}$ olduğundan Poisson olasılık fonksiyonu elde edilir. $\square$

Örnek 8.5 (Çağrı Merkezi). *Bir çağrı merkezine dakikada ortalama 4 çağrı geliyorsa bir dakikadaki çağrı sayısı yaklaşık $X \sim \text{Poisson}(4)$ ile modellenenebilir. Hiç çağrı gelmeme olasılığı*

$$P(X = 0) = e^{-4}$$

olur.

Bir üretim hattında kusur olasılığı $p = 0.01$ ve günlük üretim $n = 500$ olsun. Kusurlu parça sayısı tam olarak $\text{Bin}(500, 0.01)$ ile modellenir. Ancak n büyük, p küçük ve $np = 5$ olduğundan $\text{Poisson}(5)$ yaklaşımı pratik ve hızlıdır.

Teorem 8.6 (Bağımsız Poisson Toplamı). $X \sim \text{Poisson}(\lambda_1)$ ve $Y \sim \text{Poisson}(\lambda_2)$ bağımsız ise

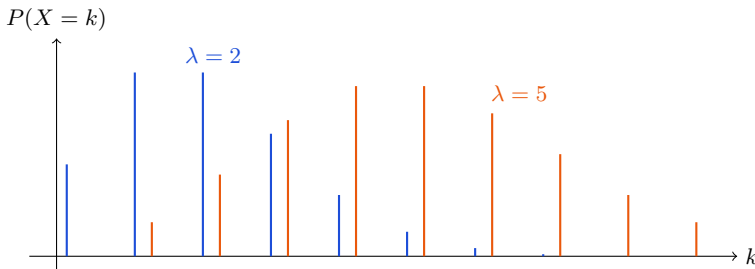
$$X + Y \sim \text{Poisson}(\lambda_1 + \lambda_2)$$

olur.

Kanıt. Poisson'un MGF'si $M(t) = \exp(\lambda(e^t - 1))$ 'dir. Bağımsızlık nedeniyle

$$M_{X+Y}(t) = e^{\lambda_1(e^t - 1)} e^{\lambda_2(e^t - 1)} = e^{(\lambda_1 + \lambda_2)(e^t - 1)}$$

olur; bu $\text{Poisson}(\lambda_1 + \lambda_2)$ 'nin MGF'sidir. Tekillik gereği sonuç çıkar. Sezgisel olarak, iki bağımsız nadir olay akışının birleşimi yine bir Poisson akışıdır. $\square$



Şekil 8.2 Poisson dağılımı: λ büyüdükçe dağılım sağa kayar, yayılır ve daha simetrik görünür.

8.7 Kesikli Düzgün Dağılım

Eş olasılıklı sonlu sonuçlar söz konusu olduğunda kesikli düzgün (uniform) dağılım kullanılır. Zar atma, rastgele numara seçme ve simülasyonlarda başlangıç modelleri bu sınıfa girer.

Tanım 8.8 (Kesikli Düzgün Dağılım) X değişkeni $a, a + 1, \dots, b$ tam sayı değerlerini eş olasılıkla alıyorsa

$$P(X = k) = \frac{1}{b - a + 1}, \quad k = a, a + 1, \dots, b$$

ve X kesikli düzgün dağılımlıdır denir.

Örnek 8.6 (Adil Zar). *Adil bir zar için $X \in \{1, 2, 3, 4, 5, 6\}$ ve $P(X = k) = 1/6$ olur. Eş olasılıklılık varsayımı fiziksel simetriye dayanır; zar hileliyse düzgün model uygun değildir.*

8.8 Sürekli Düzgün Dağılım

Sürekli düzgün (uniform) dağılım, belirli bir aralıkta bütün alt aralıkların uzunlukları oranında olasılık aldığı modeli temsil eder. Simülasyonun temel dağılımıdır; pek çok rassal sayı üretme yöntemi düzgün değişkenden başlar.

Tanım 8.9 (Sürekli Düzgün Dağılım) $X \sim \text{Uniform}(a, b)$ ise yoğunluğu

$$f(x) = \frac{1}{b - a}, \quad a < x < b$$

şeklindedir. Bu durumda

$$E(X) = \frac{a + b}{2}, \quad \text{Var}(X) = \frac{(b - a)^2}{12}.$$

Örnek 8.7 (Rastgele Varış Anı). *Bir otobüsün 10 dakikalık aralıklarla geldiği ve yolcunun durağa geliş anının bu aralık içinde rastgele olduğu varsayılın. Bekleme süresi yaklaşık $\text{Uniform}(0, 10)$ ile modellenenebilir.*

8.9 Üstel ve Gamma Dağılımları

Üstel dağılım sürekli bekleme süresi modellerinin temelidir. Gamma dağılımı ise birden fazla bağımsız üstel bekleme süresinin toplamını modellemek için kullanılır.

Tanım 8.10 (Üstel Dağılım) $\lambda > 0$ için

$$f(x) = \lambda e^{-\lambda x}, \quad x \geq 0$$

yoğunluğuna sahip değişkene üstel dağılımlı denir ve $X \sim \text{Exp}(\lambda)$ yazılır. Bu durumda

$$E(X) = \frac{1}{\lambda}, \quad \text{Var}(X) = \frac{1}{\lambda^2}.$$

Teorem 8.7 (Üstel Dağılımın Belleksizlik Özelliği). $X \sim \text{Exp}(\lambda)$ ise her $s, t \geq 0$ için

$$P(X > s + t \mid X > s) = P(X > t)$$

olur.

Kanıt. Üstel dağılımın sağkalım fonksiyonu $P(X > x) = e^{-\lambda x}$ olduğundan

$$P(X > s + t \mid X > s) = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = P(X > t).$$

□

Tanım 8.11 (Gamma Dağılımı) $\alpha > 0$ ve $\lambda > 0$ için

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, \quad x > 0$$

yoğunluğuna sahip değişkene gamma dağılımlı denir. Bu durumda

$$E(X) = \frac{\alpha}{\lambda}, \quad \text{Var}(X) = \frac{\alpha}{\lambda^2}.$$

Üstel dağılım, $\alpha = 1$ özel durumudur.

Tanım 8.12 (Poisson Süreci) Olayların zaman içinde tek tek gerçekleştiği bir sayım sürecinde, ayrık zaman aralıklarındaki olay sayıları bağımsız ve uzunluğu t olan her aralıktaki olay sayısı $\text{Poisson}(\lambda t)$ dağılımlı ise bu sürece hızı λ olan *Poisson süreci* denir. Bu süreçte ardışık olaylar arası bekleme süreleri bağımsız $\text{Exp}(\lambda)$ dağılımlıdır; ilk α olaya kadar geçen toplam süre ise $\text{Gamma}(\alpha, \lambda)$ dağılımına sahiptir.

Örnek 8.8 (Birden Fazla Olayı Bekleme). *Bir Poisson sürecinde olay hızı saatte $\lambda = 3$ olsun. İlk olay bekleme süresi $\text{Exp}(3)$ ile, üçüncü olaya kadar geçen süre ise şekil parametresi $\alpha = 3$ olan gamma dağılımıyla modellenir.*

Teorem 8.8 (Üstel Toplamı Gamma'dır). $X_1, \dots, X_n \sim \text{Exp}(\lambda)$ bağımsız ise

$$\sum_{i=1}^n X_i \sim \text{Gamma}(n, \lambda)$$

olur.

Kanıt. Üstel dağılımın MGF'si $\lambda/(\lambda - t)$ 'dir ($t < \lambda$). Bağımsızlık nedeniyle toplamın MGF'si bunların çarpımı, yani $(\lambda/(\lambda - t))^n$ olur; bu $Gamma(n, \lambda)$ 'nın MGF'sidir. Tekillik teoremi sonucu verir. Bu, bir Poisson sürecinde n 'inci olaya kadar geçen toplam sürenin gamma dağıldığını açıklar. $\square$

Teorem 8.9 (Bağımsız Üstelerin Minimumu). $X_1, \dots, X_n$ bağımsız ve $X_i \sim Exp(\lambda_i)$ olsun. Bu durumda

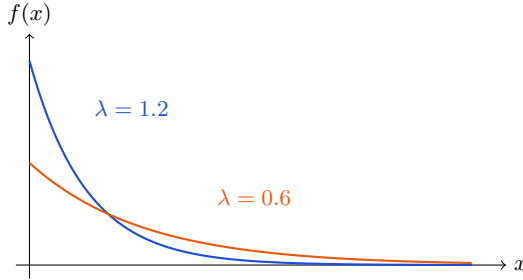
$$\min(X_1, \dots, X_n) \sim Exp\left(\sum_{i=1}^n \lambda_i\right)$$

olur.

Kanıt. Minimumun x 'i aşması tüm değişkenlerin x 'i aşması demektir:

$$P(\min_i X_i > x) = \prod_{i=1}^n P(X_i > x) = \prod_{i=1}^n e^{-\lambda_i x} = e^{-(\sum_i \lambda_i)x}.$$

Bu, hızı $\sum_i \lambda_i$ olan üstel dağılımın sağkalım fonksiyonudur. $\square$



Şekil 8.3 Üstel yoğunluk: hız λ büyüdükçe tepe yükselir ve dağılım orijine doğru sıkışır.

8.10 Beta Dağılımı

Beta dağılımı $[0, 1]$ aralığında tanımlı sürekli değişkenler için esnek bir ailedir. Oranlar, olasılıklar ve belirsiz başarı olasılığı modellerinde sık kullanılır.

Tanım 8.13 (Beta Dağılımı) $\alpha > 0$ ve $\beta > 0$ için

$$f(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}, \quad 0 < x < 1$$

yoğunluğuna sahip değişkene beta dağılımlı denir. Bu durumda

$$E(X) = \frac{\alpha}{\alpha + \beta}.$$

Örnek 8.9 (Başarı Olasılığı Belirsizliği). *Yeni bir ürünün beğenilme oranı p bilinmiyorsa ve p için değerler 0 ile 1 arasında değişiyorsa beta dağılımı doğal bir ön modeldir. α ve β parametreleri önceki başarı ve başarısızlık bilgisini temsil edebilir.*

8.11 Normal Dağılım

Normal dağılım, simetrik ölçüm hataları, toplam etkiler ve merkezi limit yaklaşımı nedeniyle olasılık ve istatistikte merkezi bir yere sahiptir. Tek tek bileşenlerin dağılımı normal olmasa bile, çok sayıda küçük bağımsız etkinin toplamı çoğu zaman yaklaşık normal davranır.

Tanım 8.14 (Normal Dağılım) $\mu \in \mathbb{R}$ ve $\sigma > 0$ için

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left\{ -\frac{(x - \mu)^2}{2\sigma^2} \right\}, \quad -\infty < x < \infty$$

yoğunluğuna sahip değişkene normal dağılımlı denir ve $X \sim N(\mu, \sigma^2)$ yazılır.

Teorem 8.10 (Standartlaştırma). $X \sim N(\mu, \sigma^2)$ ise

$$Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$$

olur.

Kanıt. Doğrusal dönüşüm yoğunluk dönüşümüyle uygulanır. $X = \mu + \sigma Z$ yazıldığında Jacobian çarpanı σ olur ve elde edilen yoğunluk standart normal yoğunluğudur. $\square$

Örnek 8.10 (Sınav Puanı). *Bir sınav puanı yaklaşık $N(70, 10^2)$ dağılımına sahipse 85'in üzerindeki puan olasılığı*

$$P(X > 85) = P\left(Z > \frac{85 - 70}{10}\right) = P(Z > 1.5)$$

şeklinde standart normal tabloyla bulunur.

Teorem 8.11 (Bağımsız Normallerin Toplamı). $X \sim N(\mu_1, \sigma_1^2)$ ve $Y \sim N(\mu_2, \sigma_2^2)$ bağımsız ise

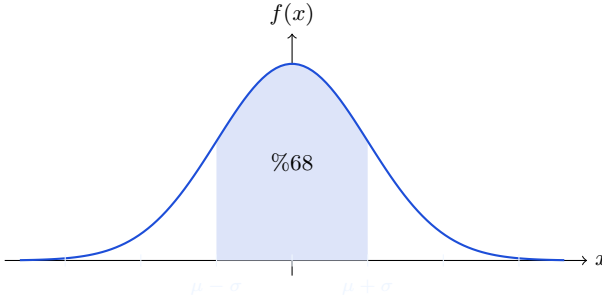
$$X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

olur.

Kanıt. Normal dağılımın MGF'si $\exp(\mu t + \sigma^2 t^2/2)$ 'dir. Bağımsız toplamda MGF'ler çarpılır; üsteller toplanarak $\exp((\mu_1 + \mu_2)t + (\sigma_1^2 + \sigma_2^2)t^2/2)$ elde edilir. Bu, belirtilen parametrelili normalin MGF'sidir. $\square$

Yorum:

Ampirik kural (68–95–99.7): Normal dağılımda gözlemlerin yaklaşık %68'i $\mu \pm \sigma$, %95'i $\mu \pm 2\sigma$ ve %99.7'si $\mu \pm 3\sigma$ aralığındadır. Bu kural, normal verinin hızlı yorumlanmasında ve aykırı değer taramasında pratik bir ölçüttür.



Şekil 8.4 Normal yoğunluk ve ampirik kural: $\mu \pm \sigma$ aralığı toplam alanın yaklaşık %68'ini kapsar.

8.12 Normal Dağılımdan Türeyen Dağılımlar

Ki-kare, Student t ve F dağılımları normal örneklem kuramından doğar. Bu dağılımlar özellikle varyans tahmini, küçük örneklem ortalama çıkarımı ve varyans oranı problemlerinde kullanılır.

Tanım 8.15 (Ki-Kare Dağılımı) $Z_1, \dots, Z_k$ bağımsız standart normal değişkenler ise

$$X = Z_1^2 + \dots + Z_k^2$$

değişkeni k serbestlik dereceli ki-kare dağılımına sahiptir ve $X \sim \chi_k^2$ yazılır.

Tanım 8.16 (Student t Dağılımı) $Z \sim N(0, 1)$, $U \sim \chi_\nu^2$ ve Z ile U bağımsız olsun. Bu durumda

$$T = \frac{Z}{\sqrt{U/\nu}}$$

değişkeni ν serbestlik dereceli Student t dağılımına sahiptir.

Tanım 8.17 (F Dağılımı) $U \sim \chi_{\nu_1}^2$ ve $V \sim \chi_{\nu_2}^2$ bağımsız ise

$$F = \frac{U/\nu_1}{V/\nu_2}$$

değişkeni (ν_1, ν_2) serbestlik dereceli F dağılımına sahiptir.

Örnek 8.11 (Küçük Örneklem Ortalaması). *Normal anakütleden gelen küçük örnekte σ bilinmiyorsa*

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

istatistiği t_{n-1} dağılımına sahiptir. Bu nedenle küçük örneklem güven aralıklarında normal tablo yerine t tablosu kullanılır.

Teorem 8.12 (Ki-Kare Gamma'nın Özel Durumudur). $X \sim \chi_k^2$ ise $X \sim \text{Gamma}(\alpha = k/2, \lambda = 1/2)$ olur. Dolayısıyla

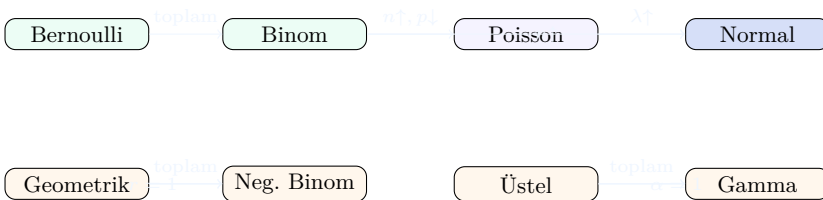
$$E(X) = k, \quad \text{Var}(X) = 2k.$$

Kanıt. Ki-kare yoğunluğu $f(x) \propto x^{k/2-1}e^{-x/2}$ biçimindedir; bu, şekil $k/2$ ve hız $1/2$ olan gamma yoğunluğudur. Gamma ortalaması $\alpha/\lambda = (k/2)/(1/2) = k$ ve varyansı $\alpha/\lambda^2 = (k/2)/(1/4) = 2k$ olur. $\square$

8.13 Dağılımlar Arası İlişkiler

Özel dağılımlar birbirinden kopuk değildir. Bir dağılım başka bir dağılımın özel durumu, limiti veya toplamı olarak ortaya çıkabilir. Bu ilişkiler model seçimini ve yaklaşık hesaplamayı kolaylaştırır.

- Bernoulli değişkenlerinin toplamı binom dağılımıdır.
- Geometrik dağılım negatif binom dağılımının $r = 1$ özel durumudur.
- Üstel dağılım gamma dağılımının $\alpha = 1$ özel durumudur.
- Bağımsız üstel bekleme sürelerinin toplamı gamma dağılımıdır.
- n büyük ve p küçük olduğunda binom dağılımı Poisson ile yaklaştırılabilir.
- n büyük olduğunda binom ve Poisson dağılımları uygun koşullarda normal ile yaklaştırılabilir.



Şekil 8.5 Özel dağılımlar arası ilişkiler: toplama, limit ve özel durum bağlantıları dağılımları tek bir harita üzerinde birleştirir.

8.14 Dağılım Seçim Kontrol Listesi

Bir modelleme probleminde dağılım seçerken aşağıdaki sorular yol göstericidir:

1. Değişken kesikli mi, sürekli mi?
2. Değerler sınırlı mı, sınırsız mı?
3. Değişken bir sayım mı, bekleme süresi mi, oran mı, ölçüm hatası mı?
4. Denemeler bağımsız mı, yoksa iadesiz seçim nedeniyle bağımlı mı?
5. Ortalama ile varyans arasında özel bir ilişki var mı?
6. Süreçte bellek var mı, yoksa kalan bekleme süresi geçmişten bağımsız mı?
7. Büyük örneklem veya nadir olay yaklaşımı kullanılabilir mi?

Yaygın Hata Kalıbı:

En yaygın hata, veri tipini dikkate almadan normal dağılım kullanmaktır. Negatif değer alamayan bekleme süreleri, sınırlı oranlar veya küçük sayım verileri çoğu zaman normal dağılımdan farklı modeller gerektirir.

8.15 Örnek Veri Tablosu ve Yazılım Çözümleri

Aşağıdaki tablo, bir üretim hattında 10 gün boyunca gözlenen kusurlu ürün sayılarını göstermektedir.

Gün	1	2	3	4	5	6	7	8	9	10
Kusurlu Sayısı	3	4	2	5	1	4	3	6	2	4

Bu veri sayım verisidir ve negatif değer alamaz. Günlük ortalama kusurlu sayısı yaklaşık Poisson parametresi için ilk tahmin olarak kullanılabilir.

```
x <- c(3, 4, 2, 5, 1, 4, 3, 6, 2, 4)
lambda_hat <- mean(x)
lambda_hat
ppois(2, lambda=lambda_hat)
```

Python Çözümü

```
import numpy as np
from math import exp, factorial

x = np.array([3, 4, 2, 5, 1, 4, 3, 6, 2, 4])
lambda_hat = np.mean(x)
prob_le_2 = sum(exp(-lambda_hat) * lambda_hat**k / factorial(k)
                 for k in range(3))

print(lambda_hat, prob_le_2)
```

8.16 Görselleştirme ve Grafikler

Farklı parametre değerleri dağılımın şeklini değiştirir. Örneğin Poisson dağılımında λ arttıkça dağılım sağa kayar ve daha simetrik görünmeye başlar.

```
import numpy as np
import matplotlib.pyplot as plt
from math import exp, factorial

k = np.arange(0, 16)
for lam in [1, 4, 9]:
    p = [exp(-lam) * lam**i / factorial(i) for i in k]
    plt.plot(k, p, marker="o", label=f"lambda={lam}")

plt.xlabel("k")
plt.ylabel("P(X=k)")
plt.title("Poisson Dağılımında Parametre Etkisi")
plt.legend()
plt.show()
```

Dikkat: Kavram Yanılgısı

Normal dağılım her sürekli veri için varsayılan model değildir. Pozitif, sağa çarpık ve uç değer üretebilen verilerde log-normal, gamma, Pareto veya Cauchy gibi ağır kuyruklu aileler daha uygun olabilir.

Ufuk Notu:

Ağır kuyruklar ve risk yönetimi: Finansal kayıp, sigorta hasarı ve internet trafiği gibi alanlarda uç değerler ortalamayı ve riski belirgin biçimde etkiler. Pareto dağılımı “az sayıda büyük olay” yapısını; log-normal dağılım ise çarpımsal büyüme süreçlerini modellemek için doğal adaydır. Cauchy dağılımında ortalamanın tanımsız olması, uç değerlerin klasik özetleri nasıl kırılganlaştırdığını gösterir.

Kendini Dene:

Ara Kontrol: Bir veri kümesinde birkaç çok büyük gözlem ortalamayı dramatik biçimde değiştiriyorsa model seçerken hangi kuyruk özelliğine bakarsınız?

8.17 Bölüm Sonu Çözümlü Problemler

1. Bir ürünün kusurlu olma olasılığı 0.02 olsun. Bağımsız 50 üründe tam 2 kusurlu çıkma olasılığını yazın.

Çözüm. Kusurlu sayısı $X \sim \text{Bin}(50, 0.02)$ olur. Bu nedenle

$$P(X = 2) = \binom{50}{2} (0.02)^2 (0.98)^{48}.$$

2. Bir çağrı merkezine dakikada ortalama 3 çağrı gelmektedir. Bir dakikada hiç çağrı gelmeme olasılığı nedir?

Çözüm. $X \sim \text{Poisson}(3)$ alınırsa

$$P(X = 0) = e^{-3} \frac{3^0}{0!} = e^{-3}.$$

3. Başarı olasılığı $p = 0.25$ olan bağımsız denemelerde ilk başarının dördüncü denemede gelme olasılığını bulun.

Çözüm. $T \sim \text{Geom}(0.25)$ olduğundan

$$P(T = 4) = (0.75)^3 (0.25).$$

4. $X \sim \text{Exp}(2)$ olsun. $P(X > 1)$ değerini hesaplayın.

Çözüm. Üstel dağılımın sağkalım fonksiyonu $P(X > x) = e^{-\lambda x}$ olduğundan

$$P(X > 1) = e^{-2}.$$

5. $X \sim N(100, 15^2)$ olsun. X 'in standartlaştırılmış değişkenini yazın.

Çözüm.

$$Z = \frac{X - 100}{15} \sim N(0, 1).$$

6. $Z_1, \dots, Z_5$ bağımsız standart normal değişkenler olsun. $Z_1^2 + \dots + Z_5^2$ dağılımı nedir?

Çözüm. Tanım gereği

$$Z_1^2 + \dots + Z_5^2 \sim \chi_5^2.$$

7. $X \sim \text{Poisson}(2)$ ve $Y \sim \text{Poisson}(3)$ bağımsız ise $X + Y$ dağılımı nedir?

Çözüm. Bağımsız Poisson toplamı gereği $X + Y \sim \text{Poisson}(5)$.

8. $X_1, X_2, X_3 \sim \text{Exp}(\lambda)$ bağımsız ise toplamlarının dağılımı nedir?

Çözüm. Üstel toplamı gamma'dır: $X_1 + X_2 + X_3 \sim \text{Gamma}(3, \lambda)$.

9. İki bağımsız sistem üstel ömürlü, hızları $\lambda_1 = 0.1$ ve $\lambda_2 = 0.2$ olsun. İlk arızaya kadar geçen sürenin dağılımı nedir?

Çözüm. Minimum üsteldir: $\min(X_1, X_2) \sim \text{Exp}(0.1 + 0.2) = \text{Exp}(0.3)$.

10. r 'inci başarıya kadar deneme sayısı $T \sim \text{NegBin}(r, p)$ ise $E(T)$ nedir?

Çözüm. r bağımsız geometriğin toplamı olduğundan $E(T) = r/p$.

11. $X \sim \chi_8^2$ için beklenen değer ve varyans nedir?

Çözüm. Ki-kare gamma'nın özel durumu olduğundan $E(X) = 8$ ve $\text{Var}(X) = 2 \cdot 8 = 16$.

8.18 Bölüm Sonu Alıştırımları

- E1. Binom dağılımıyla modellenebilecek gerçek bir durum belirleyin ve varsayımları yazın.
 E2. Geometrik ve negatif binom dağılımlarını bekleme zamanı açısından karşılaştırın.
 E3. Üstel dağılımın neden hafızasız olduğunu açıklayın.
 E4. Normal dağılımın uygulamalarda sık görülmesinin bir nedenini belirtin.

8.19 Kaynak Notu

Bu bölümdeki özel dağılımlar, dağılım aileleri, bekleme süresi modelleri, Poisson yaklaşımı ve normal aileden türeyen dağılımlar; Ross (2019), Hogg, McKean ve Craig (2019), Casella ve Berger (2002), Montgomery ve Runger (2018), Grimmer ve Stirzaker (2020) ve Hsu (1997) gibi klasik olasılık ve istatistik kaynaklarında yer alan standart çerçeveye uyumludur. Moment üreten fonksiyonlarla toplam dağılımları, üstel ve gamma ilişkileri, hipergeometrik ile binom bağlantısı ve dağılımlar arası ilişki haritası için Feller (1968), Casella ve Berger (2002) ve Mukhopadhyay (2000) izlenmiştir.

Özet:

Dağılım seçimi, problemin ölçüm tipine ve süreç yapısına bağlıdır. İki sonuçlu deneyler Bernoulli ve binom; ilk başarıya kadar bekleme geometrik; nadir olay sayıları Poisson; sürekli bekleme süreleri üstel ve gamma; oranlar beta; simetrik ölçüm hataları ve toplam etkiler normal dağılımla modellenir. Ki-kare, t ve F dağılımları normal örneklem kuramından türeyerek çıkarımsal istatistikte temel rol oynar.

Bölüm 9

Koşullu Dağılımlar ve Beklenti

9.1 Gerçek Dünya Problemi (Giriş)

Bir bankanın kredi riski modeli, müşterinin gelir düzeyi, ödeme geçmişi ve mevcut borç oranı bilindiğinde temerrüt olasılığını yeniden hesaplar. Bir hastanede tanı testi pozitif çıkan hastanın gerçek hastalık olasılığı, hastalığın toplumdaki yaygınlığına ve testin hata oranlarına bağlıdır. Bir sigorta şirketinde beklenen hasar, müşterinin yaş grubu, araç tipi ve geçmiş kaza bilgisi öğrenildiğinde değişir.

Bu örneklerin ortak noktası yeni bilginin belirsizliği değiştirmesidir. Koşullu dağılımlar ve koşullu beklenti, “bilgi geldikten sonra olasılık modeli nasıl güncellenir?” sorusuna sistematik yanıt verir. Bu bölüm, olaylar için koşullu olasılık fikrini rassal değişkenlere ve beklentiye taşır.

Bu Bölüm Ne Sağlar?

Bu bölüm koşullu olasılıktan başlayarak koşullu PMF, koşullu yoğunluk, koşullu dağılım fonksiyonu, koşullu beklenti, koşullu varyans, toplam beklenti yasası, toplam varyans yasası, Bayes kuralı, koşullama ile parçalama ve koşullu beklentinin en iyi karesel tahmin yorumu konularını ele alır. Amaç, koşullamayı yalnız formül değil; bilgi güncelleme, tahmin ve karmaşık modelleri alt gruplara ayırma yöntemi olarak kurmaktır.

Amaç ve Öğrenme Çıktıları

Bu bölümün sonunda okuyucu

- koşullu olasılık, koşullu PMF ve koşullu yoğunluk fonksiyonlarını yazabilir,
- koşullu dağılımdan koşullu beklenen değer ve koşullu varyans hesaplayabilir,
- toplam beklenti ve toplam varyans yasalarını uygulayabilir,
- Bayes kuralını dağılım ve hipotez güncelleme problemlerinde kullanabilir,
- koşullu beklentiye bir rassal değişken ve en iyi karesel tahmin olarak yorumlayabilir,
- koşullama ile karmaşık olasılık ve beklenti problemlerini daha küçük parçalara ayırabilir.

9.2 Koşullama Fikrinin Temeli

Koşullu olasılık, örnek uzayın yeni bilgiyle daraltılmasıdır. B olayının gerçekleştiği biliniyorsa artık olasılıklar tüm Ω üzerinde değil, B üzerinde yeniden ölçeklenerek hesaplanır. Bu fikir, Bayes düşüncesinin ve koşullu dağılımların temelidir.

Tanım 9.1 (Koşullu Olasılık) $P(B) > 0$ ise A olayının B koşulu altındaki olasılığı

$$P(A | B) = \frac{P(A \cap B)}{P(B)}$$

olarak tanımlanır.

Teorem 9.1 (Çarpım Kuralı). $P(B) > 0$ ise

$$P(A \cap B) = P(B)P(A | B)$$

olur. Uygun pozitif olasılıklar altında genel biçim

$$P(A_1 \cap \cdots \cap A_n) = P(A_1)P(A_2 | A_1) \cdots P(A_n | A_1 \cap \cdots \cap A_{n-1})$$

şeklinde dir.

Kanıt. İlk eşitlik koşullu olasılık tanımının yeniden düzenlenmesidir. Genel eşitlik aynı düzenlemenin ardışık kesişimlere uygulanmasıyla elde edilir. $\square$

Örnek 9.1 (İadesiz Çekim). Bir torbada 5 kırmızı ve 3 mavi top olsun. İadesiz iki top çekildiğinde ikisinin de kırmızı olma olasılığı

$$P(K_1 \cap K_2) = P(K_1)P(K_2 | K_1) = \frac{5}{8} \cdot \frac{4}{7}$$

olur.

Koşullu olasılıkta koşul değiştiğinde olasılık uzayı değişir. $P(A | B)$ ile $P(B | A)$ genellikle aynı değildir. Bu karışıklık tıbbi test, kredi riski ve sınıflandırma problemlerinde en yaygın yorum hatalarından biridir.

Koşullamanın en güçlü kullanımlarından biri, karmaşık bir olayı ayrık senaryolara bölüp her senaryonun katkısını toplamaktır. Bu fikir toplam olasılık yasasıyla ifade edilir (Ross, 2019; Bertsekas & Tsitsiklis, 2008).

Teorem 9.2 (Toplam Olasılık Yasası). $A_1, \dots, A_n$ örnek uzayın bir ayrışımı ($P(A_i) > 0$) ve B herhangi bir olay olsun. Bu durumda

$$P(B) = \sum_{i=1}^n P(B | A_i)P(A_i)$$

olur.

Kanıt. A_i olayları ayrık ve birleşimleri tüm örnek uzay olduğundan $B = \bigcup_i (B \cap A_i)$ ayrık birleşimdir. Olasılıkların toplanabilirliğiyle $P(B) = \sum_i P(B \cap A_i)$ olur. Çarpım kuralıyla her terim $P(B | A_i)P(A_i)$ 'ye eşittir. $\square$

Örnek 9.2 (İki Makineli Üretim). *Bir fabrikada parçaların %60'ı birinci makineden (%2 kusur), %40'ı ikinci makineden (%5 kusur) geliyor. Rastgele bir parçanın kusurlu olma olasılığı*

$$P(\text{kusur}) = 0.02 \cdot 0.60 + 0.05 \cdot 0.40 = 0.032$$

olur. Senaryolara bölüp ağırlıklı toplama klasik bir koşullama uygulamasıdır.

9.3 Kesikli Koşullu Dağılımlar

Kesikli rassal değişkenlerde koşullu dağılım, ortak olasılık fonksiyonunun ilgili marjinal olasılığa bölünmesiyle elde edilir. Bu işlem, $Y = y$ bilgisi altında X değişkeninin olasılıklarını yeniden ölçekler.

Tanım 9.2 (Koşullu PMF) Kesikli X ve Y rassal değişkenleri için $P(Y = y) > 0$ olduğunda

$$P(X = x | Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)}$$

ifadesi X 'in $Y = y$ koşulu altındaki koşullu olasılık kütle fonksiyonudur.

Teorem 9.3 (Koşullu PMF'nin Normalize Olması). $P(Y = y) > 0$ ise

$$\sum_x P(X = x \mid Y = y) = 1$$

olur.

Kanıt. Tanımdan

$$\sum_x P(X = x \mid Y = y) = \frac{1}{P(Y = y)} \sum_x P(X = x, Y = y) = \frac{P(Y = y)}{P(Y = y)} = 1.$$

□

Örnek 9.3 (Müşteri Sınıfı ve Temerrüt). Y müşteri sınıfını, X temerrüt durumunu gösterebilir. $P(X = 1, Y = y)$ ortak olasılığı ve $P(Y = y)$ sınıf olasılığı biliniyorsa

$$P(X = 1 \mid Y = y) = \frac{P(X = 1, Y = y)}{P(Y = y)}$$

o sınıftaki temerrüt oranını verir.

Örnek 9.4 (Ortak Tablodan Koşullu Dağılım). İki zar atımında X birinci zar, $Y = X_1 + X_2$ toplam olsun. $Y = 4$ verildiğinde X 'in koşullu dağılımı aranıyor. $Y = 4$ veren ortak sonuçlar $(1, 3), (2, 2), (3, 1)$ olup her biri $1/36$ olasılıktadır; dolayısıyla $P(Y = 4) = 3/36$. Koşullu olasılıklar

$$P(X = 1 \mid Y = 4) = P(X = 2 \mid Y = 4) = P(X = 3 \mid Y = 4) = \frac{1/36}{3/36} = \frac{1}{3}$$

olur. Toplamın bilinmesi, birinci zarı üç eş olasılıklı değere indirger.

9.4 Sürekli Koşullu Yoğunluk

Sürekli rassal değişkenlerde tek bir noktaya koşullamanın olasılığı sıfırdır; buna rağmen yoğunluklar aracılığıyla $Y = y$ bilgisi altında X 'in koşullu yoğunluğu tanımlanabilir.

Tanım 9.3 (Koşullu Yoğunluk) (X, Y) ortak yoğunluğu $f_{X,Y}$ ve Y marjinal yoğunluğu f_Y olsun. $f_Y(y) > 0$ ise

$$f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$$

X 'in $Y = y$ koşulu altındaki koşullu yoğunluğudur.

Teorem 9.4 (Koşullu Yoğunluğun Normalize Olması). $f_Y(y) > 0$ ise

$$\int_{-\infty}^{\infty} f_{X|Y}(x | y) dx = 1$$

olur.

Kanıt. Koşullu yoğunluk tanımını kullanılır:

$$\int f_{X|Y}(x | y) dx = \frac{1}{f_Y(y)} \int f_{X,Y}(x, y) dx = \frac{f_Y(y)}{f_Y(y)} = 1.$$

□

Tanım 9.4 (Koşullu Dağılım Fonksiyonu) $Y = y$ bilgisi altında X 'in koşullu dağılım fonksiyonu

$$F_{X|Y}(x | y) = P(X \leq x | Y = y)$$

şeklinde tanımlanır. Sürekli durumda

$$F_{X|Y}(x | y) = \int_{-\infty}^x f_{X|Y}(u | y) du$$

olur.

Örnek 9.5 (Sıcaklığa Göre Enerji Tüketimi). X günlük enerji tüketimi, Y dış sıcaklık olsun. $f_{X|Y}(x | 5)$, dış sıcaklık 5 derece olduğunda enerji tüketiminin dağılımını gösterir. Bu koşullu dağılım, genel tüketim dağılımından daha dar veya farklı konumda olabilir.

Örnek 9.6 (Üçgen Bölgede Koşullu Dağılım). (X, Y) çifti $0 < x < y < 1$ üçgeninde düzgün dağılsın; ortak yoğunluk $f_{X,Y}(x, y) = 2$. Y 'nin marjinal yoğunluğu $f_Y(y) = \int_0^y 2 dx = 2y$ olur. Buradan

$$f_{X|Y}(x | y) = \frac{2}{2y} = \frac{1}{y}, \quad 0 < x < y,$$

yani $Y = y$ verildiğinde $X \sim U(0, y)$ olur. Dolayısıyla $E[X | Y = y] = y/2$ bulunur; bu da $E[X | Y] = Y/2$ rassal değişkenini verir.

9.5 Bağımsızlık ve Koşullama

Koşullama yeni bilgi ekler. Eğer bu bilgi dağılımı değiştirmiyorsa, ilgili değişkenler bağımsızdır. Bu nedenle bağımsızlık koşullu dağılım dilinde çok doğal biçimde ifade edilir.

Teorem 9.5 (Bağımsızlık ve Koşullu Dağılım). X ve Y bağımsız ise $P(Y = y) > 0$ olan kesikli durumda

$$P(X = x \mid Y = y) = P(X = x)$$

olur. Sürekli durumda ise $f_Y(y) > 0$ için

$$f_{X|Y}(x \mid y) = f_X(x)$$

geçerlidir.

Kanıt. Kesikli durumda bağımsızlık $P(X = x, Y = y) = P(X = x)P(Y = y)$ verir. Tanıma yerleştirildiğinde $P(X = x \mid Y = y) = P(X = x)$ bulunur. Sürekli durumda aynı argüman ortak yoğunluğun $f_X(x)f_Y(y)$ biçiminde çarpanlara ayrılmasıyla yapılır. $\square$

Örnek 9.7 (Bilgi Değer Taşımadığında). *Bir zarın sonucu X , aynı anda atılan bağımsız bir paranın sonucu Y olsun. Paranın yazı geldiğini bilmek zarın dağılımını değiştirmez; bu nedenle $P(X = k \mid Y = yazı) = 1/6$ olur.*

9.6 Koşullu Beklenti

Koşullu beklenti, yeni bilgi geldikten sonra rassal değişkenin ortalamasıdır. $E[X \mid Y = y]$ sayısal bir değerdir; $E[X \mid Y]$ ise Y 'nin değerine göre değişen bir rassal değişkendir.

Tanım 9.5 (Koşullu Beklenti) Kesikli durumda

$$E[X \mid Y = y] = \sum_x xP(X = x \mid Y = y),$$

sürekli durumda ise

$$E[X \mid Y = y] = \int_{-\infty}^{\infty} x f_{X|Y}(x \mid y) dx$$

olarak tanımlanır.

Tanım 9.6 (Koşullu Beklenti Bir Rassal Değişkendir) $E[X \mid Y = y]$ değeri y 'nin bir fonksiyonudur. Bu fonksiyon $m(y)$ ile gösterilirse

$$E[X \mid Y] = m(Y)$$

bir rassal değişkendir.

Teorem 9.6 (Koşullu Beklentinin Doğrusallığı). *Beklentiler tanımlıysa*

$$E[aX + bZ \mid Y] = aE[X \mid Y] + bE[Z \mid Y]$$

geçerlidir.

Kant. Her $Y = y$ koşulu altında beklenti, koşullu dağılıma göre alınan sıradan beklentidir. Beklenti operatörünün doğrusallığı koşullu dağılım altında da geçerlidir. $\square$

Örnek 9.8 (Gelire Göre Harcama). X aylık harcama, Y gelir grubu olsun. $E[X | Y = \text{yüksek}]$, yüksek gelir grubundaki ortalama harcamayı verir. $E[X | Y]$ ise her müşteri için gelir grubuna bağlı tahmin değeridir.

Koşullu beklentinin en kullanışlı özelliklerinden biri, koşullanan değişkenin fonksiyonlarının sabit gibi dışarı çıkarılabilmesidir. Bu kural çok sayıda hesaba kısaltır (Williams, 1991; Grimmett & Stirzaker, 2020).

Teorem 9.7 (Bilineni Dışarı Alma). $g(Y)$ uygun bir fonksiyon ve beklentiler tanımlı olsun. Bu durumda

$$E[g(Y)X | Y] = g(Y) E[X | Y]$$

olur. Özel olarak X ile Y bağımsızsa $E[X | Y] = E[X]$ sabittir.

Kant. $Y = y$ koşulu altında $g(Y) = g(y)$ sabittir; koşullu beklenti sıradan beklenti gibi davrandığından sabit dışarı çıkar: $E[g(y)X | Y = y] = g(y)E[X | Y = y]$. Bu her y için geçerli olduğundan rassal değişken eşitliği elde edilir. Bağımsızlık durumunda koşullu dağılım marjinal dağılıma eşit olduğundan $E[X | Y] = E[X]$ olur. $\square$

Örnek 9.9 (Çarpımın Koşullu Beklentisi). $E[Y^2 | Y] = Y^2$ olur; çünkü Y koşulu altında Y bilinir. Benzer biçimde $E[YX | Y] = Y E[X | Y]$ yazılır. Bu, regresyon ve filtreleme hesaplarında sık kullanılan bir sadeleştirmedir.

9.7 Toplam Beklenti Yasası

Toplam beklenti yasası, genel ortalamanın koşullu ortalamaların ağırlıklı ortalaması olduğunu söyler. Karmaşık modellerde önce alt gruplara göre ortalama almak, sonra bu ortalamaları grup olasılıklarıyla birleştirmek çoğu zaman en kolay yoldur.

Teorem 9.8 (Toplam Beklenti Yasası). $E|X| < \infty$ ise

$$E[X] = E\{E[X | Y]\}$$

geçerlidir.

Kant. Kesikli durumda

$$E\{E[X | Y]\} = \sum_y E[X | Y = y]P(Y = y) = \sum_y \sum_x xP(X = x | Y = y)P(Y = y).$$

Çarpım kuralı ile $P(X = x \mid Y = y)P(Y = y) = P(X = x, Y = y)$ olur. Böylece

$$\sum_y \sum_x xP(X = x, Y = y) = \sum_x xP(X = x) = E[X]$$

elde edilir. Sürekli durumda toplamların yerini integraller alır. $\square$

Örnek 9.10 (Sigorta Hasarı). *Bir portföyde müşteri tipi Y olsun. Genç sürücüler ve deneyimli sürücüler farklı hasar ortalamalarına sahipse genel beklenen hasar*

$$E[X] = \sum_y E[X \mid Y = y]P(Y = y)$$

şeklinde tip bazlı ortalamaların ağırlıklı ortalamasıdır.

9.8 Koşullu Varyans ve Toplam Varyans Yasası

Koşullu beklenti ortalamasının bilgiyle nasıl değiştiğini gösterir. Koşullu varyans ise bilgi geldikten sonra kalan belirsizliği ölçer. Toplam varyans yasası, toplam belirsizliği iki parçaya ayırır: grupların içindeki belirsizlik ve grup ortalamaları arasındaki belirsizlik.

Tanım 9.7 (Koşullu Varyans) $E[X^2] < \infty$ ise

$$\text{Var}(X \mid Y) = E[X^2 \mid Y] - \{E[X \mid Y]\}^2$$

olarak tanımlanır.

Teorem 9.9 (Toplam Varyans Yasası). $E[X^2] < \infty$ ise

$$\text{Var}(X) = E[\text{Var}(X \mid Y)] + \text{Var}(E[X \mid Y])$$

geçerlidir.

Kanıt. Varyans $\text{Var}(X) = E[X^2] - \{E[X]\}^2$ biçimindedir. Koşullu varyans tanımından

$$E[X^2 \mid Y] = \text{Var}(X \mid Y) + \{E[X \mid Y]\}^2$$

yazılır. Beklenti alınırsa

$$E[X^2] = E[\text{Var}(X \mid Y)] + E[\{E[X \mid Y]\}^2].$$

Ayrıca toplam beklenti yasasıyla $E[X] = E[E[X \mid Y]]$ olur. Terimler düzenlenince sonuç elde edilir. $\square$

Örnek 9.11 (Sınıf İçi ve Sınıflar Arası Değişkenlik). *Öğrenci puanı X , okul türü Y olsun. Toplam varyans, okul türleri içindeki puan değişkenliği ile okul türlerinin ortalamaları arasındaki değişkenliğin toplamıdır.*

Örnek 9.12 (Karışım Modelinde Varyans). *İki üretim hattı eşit olasılıkla seçiliyor; birinci hat $N(100, 4)$, ikinci hat $N(106, 4)$ ürettiyor. Y hattı gösterebilir. $E[\text{Var}(X | Y)] = 4$ (grup içi sabit). Grup ortalamaları 100 ve 106 eşit olasılıkla olduğundan $\text{Var}(E[X | Y]) = (106 - 100)^2/4 = 9$. Toplam varyans $4 + 9 = 13$ olur; tek bir normal model bu çift-tepeli yapıyı yakalayamazdı.*

Yorum:

Toplam varyans yasası, belirsizliğin nereden geldiğini gösterir. Eğer $E[\text{Var}(X | Y)]$ büyükse grupların içinde çok değişkenlik vardır. Eğer $\text{Var}(E[X | Y])$ büyükse grupların ortalamaları birbirinden belirgin biçimde farklıdır.

9.9 Bayes Kuralı ve Güncelleme

Bayes kuralı, gözlenen veriden nedene doğru akıl yürütmeyi sağlar. Önsel olasılık, veri olasılığı ve toplam olasılık birlikte kullanılarak sonsal olasılık elde edilir.

Teorem 9.10 (Bayes Kuralı). $A_1, \dots, A_n$ örnek uzayın bir ayrışımı ve $P(B) > 0$ olsun. Bu durumda

$$P(A_i | B) = \frac{P(B | A_i)P(A_i)}{\sum_{j=1}^n P(B | A_j)P(A_j)}$$

geçerlidir.

Kanıt. Koşullu olasılık tanımından

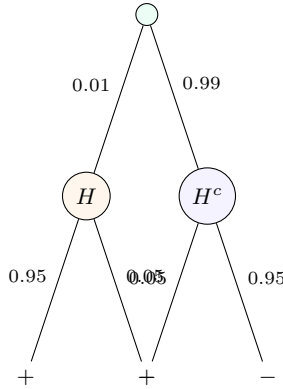
$$P(A_i | B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(B | A_i)P(A_i)}{P(B)}.$$

Payda toplam olasılık yasasıyla $P(B) = \sum_j P(B | A_j)P(A_j)$ biçiminde yazılır. $\square$

Örnek 9.13 (Tıbbi Test). *Hastalık oranı $P(H) = 0.01$, duyarlılık $P(+ | H) = 0.95$ ve yanlış pozitif oranı $P(+ | H^c) = 0.05$ olsun. Pozitif test sonrası hastalık olasılığı*

$$P(H | +) = \frac{0.95 \cdot 0.01}{0.95 \cdot 0.01 + 0.05 \cdot 0.99}$$

olur. Hastalık nadir olduğu için pozitif sonuç sonrası olasılık 0.95 değildir.



Şekil 9.1 Bayes ağacı: önsel dallar (hastalık/sağlam) test sonuçlarına ayrılır. Pozitif yolların ağırlıkları sonsal olasılığı verir.

Bayes kuralı yalnız sonlu hipotezler için değil, sürekli parametreler için de geçerlidir; önsel yoğunluk olabilirlikle çarpılıp normalize edilir (Casella & Berger, 2002; Jaynes, 2003).

Teorem 9.11 (Sürekli Bayes Kuralı). *Parametre θ için önsel yoğunluk $\pi(\theta)$ ve veri olasılığı $f(x | \theta)$ olsun. Sonsal yoğunluk*

$$\pi(\theta | x) = \frac{f(x | \theta)\pi(\theta)}{\int f(x | \vartheta)\pi(\vartheta) d\vartheta}$$

olur; yani sonsal, olabilirlik çarpı önselle orantılıdır.

Kanıt. Ortak yoğunluk $f(x | \theta)\pi(\theta)$ 'dir. Koşullu yoğunluk tanımıyla $\pi(\theta | x)$, bu ortak yoğunluğun x 'in marjinaline bölünmesidir; marjinal, paydadaki integraldir. $\square$

Örnek 9.14 (Beta-Binom Güncellemesi). *Başarı olasılığı θ için önsel $\theta \sim \text{Beta}(a, b)$ ve n denemede k başarı gözlenirse, sonsal yine beta olur: $\theta | k \sim \text{Beta}(a+k, b+n-k)$. Beta ailesinin binom için eşlenik önsel olması, güncellemeyi tek satıra indirir.*

9.10 Koşullu Beklenti ve En İyi Tahmin

Koşullu beklenti yalnızca koşullu ortalama değildir; aynı zamanda verilen bilgiye dayalı en iyi karesel tahmindir. Bu özellik regresyon, makine öğrenmesi ve sinyal tahmininin teorik temelini oluşturur.

Teorem 9.12 (En İyi Kare Hata Tahmini). *$E[Y^2] < \infty$ olsun. Y değişkeni yalnız X bilgisine dayalı $g(X)$ biçimindeki tahminlerle tahmin edilecekse, ortalama*

kare hatayı en küçük yapan tahmin

$$g(X) = E[Y | X]$$

olur.

Kanıt. Herhangi bir $g(X)$ için

$$Y - g(X) = \{Y - E[Y | X]\} + \{E[Y | X] - g(X)\}$$

yazılır. Kare alınıp beklenti uygulandığında çapraz terim sıfır olur; çünkü

$$E[(Y - E[Y | X])h(X)] = 0$$

her uygun h için geçerlidir. Böylece hata, $g(X) = E[Y | X]$ seçildiğinde en küçük olur. $\square$

Teorem 9.13 (Ortogonalite Özelliği). $E[Y^2] < \infty$ ise ve $h(X) = E[Y | X]$ tanımlanırsa, X 'in uygun her fonksiyonu $g(X)$ için

$$E[(Y - E[Y | X])g(X)] = 0$$

olur.

Kanıt. Koşullu beklenti tanımı gereği

$$E[Y - E[Y | X] | X] = 0.$$

Toplam beklenti yasasıyla

$$E[(Y - E[Y | X])g(X)] = E\{g(X)E[Y - E[Y | X] | X]\} = 0$$

elde edilir. $\square$

Örnek 9.15 (Kredi Harcaması Tahmini). *Bir müşterinin gelecek ayki harcaması Y , gelir grubu ve geçmiş harcama özeti X ile tahmin edilecek olsun. $E[Y | X]$, aynı bilgiye dayalı tahminler arasında ortalama kare hatayı en aza indiren teorik tahmindir.*

Koşullu beklenti en iyi tahmindir ama her zaman kapalı biçimde bulunamaz. Tahmini doğrusal fonksiyonlarla sınırlarsak, en iyi doğrusal tahmin korelasyon ve standart sapmalarla açıkça verilir; bu, klasik regresyonun temelidir (Casella & Berger, 2002).

Teorem 9.14 (En İyi Doğrusal Tahmin). $E[X^2], E[Y^2] < \infty$, $\rho = \text{Corr}(X, Y)$, σ_X, σ_Y standart sapmalar olsun. Y 'yi $a + bX$ biçiminde tahmin eden ve ortalama kare hatayı en küçük yapan katsayılar

$$b = \rho \frac{\sigma_Y}{\sigma_X}, \quad a = E[Y] - b E[X]$$

olur. Bu tahminin kalan ortalama kare hatası $\sigma_Y^2(1 - \rho^2)$ 'dir.

Kanıt. $E[(Y - a - bX)^2]$ ifadesinin a ve b 'ye göre kısmi türevleri sıfıra eşitlenir. a türevi $a = E[Y] - bE[X]$ verir. Bunu yerine koyup b türevi alındığında $b = \text{Cov}(X, Y) / \text{Var}(X) = \rho\sigma_Y / \sigma_X$ bulunur. Bu değerler yerine konduğunda kalan hata $\sigma_Y^2(1 - \rho^2)$ olur. $\square$

Örnek 9.16 (Korelasyon ve Açıklanan Değişkenlik). $\rho = 0.8$ ise en iyi doğrusal tahminin kalan hatası $\sigma_Y^2(1 - 0.64) = 0.36\sigma_Y^2$ olur. Yani doğrusal tahmin toplam değişkenliğin yaklaşık %64'ünü açıklar. ρ^2 değeri “açıklanan değişkenlik oranı” olarak yorumlanır.

9.11 Kule Özelliği ve Bilgi Düzeyleri

Koşullu beklenti bilgi düzeyine göre değişir. Daha ayrıntılı bilgiyle yapılan tahmin, daha kaba bilgiye indirgenirse doğrudan kaba bilgiyle yapılan tahmin elde edilir. Bu özelliğe kule özelliği denir.

Teorem 9.15 (Kule Özelliği). $\mathcal{G} \subseteq \mathcal{F}$ iki bilgi kümesi olsun. Uygun integrallelenebilir X için

$$E\{E[X \mid \mathcal{F}] \mid \mathcal{G}\} = E[X \mid \mathcal{G}]$$

geçerlidir.

Kanıt. $\mathcal{F}$ daha ayrıntılı bilgi, $\mathcal{G}$ daha kaba bilgi içerir. Önce ayrıntılı bilgiye göre ortalama alıp sonra kaba bilgiye indirgemek, doğrudan kaba bilgiye göre ortalama almakla aynıdır. Bu, koşullu beklentinin tanımlayıcı özelliklerinden biridir. $\square$

Örnek 9.17 (Gün İçi ve Gün Sonu Bilgi). Bir hisse getirisi için gün içi işlem bilgisi $\mathcal{F}$, yalnız gün sonu özet bilgisi $\mathcal{G}$ olsun. Gün içi bilgiye göre yapılan tahmin daha sonra gün sonu bilgiye indirgenirse, doğrudan gün sonu bilgisiyle yapılan tahmine eşit olur.

9.12 Koşullama ile Parçalama Stratejisi

Bir olasılık veya beklenti doğrudan hesaplanamıyorsa, uygun bir Y değişkenine koşullamak problemi basitleştirebilir. Bu strateji özellikle karışım modellerinde, bekleme sürelerinde, stokastik süreçlerde ve simülasyon analizinde kullanılır.

- Önce problemi kolaylaştıran bilgi değişkenini seç.
- $Y = y$ koşulu altında olasılığı veya beklentiyi hesapla.
- Sonra Y 'nin dağılımına göre ortalama al.

Örnek 9.18 (Rassal Sayıda Talep). *Bir günde gelen müşteri sayısı N ve her müşterinin ortalama harcaması μ olsun. N bilindiğinde toplam harcama S için $E[S | N] = N\mu$ olur. Toplam beklenti yasasıyla*

$$E[S] = E[E[S | N]] = E[N\mu] = \mu E[N]$$

bulunur.

Rassal sayıda bağımsız terimin toplamı, koşullamanın en zarif uygulamalarından biridir. Wald eşitliği bu toplamların ortalama ve varyansını verir (Ross, 2019).

Teorem 9.16 (Rassal Toplam: Wald Eşitliği). *$X_1, X_2, \dots$ bağımsız aynı dağılımlı, $E(X_i) = \mu$, $\text{Var}(X_i) = \sigma^2$ ve N bunlardan bağımsız negatif olmayan tam sayı değerli bir değişken olsun. $S = \sum_{i=1}^N X_i$ için*

$$E[S] = \mu E[N], \quad \text{Var}(S) = \mu^2 \text{Var}(N) + \sigma^2 E[N]$$

olur.

Kanıt. $N = n$ koşulu altında $E[S | N = n] = n\mu$ ve $\text{Var}(S | N = n) = n\sigma^2$ 'dir. Dolayısıyla $E[S | N] = N\mu$ ve $\text{Var}(S | N) = N\sigma^2$. Toplam beklenti yasasıyla $E[S] = \mu E[N]$. Toplam varyans yasasıyla

$$\text{Var}(S) = E[\text{Var}(S | N)] + \text{Var}(E[S | N]) = \sigma^2 E[N] + \mu^2 \text{Var}(N).$$

□

Örnek 9.19 (Toplam Hasar Tutarı). *Bir sigorta portföyünde yıllık hasar sayısı $N \sim \text{Poisson}(\lambda)$ ve her hasar ortalama μ , varyans σ^2 olsun. Poisson için $E[N] = \text{Var}(N) = \lambda$ olduğundan toplam hasar S için $E[S] = \mu\lambda$ ve $\text{Var}(S) = \lambda(\mu^2 + \sigma^2)$ bulunur. Bu, birleşik Poisson modelinin temel sonucudur.*

9.13 Örnek Veri Tablosu ve Yazılım Çözümleri

Aşağıdaki tablo, iki müşteri grubunda gözlenen ortalama aylık hasar tutarını ve grup payını temsil etsin.

Grup	Grup Payı Koşullu Ortalama Hasar
Düşük risk	0.70 1200
Yüksek risk	0.30 3600

Genel beklenen hasar, koşullu ortalamaların grup paylarıyla ağırlıklı ortalamasıdır:

$$E[X] = 1200(0.70) + 3600(0.30) = 1920.$$

```
pay <- c(0.70, 0.30)
kosullu_ortalama <- c(1200, 3600)
sum(pay * kosullu_ortalama)
```

Python Çözümü

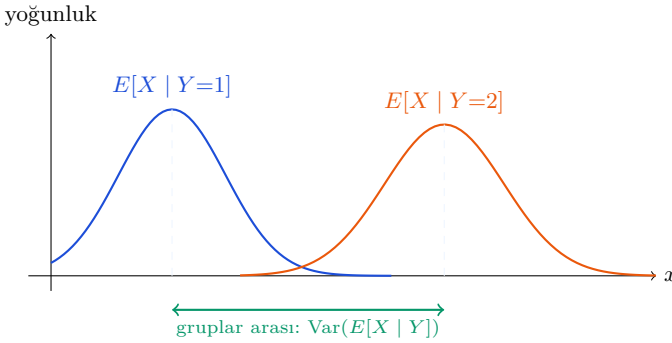
```
import numpy as np

pay = np.array([0.70, 0.30])
kosullu_ortalama = np.array([1200, 3600])
genel_ortalama = np.sum(pay * kosullu_ortalama)

print(genel_ortalama)
```

9.14 Görselleştirme ve Grafikler

Koşullu beklenti, gruplara göre farklı ortalamaları görselleştirmek için kullanılabilir. Aşağıdaki grafik, toplam belirsizliğin grup içi ve gruplar arası iki bileşene ayrıştığını gösterir.



Şekil 9.2 Toplam varyans yasası: her grubun içindeki yayılım ($E[\text{Var}(X | Y)]$) ile grup ortalamaları arasındaki uzaklık ($\text{Var}(E[X | Y])$) toplam varyansı oluşturur.

Aşağıdaki kod basit bir grup-ortalama grafiği üretir.

```
import matplotlib.pyplot as plt

gruplar = ["Düşük risk", "Yüksek risk"]
ortalamalar = [1200, 3600]
```

```
plt.bar(gruplar, ortalamalar, edgecolor="black")
plt.ylabel("E[X | Grup]")
plt.title("Gruba Göre Koşullu Ortalama Hasar")
plt.show()
```

Dikkat: Kavram Yanılgısı

Koşullu beklenti tek bir sayı olmak zorunda değildir. $E[X | Y]$, Y 'nin aldığı bilgiye göre değişen bir fonksiyondur; $E[X | Y = y]$ ise bu fonksiyonun belirli bir y değerindeki sayısal karşılığıdır.

Ufuk Notu:

Makine öğrenmesine köprü: Regresyonda kare kaybını en aza indiren tahmin fonksiyonu koşullu beklentidir: en iyi tahmin $m(y) = E[X | Y = y]$ biçimindedir. Mutlak kayıp altında koşullu medyan, sınıflandırmada ise koşullu sınıf olasılıkları öne çıkar. Bu nedenle koşullu beklenti, modern kayıp fonksiyonları ve denetimli öğrenme için temel bir olasılık kavramıdır.

Kendini Dene:

Ara Kontrol: Kare kaybı yerine mutlak hata kaybı kullanılırsa “en iyi tahmin” neden ortalama yerine medyana yaklaşır?

9.15 Bölüm Sonu Çözümlü Problemler

1. $P(A) = 0.4$, $P(B) = 0.5$ ve $P(A \cap B) = 0.2$ olsun. $P(A | B)$ değerini bulun.
Çözüm.

$$P(A | B) = \frac{P(A \cap B)}{P(B)} = \frac{0.2}{0.5} = 0.4.$$

2. $P(Y = 1) = 0.3$, $P(X = 1, Y = 1) = 0.12$ olsun. $P(X = 1 | Y = 1)$ değerini hesaplayın.
Çözüm.

$$P(X = 1 | Y = 1) = \frac{0.12}{0.3} = 0.4.$$

3. $E[X | Y = 0] = 10$, $E[X | Y = 1] = 20$ ve $P(Y = 1) = 0.25$ olsun. $E[X]$ değerini bulun.
Çözüm. $P(Y = 0) = 0.75$ olduğundan

$$E[X] = 10(0.75) + 20(0.25) = 12.5.$$

4. $\text{Var}(X | Y) = 4$ sabit ve $\text{Var}(E[X | Y]) = 9$ olsun. $\text{Var}(X)$ değerini bulun.

Çözüm. Toplam varyans yasasıyla

$$\text{Var}(X) = E[\text{Var}(X | Y)] + \text{Var}(E[X | Y]) = 4 + 9 = 13.$$

5. Bir hastalık oranı 0.02, test duyarlılığı 0.90, yanlış pozitif oranı 0.04 olsun. Pozitif test sonrası hastalık olasılığını yazın.

Çözüm.

$$P(H | +) = \frac{0.90 \cdot 0.02}{0.90 \cdot 0.02 + 0.04 \cdot 0.98}.$$

6. $E[Y | X] = 2X + 1$ ve $E[X] = 3$ olsun. $E[Y]$ değerini bulun.

Çözüm. Toplam beklenti yasasıyla

$$E[Y] = E[E[Y | X]] = E[2X + 1] = 2E[X] + 1 = 7.$$

7. Bir fabrikada parçaların %60'ı A makinesinden (%2 kusur), %40'ı B makinesinden (%5 kusur) geliyor. Rastgele bir parçanın kusurlu olma olasılığı nedir?

Çözüm. Toplam olasılık yasasıyla $P(\text{kusur}) = 0.02(0.6) + 0.05(0.4) = 0.032$.

8. Hasar sayısı $N \sim \text{Poisson}(4)$ ve her hasar ortalama $\mu = 500$ olsun. Toplam beklenen hasar nedir?

Çözüm. Wald eşitliğiyle $E[S] = \mu E[N] = 500 \cdot 4 = 2000$.

9. Yukarıdaki problemde her hasarın varyansı $\sigma^2 = 10000$ ise toplam hasarın varyansı nedir?

Çözüm. Poisson için $E[N] = \text{Var}(N) = 4$. $\text{Var}(S) = \mu^2 \text{Var}(N) + \sigma^2 E[N] = 500^2 \cdot 4 + 10000 \cdot 4 = 1.04 \times 10^6$.

10. $E[Y^2 | X]$ ifadesi $Y = X$ olduğunda neye eşittir?

Çözüm. X koşulu altında $Y = X$ bilindiğinden $E[X^2 | X] = X^2$ olur (bilineni dışarı alma).

11. Başarı olasılığı için önsel $\theta \sim \text{Beta}(2, 2)$ ve 10 denemede 7 başarı gözleniyor. Sonsal dağılım nedir?

Çözüm. Beta-binom eşlenikliğiyle $\theta | \text{veri} \sim \text{Beta}(2 + 7, 2 + 3) = \text{Beta}(9, 5)$.

9.16 Bölüm Sonu Alıştırmaları

- E1. Koşullu beklentiye sözel olarak tanımlayın ve neden kendisinin de bir rassal değişken olduğunu açıklayın.
- E2. Toplam beklenti yasasını kullanarak $E(X)$ değerini $A_1, \dots, A_k$ ayrışımı cinsinden yazın.
- E3. Bayes teoreminin yeni kanıtla başlangıç olasılığını güncellediği uygulamalı bir örnek verin.
- E4. Koşullu bağımsızlığın sıradan bağımsızlıktan nasıl ayrıldığını açıklayın.

9.17 Kaynak Notu

Bu bölümdeki koşullu dağılım, koşullu yoğunluk, koşullu beklenti, toplam beklenti yasası, toplam varyans yasası, Bayes kuralı, kule özelliği ve en iyi karesel tahmin yorumu; Ross (2019), Hogg, McKean ve Craig (2019), Casella ve Berger (2002), Grimmett ve Stirzaker (2020), Bertsekas ve Tsitsiklis (2008) ve Billingsley (1995) çizgisindeki klasik olasılık ve istatistik kaynaklarıyla uyumludur. Bilineni dışarı alma ve kule özelliğinin ölçü-teorik temeli için Williams (1991); Wald eşitliği ve birleşik Poisson modeli için Ross (2019); sürekli Bayes ve beta-binom eşlenikliği için Jaynes (2003) ile Casella ve Berger (2002) izlenmiştir.

Özet:

Koşullu dağılımlar, yeni bilgi geldiğinde rassal değişkenin dağılımının nasıl değiştiğini gösterir. Koşullu beklenti, bu yeni bilgi altındaki ortalamadır ve aynı zamanda en iyi karesel tahmin aracıdır. Toplam beklenti yasası genel ortalamayı koşullu ortalamaların ağırlıklı ortalaması olarak, toplam varyans yasası ise toplam belirsizliği grup içi ve gruplar arası bileşenler olarak ifade eder. Bayes kuralı gözlemden nedene doğru olasılık güncellemesi sağlar.

Bölüm 10

Olasılık Eşitsizlikleri ve Yakınsama

10.1 Gerçek Dünya Problemi (Giriş)

Bir mühendislik sisteminde hata değişkeninin tam dağılımı bilinmeyebilir; fakat ortalaması ve varyansı yaklaşık olarak biliniyor olabilir. Bir finans portföyünde kaybın bütün dağılımını modellemek zor olabilir; yine de kaybın belirli bir eşiği aşma olasılığı için güvenli bir üst sınır istenir. Bir simülasyonda tahminlerin örneklem sayısı arttıkça gerçek değere yaklaşıp yaklaşmadığı sorgulanır.

Olasılık eşitsizlikleri bu tür durumlarda dağılımın tamamını bilmeden risk sınırı verir. Yakınsama kavramları ise rassal büyüklüklerin uzun dönemde nasıl kararlılaştığını açıklar. Böylece önceki bölümlerde kullanılan büyük sayılar yasası ve merkezi limit teoremi daha sağlam bir matematiksel zemine oturur.

Bu Bölüm Ne Sağlar?

Bu bölüm Markov, Çebişev, Jensen, Cauchy–Schwarz ve Boole eşitsizliklerini tanıtır; moment ve kuyruk sınırlarının nasıl üretildiğini gösterir. Ardından olasılıkta, dağılımda, ortalamada, ortalama karede ve hemen hemen kesin yakınsama kavramları karşılaştırılır. Son kısımda Borel–Cantelli lemması, yakınsama ilişkileri ve eşitsizliklerin büyük sayılar yasasıyla bağlantısı ele alınır.

Amaç ve Öğrenme Çıktıları

Bu bölümün sonunda okuyucu

- Markov ve Çebişev eşitsizliklerini kullanarak kuyruk olasılıklarına üst sınır verebilir,
- Jensen ve Cauchy–Schwarz eşitsizliklerini beklenti ve moment problemlerine uygulayabilir,
- Boole eşitsizliği ile birleşim olaylarını sınırlayabilir,
- olasılıkta, dağılımda, L^1 , L^2 ve hemen hemen kesin yakınsama türlerini ayırt edebilir,
- yakınsama türleri arasındaki temel ilişkileri açıklayabilir,
- eşitsizliklerin zayıf büyük sayılar yasasına nasıl temel oluşturduğunu gösterebilir.

10.2 Neden Olasılık Eşitsizlikleri?

Tam dağılım bilindiğinde istenen olasılıklar doğrudan hesaplanabilir. Fakat uygulamada çoğu zaman yalnızca ortalama, varyans veya birkaç moment bilinir. Olasılık eşitsizlikleri, sınırlı bilgiyle güvenli ve genellikle dağılımdan bağımsız sınırlar üretir.

Örnek 10.1 (Dağılım Bilinmediğinde Risk). *Bir sensör hatasının ortalaması 0 ve varyansı 4 olsun. Hata dağılımı bilinmese bile $|X| \geq 10$ olasılığı Çebişev eşitsizliğiyle en fazla $4/100 = 0.04$ olarak sınırlanabilir.*

Eşitsizlik sınırları çoğu zaman kaba fakat güvenlidir. Amaç tam olasılığı bulmak değil, eldeki az bilgiyle kesin geçerli bir üst sınır üretmektir.

10.3 Markov Eşitsizliği

Markov eşitsizliği, negatif olmayan bir rassal değişkenin büyük değer alma olasılığını yalnız beklenen değeriyle sınırlar. Pek çok daha güçlü eşitsizlik Markov eşitsizliğinin uygun dönüşümlere uygulanmasıyla elde edilir.

Teorem 10.1 (Markov Eşitsizliği). $X \geq 0$ ve $a > 0$ ise

$$P(X \geq a) \leq \frac{E[X]}{a}.$$

Kanıt. $X \geq aI_{\{X \geq a\}}$ eşitsizliği her örnek nokta için doğrudur. Beklenti alınırsa

$$E[X] \geq aP(X \geq a)$$

olur. Her iki taraf a 'ya bölünür. $\square$

Örnek 10.2 (Kayıp Sınırı). *Negatif olmayan günlük kayıp X için $E[X] = 2$ milyon TL olsun. Kaybın 20 milyon TL'yi aşma olasılığı*

$$P(X \geq 20) \leq \frac{2}{20} = 0.10$$

ile sınırlanır.

Markov eşitsizliğini doğrudan X 'e değil, e^{tX} üstel dönüşümüne uygulamak çok daha keskin, üstel hızda azalan kuyruk sınırları verir. Bu fikir Chernoff sınırının temelidir ve modern olasılıkta yoğunlaşma eşitsizliklerinin çıkış noktasıdır (Billingsley, 1995; Vershynin, 2018).

Teorem 10.2 (Chernoff Sınırı). *X değişkeninin moment üreten fonksiyonu $M_X(t) = E[e^{tX}]$ var olsun. Her a ve her $t > 0$ için*

$$P(X \geq a) \leq e^{-ta} M_X(t)$$

olur. En iyi sınır, sağ tarafı $t > 0$ üzerinde en küçük yapan t seçimiyle elde edilir.

Kanıt. $t > 0$ için $X \geq a$ olayı $e^{tX} \geq e^{ta}$ olayına denktir. Negatif olmayan e^{tX} değişkenine Markov eşitsizliği uygulanır:

$$P(X \geq a) = P(e^{tX} \geq e^{ta}) \leq \frac{E[e^{tX}]}{e^{ta}} = e^{-ta} M_X(t).$$

Eşitsizlik her $t > 0$ için geçerli olduğundan en küçük üst sınır alınabilir. $\square$

Örnek 10.3 (Standart Normal Kuyruğu). *$Z \sim N(0, 1)$ için $M_Z(t) = e^{t^2/2}$ olduğundan $P(Z \geq a) \leq e^{-ta+t^2/2}$. Sağ taraf $t = a$ 'da en küçüktür ve*

$$P(Z \geq a) \leq e^{-a^2/2}$$

sınırını verir. Bu, Çebişev'in verdiği $1/a^2$ tipi sınırdan çok daha hızlı azalır.

10.4 Çebişev Eşitsizliği

Çebişev eşitsizliği, sonlu varyansa sahip herhangi bir dağılımda ortalamadan büyük sapmaların olasılığını sınırlar. Dağılımın normal, simetrik veya kesikli olması gerekmez.

Teorem 10.3 (Çebişev Eşitsizliği). $E[X] = \mu$ ve $\text{Var}(X) = \sigma^2 < \infty$ ise her $\varepsilon > 0$ için

$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}.$$

Kanıt. Negatif olmayan $Y = (X - \mu)^2$ değişkenine Markov eşitsizliği uygulanır:

$$P(|X - \mu| \geq \varepsilon) = P((X - \mu)^2 \geq \varepsilon^2) \leq \frac{E[(X - \mu)^2]}{\varepsilon^2} = \frac{\sigma^2}{\varepsilon^2}.$$

□

Örnek 10.4 (Ölçüm Hatası). Bir ölçüm hatasının varyansı 9 ise hatanın ortalamadan en az 15 birim sapma olasılığı

$$P(|X - \mu| \geq 15) \leq \frac{9}{15^2} = 0.04$$

olur.

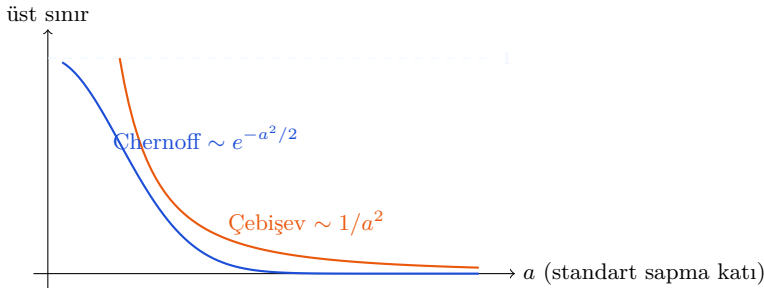
Çebişev eşitsizliği iki yönlü sapmayı sınırlar. Yalnız tek yönde sapma ilgilenildiğinde, daha keskin olan tek taraflı (Cantelli) eşitsizliği kullanılır (Grimmett & Stirzaker, 2020).

Teorem 10.4 (Tek Taraflı Çebişev (Cantelli)). $E[X] = \mu$, $\text{Var}(X) = \sigma^2$ ve $a > 0$ ise

$$P(X - \mu \geq a) \leq \frac{\sigma^2}{\sigma^2 + a^2}$$

olur.

Kanıt. $Y = X - \mu$ olsun; $E[Y] = 0$, $\text{Var}(Y) = \sigma^2$. Her $t > 0$ için $P(Y \geq a) = P(Y + t \geq a + t) \leq P((Y + t)^2 \geq (a + t)^2)$. Markov eşitsizliğiyle bu $E[(Y + t)^2]/(a + t)^2 = (\sigma^2 + t^2)/(a + t)^2$ ile sınırlıdır. Sağ taraf $t = \sigma^2/a$ 'da en küçüktür ve $\sigma^2/(\sigma^2 + a^2)$ değerini verir. □



Şekil 10.1 Kuyruk sınırlarının karşılaştırması: Chernoff (üstel) sınırı, Çebişev'in polinom sınırından çok daha hızlı azalır.

10.5 Moment Sınırları ve Genelleştirme

Markov eşitsizliği yalnız X değişkenine değil, $|X|^p$ gibi negatif olmayan dönüşümlere de uygulanabilir. Böylece daha yüksek momentler bilindiğinde kuyruk olasılıkları için daha güçlü sınırlar elde edilir.

Teorem 10.5 (Moment Eşitsizliği). $p > 0$ ve $E[|X|^p] < \infty$ ise her $a > 0$ için

$$P(|X| \geq a) \leq \frac{E[|X|^p]}{a^p}.$$

Kanıt. $Y = |X|^p$ negatif olmayan bir değişkendir. Markov eşitsizliğiyle

$$P(|X| \geq a) = P(|X|^p \geq a^p) \leq \frac{E[|X|^p]}{a^p}.$$

□

Tanım 10.1 (L^p Uzayı) $p \geq 1$ için $E[|X|^p] < \infty$ koşulunu sağlayan rassal değişkenlerin sınıfına L^p uzayı denir. L^1 sonlu mutlak beklentiyi, L^2 sonlu ikinci momenti ifade eder.

10.6 Jensen Eşitsizliği

Jensen eşitsizliği, konveks fonksiyonlarla beklenti arasındaki temel ilişkiyi verir. Varyansın negatif olmaması, aritmetik-geometrik ortalama ilişkileri ve bilgi kuramındaki birçok sonuç bu eşitsizlikle bağlantılıdır.

Tanım 10.2 (Konveks Fonksiyon) Bir ϕ fonksiyonu, her $0 \leq t \leq 1$ için

$$\phi(tx + (1-t)y) \leq t\phi(x) + (1-t)\phi(y)$$

koşulunu sağlıyorsa konvektir.

Teorem 10.6 (Jensen Eşitsizliği). ϕ konveks ve ilgili beklentiler sonlu ise

$$\phi(E[X]) \leq E[\phi(X)].$$

Kanıt. Kesikli durumda $E[X]$, X değerlerinin olasılıklarla ağırlıklı ortalamasıdır. Konvekslik tanımı sonlu ağırlıklı ortalamalara uygulanınca sonuç elde edilir. Genel durum, kesikli yaklaşımlar ve limit argümanı ile tamamlanır. □

Örnek 10.5 (Varyansın Negatif Olmaması). $\phi(x) = x^2$ konvektir. Jensen eşitsizliği

$$(E[X])^2 \leq E[X^2]$$

verir. Dolayısıyla

$$\text{Var}(X) = E[X^2] - (E[X])^2 \geq 0.$$

Örnek 10.6 (Aritmetik-Geometrik Ortalama). $\phi(x) = -\log x$ konvektir. Pozitif değerler alan X için Jensen eşitsizliği $-\log E[X] \leq E[-\log X]$, yani $E[\log X] \leq \log E[X]$ verir. Sonlu eşit ağırlıklı durumda bu, geometrik ortalamanın aritmetik ortalamaı aşamayacağını söyler:

$$(x_1 \cdots x_n)^{1/n} \leq \frac{x_1 + \cdots + x_n}{n}.$$

10.7 Cauchy–Schwarz Eşitsizliği

Cauchy–Schwarz eşitsizliği, ikinci momentleri sonlu iki değişkenin çarpım beklentisini kontrol eder. Kovaryans, korelasyon ve regresyon teorisinde temel bir araçtır.

Teorem 10.7 (Cauchy–Schwarz Eşitsizliği). $X, Y \in L^2$ ise

$$|E[XY]| \leq \sqrt{E[X^2]E[Y^2]}.$$

Kanıt. Her gerçekte t için $E[(X - tY)^2] \geq 0$ olur. Bu ifade t 'ye göre

$$E[Y^2]t^2 - 2E[XY]t + E[X^2]$$

polinomudur. Her t için negatif olmaması diskriminantın pozitif olamamasını gerektirir. Böylece $(E[XY])^2 \leq E[X^2]E[Y^2]$ elde edilir. $\square$

Teorem 10.8 (Kovaryans Sınırı). *Sonlu ikinci momentler altında*

$$|\text{Cov}(X, Y)| \leq \sqrt{\text{Var}(X) \text{Var}(Y)}.$$

Kanıt. $U = X - E[X]$ ve $V = Y - E[Y]$ olsun. Cauchy–Schwarz eşitsizliği $|E[UV]| \leq \sqrt{E[U^2]E[V^2]}$ verir. Sol taraf $|\text{Cov}(X, Y)|$, sağ taraf $\sqrt{\text{Var}(X) \text{Var}(Y)}$ olur. $\square$

Örnek 10.7 (Korelasyonun Sınırı). $\rho = \text{Cov}(X, Y)/(\sigma_X \sigma_Y)$ olduğundan kovaryans sınırı $|\rho| \leq 1$ sonucunu verir.

Cauchy–Schwarz, daha genel olan Hölder eşitsizliğinin $p = q = 2$ özel durumudur. Hölder, farklı moment mertebelerini birbirine bağlayan temel araçtır (Billingsley, 1995).

Teorem 10.9 (Hölder Eşitsizliği). $p, q > 1$ ve $1/p + 1/q = 1$ olsun. $X \in L^p$, $Y \in L^q$ ise

$$E[|XY|] \leq (E[|X|^p])^{1/p} (E[|Y|^q])^{1/q}$$

olur.

Kanıt. Young eşitsizliği gereği negatif olmayan u, v için $uv \leq u^p/p + v^q/q$ 'dur. $u = |X|/\|X\|_p$ ve $v = |Y|/\|Y\|_q$ alınıp beklenti uygulanır; sağ taraf $1/p + 1/q = 1$ olduğundan 1'e eşitlenir ve eşitsizlik elde edilir. $\square$

Örnek 10.8 (Moment Mertebeleri Arasında). *Hölder eşitsizliğinin bir sonucu, $r < s$ için $(E|X|^r)^{1/r} \leq (E|X|^s)^{1/s}$ olmasıdır. Yani yüksek mertebeden moment normu daima daha büyük veya eşittir; bu, $L^s \subseteq L^r$ içermesini açıklar.*

10.8 Boole Eşitsizliği ve Birleşim Sınırı

Birden çok risk kaynağı olduğunda en az birinin gerçekleşme olasılığı önemlidir. Boole eşitsizliği, olaylar bağımsız olmasa bile birleşim olasılığına güvenli bir üst sınır verir.

Teorem 10.10 (Boole Eşitsizliği). $A_1, A_2, \dots, A_n$ olayları için

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

Sonsuz sayıda olay için de

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

geçerlidir.

Kanıt. Ayırık olmayan olaylar ayırık parçalara ayrılabilir veya gösterge fonksiyonları kullanılabilir. $I_{\bigcup_i A_i} \leq \sum_i I_{A_i}$ olduğundan beklenti alınır ve sonuç elde edilir. $\square$

Örnek 10.9 (Çoklu Sistem Riski). *Beş alt sistemin arıza olasılıkları 0.01 ile 0.03 arasında olsun ve toplamaları 0.09 olsun. En az bir arıza olasılığı, bağımlılık yapısı bilinmese bile en fazla 0.09 olur.*

Boole eşitsizliği bir üst sınır verir; karşı yönde Bonferroni eşitsizlikleri alt sınır sağlar. Çoklu karşılaştırma ve eşzamanlı güven aralıklarında yaygın kullanılır (Casella & Berger, 2002).

Teorem 10.11 (Bonferroni Alt Sınırı). $A_1, \dots, A_n$ olayları için

$$P\left(\bigcup_{i=1}^n A_i\right) \geq \sum_{i=1}^n P(A_i) - \sum_{i < j} P(A_i \cap A_j)$$

olur. Ayrıca tümleyenler üzerinden $P(\bigcap_i A_i) \geq 1 - \sum_i P(A_i^c)$ geçerlidir.

Kanıt. İçerme-dışlama açılımı, ardışık kısmi toplamlar arasında salınır; ilk iki terimde kesildiğinde elde edilen ifade gerçek olasılık için bir alt sınırdır. İkinci eşitsizlik, $\bigcap_i A_i$ 'nin tümleyenine Boole eşitsizliği uygulanmasıyla çıkar. $\square$

Örnek 10.10 (Eşzamanlı Güven). *m ayrı parametrenin her biri için %95 güven aralığı kuruluyorsa, her birinin hata olasılığı 0.05'tir. Bonferroni gereği hepsinin aynı anda doğru olma olasılığı en az $1 - m(0.05)$ 'tir; $m = 5$ için bu en az 0.75 olur.*

10.9 Yakınsama Türleri

Rassal değişken dizileri farklı anlamlarda yakınsayabilir. Limit teoremlerini doğru yorumlamak için hangi yakınsama türünün kullanıldığını bilmek gerekir.

Tanım 10.3 (Olasılıkta Yakınsama) Her $\varepsilon > 0$ için

$$P(|X_n - X| > \varepsilon) \rightarrow 0$$

ise $X_n \xrightarrow{p} X$ yazılır.

Tanım 10.4 (Dağılımda Yakınsama) X_n dağılım fonksiyonları F_n , X dağılım fonksiyonu F olsun. F 'nin sürekli olduğu her x noktasında $F_n(x) \rightarrow F(x)$ ise

$$X_n \xrightarrow{d} X$$

yazılır.

Tanım 10.5 (L^1 ve L^2 Yakınsama) $E[|X_n - X|] \rightarrow 0$ ise X_n , X 'e ortalamada veya L^1 anlamında yakınsar. $E[(X_n - X)^2] \rightarrow 0$ ise ortalama karede veya L^2 anlamında yakınsar.

Tanım 10.6 (Hemen Hemen Kesin Yakınsama)

$$P\{\omega : X_n(\omega) \rightarrow X(\omega)\} = 1$$

ise $X_n \xrightarrow{a.s.} X$ yazılır.

10.10 Yakınsama İlişkileri

Yakınsama türleri güç bakımından farklıdır. Hemen hemen kesin yakınsama ve L^2 yakınsama genellikle olasılıkta yakınsamadan daha güçlüdür; dağılımda yakınsama ise daha zayıftır.

Teorem 10.12 (Temel Yakınsama İlişkileri). *Aşağıdaki çıkarımlar geçerlidir:*

$$X_n \xrightarrow{a.s.} X \implies X_n \xrightarrow{p} X \implies X_n \xrightarrow{d} X.$$

Ayrıca

$$X_n \xrightarrow{L^2} X \implies X_n \xrightarrow{L^1} X \implies X_n \xrightarrow{p} X.$$

Kanıt. İlk zincirde hemen hemen kesin yakınsama, sapma olaylarının olasılığının sıfıra gitmesini sağlar; olasılıkta yakınsama da dağılım fonksiyonlarının süreklilik noktalarında yakınsamasını verir. İkinci zincirde Cauchy–Schwarz eşitsizliği $E|X_n - X| \leq \sqrt{E[(X_n - X)^2]}$ sonucunu verir. L^1 yakınsamadan olasılıkta yakınsama ise Markov eşitsizliğiyle

$$P(|X_n - X| > \varepsilon) \leq \frac{E|X_n - X|}{\varepsilon}$$

şeklinde elde edilir. □

Yorum:

Merkezi limit teoreminde yakınsama dağılımda yakınsamadır. Büyük sayılar yasası ise genellikle olasılıkta veya hemen hemen kesin yakınsama biçiminde ifade edilir.

10.11 Borel–Cantelli Lemması

Borel–Cantelli lemması, olayların sonsuz kez gerçekleşip gerçekleşmeyeceğini toplam olasılıklar üzerinden inceler. Hemen hemen kesin yakınsama ispatlarında sık kullanılan temel araçlardan biridir.

Teorem 10.13 (Borel–Cantelli Lemması). *Olaylar $A_1, A_2, \dots$ için*

$$\sum_{n=1}^{\infty} P(A_n) < \infty$$

ise

$$P(A_n \text{ sonsuz kez gerçekleşir}) = 0.$$

Kanıt. $\{A_n \text{ sonsuz kez}\} = \bigcap_{m=1}^{\infty} \bigcup_{n \geq m} A_n$ yazılır. Boole eşitsizliğiyle

$$P\left(\bigcup_{n \geq m} A_n\right) \leq \sum_{n \geq m} P(A_n).$$

Seri yakınsak olduğundan kuyruk toplamı sıfıra gider. Azalan olayların sürekliliği sonucu verir. $\square$

Örnek 10.11 (Nadir Hatalar). *Bir algoritmanın n 'inci çalışmadaki hata olasılığı $1/n^2$ ise $\sum_n 1/n^2 < \infty$ olduğundan hatalar olasılık bir ile yalnız sonlu sayıda gerçekleşir.*

Birinci lemmanın tersi, olaylar bağımsız ve olasılıklar toplamı ıraksak olduğunda olayların sonsuz kez gerçekleştiğini söyler. İki lemma birlikte sıfır–bir tipi sonuçlar üretir.

Teorem 10.14 (İkinci Borel–Cantelli Lemması). *$A_1, A_2, \dots$ bağımsız olaylar ve $\sum_{n=1}^{\infty} P(A_n) = \infty$ ise*

$$P(A_n \text{ sonsuz kez gerçekleşir}) = 1$$

olur.

Kanıt. Tümleyenler üzerinden $P(\bigcap_{n \geq m} A_n^c) = \prod_{n \geq m} (1 - P(A_n))$ yazılır. $1 - x \leq e^{-x}$ eşitsizliğiyle bu çarpım $\exp(-\sum_{n \geq m} P(A_n)) = 0$ ile sınırlıdır; çünkü seri ıraksaktır. Dolayısıyla hiçbir kuyruktan itibaren olayların tümünün gerçekleşmemesi olasılığı sıfırdır. $\square$

Örnek 10.12 (Sonsuz Kez Başarı). *Bağımsız denemelerde n 'inci denemenin başarı olasılığı $1/n$ ise $\sum_n 1/n = \infty$ ıraksak olduğundan başarılar olasılık bir ile sonsuz kez gerçekleşir; oysa her tek olasılık giderek küçülür.*

10.12 Yoğunlaşma: Hoeffding Eşitsizliği

Sınırlı bağımsız değişkenlerin toplamı, ortalamasından sapma olasılığı üstel hızda azalan güçlü bir yoğunlaşma sergiler. Hoeffding eşitsizliği bunu açık bir sınırla ifade eder ve modern istatistik ile öğrenme kuramının temel araçlarından biridir (Vershynin, 2018).

Teorem 10.15 (Hoeffding Eşitsizliği). *$X_1, \dots, X_n$ bağımsız ve $a_i \leq X_i \leq b_i$ olsun. $S_n = \sum_{i=1}^n X_i$ için her $t > 0$ değerinde*

$$P(S_n - E[S_n] \geq t) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right)$$

olur.

Kanıt. Chernoff yöntemiyle $P(S_n - E[S_n] \geq t) \leq e^{-st} \prod_i E[e^{s(X_i - EX_i)}]$ yazılır. Hoeffding yardımcı önermesi, $[a_i, b_i]$ aralığındaki merkezlenmiş bir değişken için $E[e^{s(X_i - EX_i)}] \leq \exp(s^2(b_i - a_i)^2/8)$ verir. Çarpım alınıp s üzerinde en iyilenir; $s = 4t/\sum (b_i - a_i)^2$ seçimi sonucu verir. $\square$

Örnek 10.13 (Yazı-Tura Oramı). *Adil bir parada n atışta yazı oranı $\hat{p}$ için $X_i \in [0, 1]$ ve $(b_i - a_i)^2 = 1$ olur. Hoeffding gereği $P(\hat{p} - 0.5 \geq \varepsilon) \leq e^{-2n\varepsilon^2}$. $n = 1000$, $\varepsilon = 0.05$ için sınır $e^{-5} \approx 0.0067$; bu, Çebişev sınırından çok daha keskindir.*

10.13 Eşitsizliklerden Büyük Sayılar Yasasına

Çebişev eşitsizliği, zayıf büyük sayılar yasasının kısa ve güçlü bir ispatını verir. Böylece eşitsizlikler yalnız risk sınırı değil, limit teoremlerinin de temel aracıdır.

Teorem 10.16 (Çebişev ile Zayıf Büyük Sayılar Yasası). $X_1, X_2, \dots$ bağımsız ve aynı dağılımlı, $E[X_i] = \mu$ ve $\text{Var}(X_i) = \sigma^2 < \infty$ olsun. O halde

$$\bar{X}_n \xrightarrow{p} \mu.$$

Kanıt. $E[\bar{X}_n] = \mu$ ve $\text{Var}(\bar{X}_n) = \sigma^2/n$ olur. Çebişev eşitsizliğiyle

$$P(|\bar{X}_n - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2} \rightarrow 0.$$

Bu, olasılıkta yakınsama tanımıdır. □

Örnek 10.14 (Anket Oramı). $S_n \sim \text{Bin}(n, p)$ ise örnek oranı $\hat{p} = S_n/n$ için

$$P(|\hat{p} - p| \geq \varepsilon) \leq \frac{p(1-p)}{n\varepsilon^2}$$

olur. Örneklem büyüdükçe oran tahmini gerçek orana yaklaşır.

10.14 Örnek Veri Tablosu ve Yazılım Çözümleri

Aşağıdaki tablo, bir sensörün 8 ölçüm hatasını göstermektedir.

Ölçüm	1	2	3	4	5	6	7	8
Hata	-2	1	0	3	-1	2	-3	1

Bu veriden örnek ortalama ve örnek varyans hesaplanıp, Çebişev eşitsizliğiyle büyük sapma için kaba üst sınır üretilebilir.

```
x <- c(-2, 1, 0, 3, -1, 2, -3, 1)
s2 <- var(x)
epsilon <- 5
s2 / epsilon^2
```

Python Çözümü

```
import numpy as np

x = np.array([-2, 1, 0, 3, -1, 2, -3, 1])
s2 = np.var(x, ddof=1)
epsilon = 5
chebyshev_bound = s2 / epsilon**2

print(s2, chebyshev_bound)
```

10.15 Görselleştirme ve Grafikler

Çebişev sınırı, eşik büyüdükçe karesel olarak azalır. Aşağıdaki kod farklı eşikler için σ^2/ε^2 sınırını çizer.

```
import numpy as np
import matplotlib.pyplot as plt

sigma2 = 4
eps = np.linspace(1, 10, 100)
bound = sigma2 / eps**2

plt.plot(eps, bound)
plt.xlabel("epsilon")
plt.ylabel("Chebyshev üst sınırı")
plt.title("Sapma Eşiği Büyüdükçe Üst Sınır Azalır")
plt.ylim(0, 1)
plt.show()
```

Dikkat: Kavram Yanılgısı

Bir eşitsizliğin gevşek olması onun işe yaramadığı anlamına gelmez. Markov ve Çebişev eşitsizlikleri, dağılım tam bilinmediğinde yalnız az sayıda momentle güvenli ve genel sınırlar verir.

Ufuk Notu:

Veri bilimi ve PAC yorumu: Hoeffding tipi yoğunlaşma eşitsizlikleri, örneklem ortalamasının gerçek ortalamadan ne kadar sapabileceğini önceden sınırlar. Bu fikir, makine öğrenmesindeki PAC yaklaşımının temel sezgisidir: sonlu örneklemle öğrenilen modelin hata payı, belirli bir güven düzeyiyle kontrol edilebilir.

Kendini Dene:

Ara Kontrol: Aynı ortalama ve varyansa sahip iki dağılımın kuyruk davranışları farklı olabilir. Buna rağmen Çebişev eşitsizliği neden her ikisi için de geçerlidir?

10.16 Bölüm Sonu Çözümlü Problemler

1. $X \geq 0$ ve $E[X] = 6$ olsun. $P(X \geq 30)$ için Markov üst sınırını bulun.

Çözüm.

$$P(X \geq 30) \leq \frac{6}{30} = 0.2.$$

2. $E[X] = 10$ ve $\text{Var}(X) = 16$ olsun. $P(|X - 10| \geq 8)$ için Çebişev sınırını bulun.

Çözüm.

$$P(|X - 10| \geq 8) \leq \frac{16}{8^2} = 0.25.$$

3. $E[X^4] = 80$ olsun. $P(|X| \geq 4)$ için moment sınırı verin.

Çözüm.

$$P(|X| \geq 4) \leq \frac{E[X^4]}{4^4} = \frac{80}{256} = 0.3125.$$

4. $X, Y \in L^2$, $E[X^2] = 9$ ve $E[Y^2] = 16$ olsun. $|E[XY]|$ için üst sınır nedir?

Çözüm. Cauchy-Schwarz ile

$$|E[XY]| \leq \sqrt{9 \cdot 16} = 12.$$

5. $X_n \xrightarrow{L^2} X$ ise hangi yakınsama sonucu kesin olarak elde edilir?

Çözüm. L^2 yakınsama önce L^1 yakınsamayı, ardından Markov eşitsizliğiyle olasılıkta yakınsamayı verir. Dolayısıyla $X_n \xrightarrow{P} X$ elde edilir.

6. $P(A_n) = 1/n^2$ olsun. A_n olayları sonsuz kez gerçekleşir mi?

Çözüm. $\sum_n 1/n^2 < \infty$ olduğundan Borel-Cantelli lemmasına göre

$$P(A_n \text{ sonsuz kez}) = 0.$$

7. $Z \sim N(0, 1)$ için $P(Z \geq 3)$ değerine Chernoff üst sınırı verin.

Çözüm. $P(Z \geq a) \leq e^{-a^2/2}$ olduğundan $P(Z \geq 3) \leq e^{-4.5} \approx 0.011$.

8. $E[X] = \mu$, $\text{Var}(X) = \sigma^2$ olsun. $P(X - \mu \geq 2\sigma)$ için tek taraflı Çebişev sınırını bulun.

Çözüm. Cantelli ile $P(X - \mu \geq 2\sigma) \leq \sigma^2/(\sigma^2 + 4\sigma^2) = 1/5 = 0.2$. Bu, tek yönlü sapma için iki taraflı Çebişev sınırı 0.25'ten daha keskindir. İki taraflı sınırı ikiye bölerek 0.125 elde etmek ise genel olarak geçerli değildir; simetri varsayımı gerektirir.

9. 1000 adil para atışında yazı oranının 0.5'i en az 0.05 aşma olasılığına Hoeffding sınırı verin.
Çözüm. $P(\hat{p} - 0.5 \geq 0.05) \leq e^{-2(1000)(0.05)^2} = e^{-5} \approx 0.0067$.
10. Bağımsız olaylar için $P(A_n) = 1/n$ olsun. Olaylar sonsuz kez gerçekleşir mi?
Çözüm. $\sum_n 1/n = \infty$ ve olaylar bağımsız olduğundan ikinci Borel–Cantelli lemmasıyla olasılık 1 ile sonsuz kez gerçekleşir.
11. 5 ayrı %95 güven aralığı için hepsinin aynı anda doğru olma olasılığına Bonferroni alt sınırı nedir?
Çözüm. $1 - 5(0.05) = 0.75$; yani en az 0.75 olasılıkla beşi birden doğrudur.

10.17 Bölüm Sonu Alıştırmaları

Temel Düzey

- B1. $X \geq 0$ ve $E[X] = 12$ olsun. $P(X \geq 30)$ için üst sınır verin.
 B2. $E[X] = 5$ ve $\text{Var}(X) = 9$ olsun. $P(|X - 5| \geq 6)$ için üst sınır verin.
 B3. Jensen eşitsizliğini $\phi(x) = x^2$ ile kullanarak $\text{Var}(X) \geq 0$ sonucunu gösterin.
 B4. Boole eşitsizliğini yazın ve olasılıkları 0.1, 0.2 ve 0.05 olan üç olaya uygulayın.
 B5. Olasılıkta yakınsamayı tanımlayın.

Orta Düzey

- I1. $P(|X| \geq a) \leq E|X|^p/a^p$ moment eşitsizliğini ispatlayın.
 I2. L^2 yakınsamanın L^1 yakınsamayı gerektirdiğini gösterin.
 I3. Sonlu varyanslı bağımsız özdeş dağılımlı değişkenler için zayıf büyük sayılar yasasını Çebişev eşitsizliğiyle ispatlayın.
 I4. $P(A_n) = 1/n^3$ olsun. Borel–Cantelli ile uzun dönem davranışı yorumlayın.
 I5. Standart normal değişken için Chernoff sınırını türetin.

İleri Düzey

- A1. Cantelli eşitsizliğini $t > 0$ üzerinde en iyileyerek ispatlayın.
 A2. Olasılıkta yakınsamanın dağılımda yakınsamayı gerektirdiğini gösterin.
 A3. Dağılımda yakınsamanın olduğu fakat olasılıkta yakınsamanın olmadığı bir örnek verin.
 A4. Slutsky teoremini toplamalar için sürekli dönüşüm teoremiyle ispatlayın.

A5. Bernoulli denemelerinde $P(|\hat{p} - p| \geq 0.01) \leq 0.05$ için yeterli bir n değerini Hoeffding eşitsizliğiyle bulun.

10.18 Kaynak Notu

Bu bölümdeki Markov, Çebişev, Jensen, Cauchy–Schwarz ve Boole eşitsizlikleri ile yakınsama türleri ve Borel–Cantelli lemması; Billingsley (1995), Grimmett ve Stirzaker (2020), Ross (2019), Hogg, McKean ve Craig (2019), Casella ve Berger (2002) ve Shao (2003) gibi klasik olasılık ve matematiksel istatistik kaynaklarında kullanılan standart çerçeveye uyumludur. Chernoff sınırı, Hoeffding eşitsizliği ve yoğunlaşma araçları için Vershynin (2018); Hölder ve tek taraflı Çebişev için Billingsley (1995) ile Grimmett ve Stirzaker (2020); Bonferroni eşitsizlikleri için Casella ve Berger (2002) izlenmiştir.

Özet:

Olasılık eşitsizlikleri, dağılım tam bilinmediğinde güvenli sınırlar verir. Markov eşitsizliği negatif olmayan değişkenlerde beklenen değerle, Çebişev eşitsizliği ise varyansla kuyruk olasılıklarını kontrol eder. Jensen konvekslik ve beklenti ilişkisini, Cauchy–Schwarz ikinci moment sınırlarını, Boole birleşim olasılıklarını düzenler. Yakınsama kavramları rassal değişken dizilerinin uzun dönem davranışını açıklar; bu kavramlar büyük sayılar yasası ve merkezi limit teoreminin temel dilidir.

Bölüm 11

Simülasyon ve Monte Carlo Yöntemleri

11.1 Gerçek Dünya Problemi (Giriş)

Bir finansal portföyün bir ay sonraki zarar dağılımını, karmaşık bir üretim hattındaki bekleme süresini veya bir ağ sisteminin çökme olasılığını kapalı formülle hesaplamak çoğu zaman mümkün değildir. Model bilinse bile ilgili integral çok boyutlu olabilir, olasılık bölgesi karmaşık olabilir veya sistemin dinamikleri analitik çözümü engelleyebilir.

Monte Carlo yöntemi bu noktada olasılık teorisini hesaplanabilir hâle getirir. Temel fikir basittir: Modelden çok sayıda rassal senaryo üret, her senaryoda ilgilenilen çıktıyı hesapla ve bu çıktıların ortalamasıyla hedef olasılık veya beklenti tahmin et. Büyük sayılar yasası tahminin tutarlılığını, merkezi limit teoremi ise hata payını açıklar.

Bu Bölüm Ne Sağlar?

Bu bölüm sözde rassal sayı üretimi, düzgün değişkenden dağılım üretme, ters dönüşüm yöntemi, kesikli dağılım simülasyonu, kabul-red yöntemi, Monte Carlo integrasyonu, olay olasılığı tahmini, simülasyon standart hatası, güven aralığı, varyans azaltma fikirleri ve simülasyon raporlama ilkelerini ele alır. Önceki bölümlerdeki dağılım, örnekleme, yakınsama ve eşitsizlik bilgisi burada hesaplama yöntemine dönüşür.

Amaç ve Öğrenme Çıktıları

Bu bölümün sonunda okuyucu

- Monte Carlo tahmin edicisini beklenti ve olasılık tahmini bağlamında kurabilir,
- düzgün rassal sayılardan hedef dağılım üretme fikrini açıklayabilir,
- ters dönüşüm ve kabul-red yöntemlerinin temel adımlarını uygulayabilir,
- Monte Carlo standart hatasını ve yaklaşık güven aralığını hesaplayabilir,
- simülasyon sayısı ile hata payı arasındaki karekök ilişkisini yorumlayabilir,
- simülasyon sonuçlarını tekrarlanabilir ve raporlanabilir biçimde sunabilir.

11.2 Monte Carlo Düşüncesi

Monte Carlo yöntemi, hesaplanması zor bir niceliği beklenen değer biçiminde yazar ve bu beklenen değeri örneklem ortalamasıyla yaklaşıklar. Bu yaklaşım, olasılık teorisinin sayısal analizle birleştiği temel noktalardan biridir.

Tanım 11.1 (Monte Carlo Tahmini) $X_1, \dots, X_n$ bağımsız ve aynı dağılımlı, $E[g(X)] < \infty$ olsun. Hesaplanmak istenen değer

$$\theta = E[g(X)]$$

ise Monte Carlo tahmin edicisi

$$\hat{\theta}_n = \frac{1}{n} \sum_{i=1}^n g(X_i)$$

şeklindedir.

Teorem 11.1 (Monte Carlo Tutarlılığı). $Y_i = g(X_i)$ değişkenleri bağımsız ve aynı dağılımlı, $E[Y_i] < \infty$ ise

$$\hat{\theta}_n = \frac{1}{n} \sum_{i=1}^n Y_i \xrightarrow{p} E[Y_1] = \theta.$$

Kanıt. Bu sonuç zayıf büyük sayılar yasasının doğrudan uygulamasıdır. Simülasyon tekrarları bağımsız ve aynı dağılımlı kurulduğunda örneklem ortalaması hedef beklenen değere olasılıkta yakınsar. $\square$

Teorem 11.2 (Monte Carlo Tahmininin Yansızlığı). $Y_i = g(X_i)$ için $E[Y_i] < \infty$ ise Monte Carlo tahmin edicisi yansızdır:

$$E[\hat{\theta}_n] = \theta.$$

Kanıt. Beklentinin doğrusallığıyla

$$E[\hat{\theta}_n] = E\left[\frac{1}{n} \sum_{i=1}^n Y_i\right] = \frac{1}{n} \sum_{i=1}^n E[Y_i] = \frac{1}{n} \cdot n\theta = \theta.$$

Dolayısıyla tahmin, her sonlu n için doğru hedefin etrafında merkezlenmiştir; tutarlılıkla birlikte bu, Monte Carlo'nun temel güvencesidir. $\square$

Örnek 11.1 (İntegrali Beklentiye Çevirme). $U \sim \text{Uniform}(0, 1)$ ise

$$\int_0^1 g(u) du = E[g(U)]$$

olur. Bu nedenle $U_1, \dots, U_n$ bağımsız düzgün değişkenleri için

$$\frac{1}{n} \sum_{i=1}^n g(U_i)$$

integralin Monte Carlo tahminidir.

$\int_0^1 x^2 dx$ integralini Monte Carlo ile tahmin etmek için $U_i \sim \text{Uniform}(0, 1)$ üretip U_i^2 değerlerinin ortalaması alınır. Teorik değer $1/3$ 'tür; simülasyon sayısı arttıkça tahmin bu değere yaklaşır.

11.3 Sözcü Rassal Sayılar ve Tekrarlanabilirlik

Bilgisayarlar gerçek anlamda rastgele değil, deterministik algoritmalarla sözcü rassal sayılar üretir. Bu sayı dizileri uygun testlerden geçtiğinde rassal gibi davranır. Aynı başlangıç değeri, yani tohum değeri, aynı dizinin tekrar üretilmesini sağlar.

Tanım 11.2 (Sözcü Rassal Sayı) Bir algoritma tarafından üretilen ve rassal sayı gibi kullanılan sayılara sözcü rassal sayı denir. Monte Carlo simülasyonlarında temel kaynak çoğunlukla $(0, 1)$ aralığında düzgün sözcü rassal sayılardır.

Tanım 11.3 (Tohum Değeri) Sözcü rassal sayı üreticinin başlangıç durumunu belirleyen değere tohum denir. Tohum sabitlenirse simülasyon tekrarlanabilir hâle gelir.

Bilimsel raporlamada simülasyonun tekrarlanabilir olması önemlidir. Kullanılan dağılımlar, parametreler, simülasyon sayısı, tohum değeri ve yazılım sürümü mümkün olduğunca açık belirtilmelidir.

11.4 Kesikli Dağılım Simülasyonu

Kesikli bir dağılımdan örnek üretmek için $[0, 1]$ aralığı olasılık kütlelerine göre parçalara ayrılır. Düzgün sayı hangi parçaya düşerse ilgili değer seçilir.

Tanım 11.4 (Kümülatif Aralık Yöntemi) X değişkeni $x_1, \dots, x_m$ değerlerini $p_1, \dots, p_m$ olasılıklarıyla alsın. $c_j = p_1 + \dots + p_j$ tanımlansın. $U \sim \text{Uniform}(0, 1)$ için $c_{j-1} < U \leq c_j$ ise $X = x_j$ atanır.

Örnek 11.2 (Hileli Zar Simülasyonu). *Bir zarın 1, 2, 3, 4, 5, 6 değerleri için olasılıkları eşit değilse, bu olasılıkların kümülatif toplamı kullanılarak düzgün sayıdan zar sonucu üretilebilir. Örneğin U ilk olasılık aralığına düşerse sonuç 1, ikinci aralığa düşerse sonuç 2 seçilir.*

11.5 Ters Dönüşüm Yöntemi

Ters dönüşüm yöntemi, düzgün rassal değişkenden hedef dağılıma sahip değişken üretmenin temel yoludur. Dağılım fonksiyonunun tersinin hesaplanabildiği durumlarda doğrudan ve güçlü bir yöntemdir.

Teorem 11.3 (Ters Dönüşüm Yöntemi). $U \sim \text{Uniform}(0, 1)$ ve F sürekli, artan bir dağılım fonksiyonu olsun. Bu durumda

$$X = F^{-1}(U)$$

değişkeninin dağılım fonksiyonu F olur.

Kanıt. Her x için

$$P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x)$$

olur. Son eşitlik U değişkeninin $(0, 1)$ aralığında düzgün dağılmasından gelir. $\square$

Örnek 11.3 (Üstel Dağılım Üretme). $X \sim \text{Exp}(\lambda)$ için dağılım fonksiyonu $F(x) = 1 - e^{-\lambda x}$ olur. $U = F(x)$ çözülürse

$$X = -\frac{1}{\lambda} \log(1 - U)$$

elde edilir. $1 - U$ da düzgün dağıldığından çoğu zaman $X = -(1/\lambda) \log U$ kullanılır.

Ters dönüşüm sürekli durumla sınırlı değildir; dağılım fonksiyonunun genelleştirilmiş tersi kullanılarak kesikli dağılımlara da uygulanır. Kümülatif aralık yöntemi aslında bu genel ilkenin özel hâlidir.

Teorem 11.4 (Genelleştirilmiş Ters Dönüşüm). *Herhangi bir dağılım fonksiyonu F için genelleştirilmiş ters $F^{-1}(u) = \inf\{x : F(x) \geq u\}$ tanımlansın. $U \sim \text{Uniform}(0, 1)$ ise $X = F^{-1}(U)$ değişkeninin dağılım fonksiyonu F olur.*

Kanıt. Genelleştirilmiş tersin temel özelliği $F^{-1}(u) \leq x \iff u \leq F(x)$ olmasıdır. Buradan $P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x)$ elde edilir. Bu, sürekli ve kesikli tüm dağılımlar için geçerlidir. $\square$

Örnek 11.4 (Geometrik Dağılım Üretme). *Geom(p) için F^{-1} kapalı biçimde yazılabilir: $X = \lceil \log(1 - U) / \log(1 - p) \rceil$ değişkeni geometrik dağılımlıdır. Ters dönüşüm, kümülatif aralıkları tek tek taramaya gerek bırakmadan doğrudan değer üretir.*

Örnek 11.5 (Normal Üretme: Box–Muller). *Normal dağılımın F^{-1} 'i kapalı biçimde olmadığından özel bir dönüşüm kullanılır. Bağımsız $U_1, U_2 \sim \text{Uniform}(0, 1)$ için*

$$Z_1 = \sqrt{-2 \log U_1} \cos(2\pi U_2), \quad Z_2 = \sqrt{-2 \log U_1} \sin(2\pi U_2)$$

bağımsız iki standart normal değişken verir. Bu, kutupsal koordinat dönüşümünün simülasyondaki klasik uygulamasıdır.

11.6 Kabul-Red Yöntemi

Bazı dağılımlarda ters dağılım fonksiyonu kolay hesaplanamaz. Kabul-red yöntemi, kolay örneklenen bir öneri dağılımından aday üretir ve hedef yoğunluğa göre adayları kabul eder veya reddeder.

Teorem 11.5 (Kabul-Red İlkesi). *Hedef yoğunluk f , öneri yoğunluğu g ve $f(x) \leq cg(x)$ olacak biçimde $c \geq 1$ olsun. $Y \sim g$ ve $U \sim \text{Uniform}(0, 1)$ bağımsız üretilsin. Eğer*

$$U \leq \frac{f(Y)}{cg(Y)}$$

ise Y kabul edilir. Kabul edilen değerlerin dağılımı f yoğunluğuna sahiptir.

Kanıt. Kabul olayı altında Y değişkeninin yoğunluğu öneri yoğunluğu $g(y)$ ile kabul olasılığı $f(y)/(cg(y))$ çarpımıyla orantılıdır. Bu çarpım $f(y)/c$ olur. Normalize edildiğinde kabul edilen değişkenin yoğunluğu f elde edilir. $\square$

Örnek 11.6 (Karmaşık Yoğunluktan Örnekleme). *Kapalı formda ters fonksiyonu olmayan bir yoğunluk için, onu üstten saran daha basit bir yoğunluk seçilir. Adaylar kolay dağılımdan üretilir; hedef yoğunluğun göreceli yüksek olduğu bölgelerde kabul olasılığı artar.*

Kabul-red yönteminin verimliliği, üstten sarma sabiti c ile doğrudan ilişkilidir. c ne kadar küçükse (öneri hedefe ne kadar iyi uyuyorsa) o kadar az aday reddedilir.

Teorem 11.6 (Kabul Olasılığı ve Beklenen Deneme Sayısı). *Kabul-red yönteminde her bir adayın kabul olasılığı $1/c$ 'dir. Bir kabul edilen değer üretmek için gereken beklenen deneme sayısı c olur.*

Kanıt. Toplam kabul olasılığı, koşullama ile $\int \frac{f(y)}{cg(y)}g(y) dy = \frac{1}{c} \int f(y) dy = \frac{1}{c}$ 'dir. Her deneme bağımsız ve kabul olasılığı $1/c$ olduğundan, ilk kabule kadar deneme sayısı geometrik dağılır; beklenen değeri $1/(1/c) = c$ olur. $\square$

Örnek 11.7 (Verimlilik Yorumu). $c = 1.2$ ise adayların yaklaşık %83'ü ($1/1.2$) kabul edilir ve bir örnek için ortalama 1.2 deneme yeterlidir. $c = 5$ ise her örnek için ortalama 5 deneme gerekir; bu, öneri dağılımının kötü seçildiğini gösterir. İyi bir öneri dağılımı c 'yi 1'e yaklaştırır.

11.7 Olay Olasılığı Tahmini

Monte Carlo yöntemiyle bir olayın olasılığı, gösterge değişkenlerinin ortalamasıyla tahmin edilir. Bu yaklaşım özellikle karmaşık bölgelerde veya çok boyutlu olaylarda kullanışlıdır.

Teorem 11.7 (Gösterge Ortalamasıyla Olasılık Tahmini). *A bir olay ve I_i simülasyonun i 'inci tekrarında olay gerçekleşirse 1, aksi halde 0 alan gösterge değişkeni olsun. O zaman*

$$\hat{p}_n = \frac{1}{n} \sum_{i=1}^n I_i$$

$P(A)$ için tutarlı bir Monte Carlo tahminidir.

Kanıt. $E[I_i] = P(A)$ olduğundan büyük sayılar yasası $\hat{p}_n \xrightarrow{P} P(A)$ sonucunu verir. $\square$

Örnek 11.8 (Sistem Çökme Olasılığı). *Bir ağ sisteminde simülasyon tekrarı sonunda sistem çökerse $I_i = 1$, çökmezse $I_i = 0$ yazılsın. 10000 tekrarda 137 çökme gözlenirse çökme olasılığı tahmini $137/10000 = 0.0137$ olur.*

11.8 Monte Carlo Standart Hatası

Simülasyon çıktısı da rassaldır. Bu nedenle yalnız tahmin değerini vermek yeterli değildir; tahminin Monte Carlo belirsizliği de raporlanmalıdır.

Teorem 11.8 (Monte Carlo Hatasının Varyansı). $Y_1, \dots, Y_n$ bağımsız ve aynı dağılımlı, $E[Y_i] = \mu$ ve $\text{Var}(Y_i) = \sigma^2 < \infty$ ise

$$\text{Var}\left(\frac{1}{n} \sum_{i=1}^n Y_i\right) = \frac{\sigma^2}{n}.$$

Kanıt. Bağımsızlık nedeniyle toplamın varyansı varyansların toplamıdır. $1/n$ katsayısı varyansı $1/n^2$ ile çarpar; sonuç $n\sigma^2/n^2 = \sigma^2/n$ olur. $\square$

Tanım 11.5 (Monte Carlo Standart Hatası) Simülasyon çıktıları $Y_1, \dots, Y_n$ ve örnek standart sapması S olsun. Ortalama tahmininin Monte Carlo standart hatası yaklaşık

$$SE(\bar{Y}_n) = \frac{S}{\sqrt{n}}$$

şeklindedir.

Yorum:

Standart hata $1/\sqrt{n}$ hızında azalır. Hata payını yarıya indirmek için simülasyon sayısını yaklaşık dört katına çıkarmak gerekir.

11.9 Monte Carlo Güven Aralığı

Merkezi limit teoremi, bağımsız simülasyon çıktılarının ortalaması için yaklaşık normal hata modeli sağlar. Bu sayede Monte Carlo tahmini için yaklaşık güven aralığı kurulabilir.

Teorem 11.9 (Monte Carlo Güven Aralığı). *Bağımsız simülasyon çıktıları sonlu varyanslıysa, büyük n için*

$$\bar{Y}_n \pm z_{\alpha/2} \frac{S}{\sqrt{n}}$$

ifadesi $\mu = E[Y]$ için yaklaşık güven aralığı verir.

Kanıt. Merkezi limit teoremi $\sqrt{n}(\bar{Y}_n - \mu)/\sigma$ değişkeninin yaklaşık standart normal olduğunu söyler. Bilinmeyen σ yerine tutarlı örnek standart sapması S kullanıldığında Slutsky teoremiyle aynı büyük örneklem yaklaşımı korunur. $\square$

Örnek 11.9 (Olasılık Tahmini İçin Hata Payı). *Bir olayın olasılığı $\hat{p} = 0.20$ ve simülasyon sayısı $n = 10000$ olsun. Gösterge değişkeni için yaklaşık standart hata*

$$\sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{0.20 \cdot 0.80}{10000}}$$

olarak hesaplanır.

11.10 Monte Carlo İntegrasyonu

Monte Carlo integrasyonu özellikle yüksek boyutlu integrallerde kullanışlıdır. Klasik grid yöntemleri boyut arttıkça hızla zorlaşırken Monte Carlo hata hızı boyuttan ziyade varyansa ve örneklem sayısına bağlıdır.

Teorem 11.10 (Birim Aralıkta Monte Carlo İntegrasyonu). $U_i \sim \text{Uniform}(0, 1)$ bağımsız ve $E[g(U)^2] < \infty$ olsun. O halde

$$\hat{I}_n = \frac{1}{n} \sum_{i=1}^n g(U_i)$$

$$I = \int_0^1 g(u) du$$

için tutarlı bir tahmindir ve büyük n için standart hatası $S_g/\sqrt{n}$ ile tahmin edilir.

Kanıt. Düzgün dağılım altında $E[g(U)] = \int_0^1 g(u) du$ olur. Tutarlılık büyük sayılar yasasından, standart hata sonucu ise örneklem ortalamasının varyansından gelir. $\square$

Örnek 11.10 (π Sayısını Tahmin Etme). *Birim karede düzgün (U, V) noktaları üretelim. $U^2 + V^2 \leq 1$ ise nokta çeyrek dairenin içindedir. Bu olayın olasılığı $\pi/4$ olduğundan*

$$\hat{\pi} = 4 \cdot \frac{1}{n} \sum_{i=1}^n I(U_i^2 + V_i^2 \leq 1)$$

Monte Carlo tahminidir.

Teorem 11.11 (Antitetik Değişkenlerde Varyans). $\hat{\theta}_A = \frac{1}{2}\{g(U) + g(1 - U)\}$ olsun. Bu durumda

$$\text{Var}(\hat{\theta}_A) = \frac{1}{4}(\text{Var}(g(U)) + \text{Var}(g(1 - U)) + 2\text{Cov}(g(U), g(1 - U)))$$

olur. g monoton ise $\text{Cov}(g(U), g(1 - U)) \leq 0$ olduğundan varyans bağımsız çift kullanmaya göre azalır.

Kanıt. Varyans açılımı toplamın varyans formülünden gelir. Monoton g için U büyüdükçe $g(U)$ artar ama $g(1 - U)$ azalır; bu ters yönlü hareket negatif kovaryans üretir. Negatif kovaryans terimi toplam varyansı düşürür. $\square$

Teorem 11.12 (Kontrol Değişkeni). $E[h(X)] = m$ bilinen bir kontrol değişkeni olsun. Her sabit b için

$$\hat{\theta}_b = \frac{1}{n} \sum_{i=1}^n (g(X_i) - b[h(X_i) - m])$$

yansız bir tahmindir. Varyansı en küçük yapan katsayı $b^* = \text{Cov}(g(X), h(X)) / \text{Var}(h(X))$ 'dir.

Kanıt. $E[h(X) - m] = 0$ olduğundan her b için $E[\hat{\theta}_b] = E[g(X)] = \theta$; yani yansızdır. Tek terimin varyansı $\text{Var}(g) - 2b \text{Cov}(g, h) + b^2 \text{Var}(h)$ ifadesidir; b^* 'ye göre türev sıfırlanınca b^* elde edilir. Bu seçimde varyans $\text{Var}(g)(1 - \rho^2)$ olur; ρ büyüdükçe kazanç artar. $\square$

Teorem 11.13 (Önem Örnekleme). Hedef $\theta = E_f[g(X)] = \int g(x)f(x)dx$ olsun. Pozitif bir öneri yoğunluğu q için

$$\theta = E_q \left[g(X) \frac{f(X)}{q(X)} \right]$$

yazılır; dolayısıyla $X_i \sim q$ ile $\hat{\theta} = \frac{1}{n} \sum_i g(X_i) w(X_i)$, $w = f/q$ yansız bir tahmindir.

Kanıt. $E_q[g(X)f(X)/q(X)] = \int g(x) \frac{f(x)}{q(x)} q(x) dx = \int g(x)f(x) dx = \theta$ olur. q , integrandın büyük olduğu bölgelere ağırlık verecek biçimde seçilirse varyans önemli ölçüde azalır; nadir olay olasılıklarının tahmininde temel araçtır. $\square$

11.12 Rassal Süreç Simülasyonuna Giriş

Monte Carlo yalnız tek değişken üretmek için değil, zaman içinde gelişen rassal süreçleri simüle etmek için de kullanılır. Poisson süreçleri, rastgele yürüyüşler ve Markov zincirleri bunun temel örnekleridir.

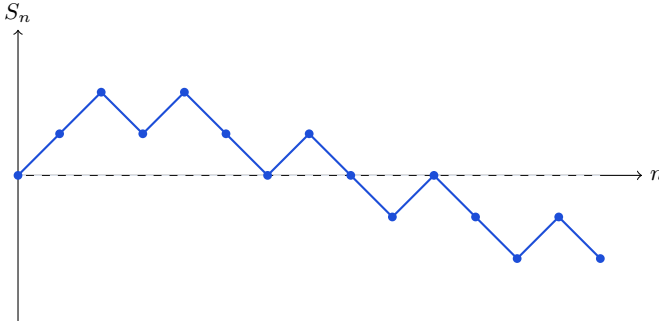
Tanım 11.6 (Rassal Süreç) Bir indeks kümesi T için $\{X(t) : t \in T\}$ biçimindeki rassal değişken ailesine rassal süreç denir. T zaman olarak yorumlandığında süreç, zaman içinde değişen belirsiz sistemi temsil eder.

Tanım 11.7 (Basit Rastgele Yürüyüş) $S_0 = 0$ ve $S_n = X_1 + \cdots + X_n$ olsun. X_i değişkenleri bağımsız olarak $+1$ veya -1 değerlerini alıyorsa S_n basit rastgele yürüyüştür.

Teorem 11.14 (Rastgele Yürüyüşte Ortalama ve Varyans). $P(X_i = 1) = p$, $P(X_i = -1) = q = 1 - p$ ise

$$E[S_n] = n(p - q), \quad \text{Var}(S_n) = 4npq.$$

Kanıt. $E[X_i] = p - q$ ve $E[X_i^2] = 1$ olduğundan $\text{Var}(X_i) = 1 - (p - q)^2 = 4pq$ olur. Bağımsız toplam için beklentiler ve varyanslar toplanır. $\square$



Şekil 11.2 Basit rastgele yürüyüşün bir gerçeğemesi: her adımda ± 1 hareket eder. Beklenen konum $n(p - q)$, yayılım $\sqrt{4npq}$ ile büyür.

Poisson süreçleri, olaylar arası bekleme sürelerinin bağımsız üstel olması özelliğiyle kolayca simüle edilir. Bu, ters dönüşümün rassal süreçlere doğal bir uygulamasıdır (Ross, 2019).

Teorem 11.15 (Poisson Sürecinin Üretimi). *Hızı λ olan bir Poisson sürecinde ardışık olaylar arası süreler bağımsız $\text{Exp}(\lambda)$ dağılımlıdır. Dolayısıyla $T_k = \sum_{i=1}^k E_i$, $E_i \sim \text{Exp}(\lambda)$ üretilerek olay zamanları elde edilir.*

Kanıt. Poisson sürecinin tanımlayıcı özelliği, ayrık aralıklardaki olay sayılarının bağımsız Poisson olması ve aralar arası sürelerin hafızasız olmasıdır. Hafızasız sürekli dağılım üstel olduğundan, ardışık üstel süreleri toplamak süreci tam olarak üretir. $\square$

Örnek 11.12 (Çağrı Zamanlarını Üretme). *Saatte ortalama $\lambda = 20$ çağrı gelen bir merkez için, $\text{Exp}(20)$ üretilip kümülatif toplam alınarak çağrı zamanları simüle*

edilir. Bir saatlik pencerede beklenen çağrı sayısı 20'dir ve süreç tekrarlanarak yoğunluk profili incelenebilir.

11.13 Simülasyon Sonucunu Raporlama

İyi bir Monte Carlo raporu yalnız sonuç sayısını vermez. Model varsayımlarını, algoritmayı, simülasyon sayısını, tohum değerini, standart hatayı ve varsa güven aralığını da sunar.

1. Hedef niceliği açık tanımla: olasılık mı, beklenti mi, integral mi?
2. Kullanılan dağılımları ve parametreleri yaz.
3. Simülasyon algoritmasını adım adım belirt.
4. Simülasyon sayısı n ve tohum değerini raporla.
5. Tahminle birlikte Monte Carlo standart hatasını ver.
6. Sonucu analitik değer, yaklaşık yöntem veya duyarlılık analiziyle kontrol et.

Yaygın Hata Kalıbı:

Monte Carlo sonucunu yalnız tek sayı olarak vermek eksiktir. Simülasyon çıktısı rassal olduğu için standart hata veya güven aralığı olmadan sonucun sayısal güvenilirliği değerlendirilemez.

11.14 Örnek Veri Tablosu ve Yazılım Çözümleri

Aşağıdaki örnek, $\int_0^1 x^2 dx$ integralini Monte Carlo ile tahmin eder. Teorik değer $1/3$ olduğundan simülasyon sonucu kolayca kontrol edilebilir.

Simülasyon Sayısı	Tahmin	Standart Hata	Teorik Değer
1 000	MC ortalaması	$S/\sqrt{1000}$	1/3
10 000	MC ortalaması	$S/\sqrt{10000}$	1/3

```
set.seed(42)
n <- 10000
u <- runif(n)
y <- u^2
mean(y)
sd(y) / sqrt(n)
```

Python Çözümü

```
import numpy as np

rng = np.random.default_rng(42)
n = 10000
u = rng.uniform(0, 1, size=n)
y = u**2

estimate = np.mean(y)
se = np.std(y, ddof=1) / np.sqrt(n)

print(estimate, se, 1/3)
```

11.15 Görselleştirme ve Grafikler

Monte Carlo tahmininin simülasyon sayısı arttıkça nasıl kararlılaştığı kümülatif ortalama grafiğiyle görülebilir.

```
import numpy as np
import matplotlib.pyplot as plt

rng = np.random.default_rng(42)
n = 5000
u = rng.uniform(0, 1, size=n)
y = u**2
running_mean = np.cumsum(y) / np.arange(1, n + 1)

plt.plot(running_mean, label="Monte Carlo tahmini")
plt.axhline(1/3, color="red", linestyle="--", label="Teorik değer")
plt.xlabel("Simülasyon sayısı")
plt.ylabel("Kümülatif ortalama")
plt.title("Monte Carlo Tahmininin Yakınsaması")
plt.legend()
plt.show()
```

Dikkat: Kavram Yanılgısı

Monte Carlo hatası, kod hatası değildir. Simülasyon sonucu rassal olduğu için aynı model doğru kurulsa bile farklı tohumlarda küçük farklar oluşur; raporda standart hata ve güven aralığı verilmesinin nedeni budur.

Mini-Proje / Vaka Çalışması:

Tekrarlanabilir Monte Carlo raporu: $P(Z > 2)$ olasılığını simülasyonla tahmin edin. Raporunuzda simülasyon sayısını, tohum değerini, tahmini, standart hatayı ve yaklaşık yüzde 95 güven aralığını verin. Sonucu tablo değeriyle karşılaştırın ve farkı Monte Carlo hatasıyla açıklayın.

Kendini Dene:

Ara Kontrol: Simülasyon sayısını 100 kat artırınca standart hatanın yaklaşık kaç kat azalmasını beklersiniz?

11.16 Bölüm Sonu Çözümlü Problemler

1. $U \sim \text{Uniform}(0, 1)$ olsun. $\int_0^1 e^{-u} du$ integralinin Monte Carlo tahmin edicisini yazın.

Çözüm. Hedef integral $E[e^{-U}]$ biçimindedir. Bağımsız $U_1, \dots, U_n$ için

$$\hat{I}_n = \frac{1}{n} \sum_{i=1}^n e^{-U_i}$$

Monte Carlo tahminidir.

2. $Y_1, \dots, Y_n$ simülasyon çıktılarının standart sapması $S = 12$ ve $n = 3600$ olsun. Monte Carlo standart hatasını bulun.

Çözüm.

$$SE = \frac{S}{\sqrt{n}} = \frac{12}{60} = 0.2.$$

3. $F(x) = 1 - e^{-2x}$, $x \geq 0$ dağılımından ters dönüşümle örnek üretme formülünü yazın.

Çözüm. $U = 1 - e^{-2x}$ çözümlürse

$$x = -\frac{1}{2} \log(1 - U)$$

olur. $1 - U$ da düzgün dağıldığından $X = -(1/2) \log U$ da kullanılabilir.

4. Bir olay 20000 simülasyonun 640 tanesinde gerçekleşmiştir. Olay olasılığı tahmini nedir?

Çözüm.

$$\hat{p} = \frac{640}{20000} = 0.032.$$

5. Monte Carlo hata payını yaklaşık yarıya indirmek için simülasyon sayısı nasıl değiştirilmelidir?

Çözüm. Standart hata $1/\sqrt{n}$ hızında azaldığından hata payını yarıya indirmek için n yaklaşık dört katına çıkarılmalıdır.

6. $P(X_i = 1) = 0.6$, $P(X_i = -1) = 0.4$ olan rastgele yürüyüşte S_{100} için beklenen değeri bulun.

Çözüm. $p - q = 0.6 - 0.4 = 0.2$ olduğundan

$$E[S_{100}] = 100(0.2) = 20.$$

7. Kabul-red yönteminde sarma sabiti $c = 2.5$ ise bir örnek üretmek için ortalama kaç deneme gerekir?

Çözüm. Beklenen deneme sayısı $c = 2.5$ 'tir; kabul olasılığı $1/2.5 = 0.4$.

8. Aynı problemde yukarıdaki rastgele yürüyüşte S_{100} 'ün varyansı nedir?

Çözüm. $\text{Var}(S_n) = 4npq = 4(100)(0.6)(0.4) = 96$.

9. $\theta = E[g(U)]$ tahmininde kontrol değişkeniyle korelasyon $\rho = 0.9$ ise varyans hangi oranda azalır?

Çözüm. Varyans $\text{Var}(g)(1 - \rho^2) = \text{Var}(g)(1 - 0.81) = 0.19 \text{Var}(g)$ 'ye iner; yaklaşık %81 azalma.

10. Hızı $\lambda = 20$ olan Poisson süreci nasıl simüle edilir?

Çözüm. Bağımsız $\text{Exp}(20)$ aralar üretilip kümülatif toplam alınır; bu toplamalar olay zamanlarını verir.

11. Box-Muller dönüşümünde $U_1, U_2 \sim \text{Uniform}(0, 1)$ kaç tane standart normal üretir?

Çözüm. İki bağımsız standart normal üretir: $Z_1 = \sqrt{-2 \log U_1} \cos(2\pi U_2)$ ve $Z_2 = \sqrt{-2 \log U_1} \sin(2\pi U_2)$.

11.17 Bölüm Sonu Alıştırımları

- E1. Sabit bir tohumla $P(Z > 1.96)$ değerini simüle edin ve Monte Carlo standart hatasını raporlayın.
- E2. $\int_0^1 \sqrt{x} dx$ integralini Monte Carlo ile tahmin edin ve kesin değerle karşılaştırın.
- E3. Simülasyon büyüklüğünü dört katına çıkarmanın standart hatayı neden yaklaşık yarıya indirdiğini açıklayın.
- E4. Bir varyans azaltma yöntemi tanımlayın ve hangi koşulda yararlı olduğunu belirtin.

11.18 Kaynak Notu

Bu bölümdeki Monte Carlo tahmini, ters dönüşüm, kabul-red, simülasyon standart hatası, güven aralığı, varyans azaltma ve rassal süreç simülasyonu sunumu;

Ross (2019), Robert ve Casella (2004), Fishman (1996), Gentle (2003), Hogg, McKean ve Craig (2019) ve Grimmett ve Stirzaker (2020) gibi olasılık, simülasyon ve Monte Carlo kaynaklarındaki standart yaklaşımla uyumludur. Genelleştirilmiş ters dönüşüm, Box–Muller ve kabul-red verimliliği için Devroye (1986) ile Gentle (2003); kontrol değişkeni, antitetik değişkenler ve önem örnekleme için Robert ve Casella (2004); Poisson süreç üretimi için Ross (2019) izlenmiştir.

Özet:

Monte Carlo yöntemleri, olasılık teorisini hesaplanabilir hâle getirir. Hedef olasılık veya beklenti, bağımsız simülasyon çıktılarının ortalamasıyla tahmin edilir. Tutarlılık büyük sayılar yasasından, hata ölçümü merkezi limit teoremi ve standart hata kavramından gelir. Ters dönüşüm ve kabul-red yöntemleri dağılım üretmenin temel araçlarıdır. İyi bir simülasyon raporu tahminle birlikte standart hata, simülasyon sayısı, tohum değeri ve model varsayımlarını da içermelidir.

Ek A

Ekler

Ek A : Dağılım ve Kritik Değer Tablo- ları

A.1 Standart Normal Dağılım Tablosu

Aşağıdaki tabloda $\Phi(z) = P(Z \leq z)$ değerleri verilmiştir. Satır ilk iki basamağı, sütun ise yüzde birlik kısmı gösterir.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

A.2 t Dağılımı Kritik Değerleri

Aşağıdaki tabloda seçilmiş serbestlik dereceleri için üst kuyruk kritik değerleri verilmiştir; yani $P(T > t_{\alpha, \nu}) = \alpha$.

sd	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.025$	$\alpha = 0.01$	$\alpha = 0.005$
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
8	1.397	1.860	2.306	2.896	3.355
10	1.372	1.812	2.228	2.764	3.169
12	1.356	1.782	2.179	2.681	3.055
15	1.341	1.753	2.131	2.602	2.947
20	1.325	1.725	2.086	2.528	2.845
24	1.318	1.711	2.064	2.492	2.797
30	1.310	1.697	2.042	2.457	2.750
40	1.303	1.684	2.021	2.423	2.704
60	1.296	1.671	2.000	2.390	2.660
120	1.289	1.658	1.980	2.358	2.617

A.3 z Dağılımı Kritik Değerleri

Standart normal dağılım için sık kullanılan tek kuyruk, çift kuyruk ve güven düzeyi kritik değerleri aşağıda özetlenmiştir.

Tek kuyruk α	Çift kuyruk α	Kritik değer z
0.10	0.20	1.282
0.05	0.10	1.645
0.025	0.05	1.960
0.01	0.02	2.326
0.005	0.01	2.576

Güven düzeyi	$z_{\alpha/2}$
80%	1.282
90%	1.645
95%	1.960
98%	2.326
99%	2.576

A.4 F Dağılımı Kritik Değerleri

Aşağıdaki tabloda $\alpha = 0.05$ için üst kuyruk kritik değerleri verilmiştir; yani $P(F > F_{0.05;\nu_1,\nu_2}) = 0.05$.

$\nu_2 \backslash \nu_1$	1	2	3	4	5	10
1	161.448	199.500	215.707	224.583	230.162	241.882
2	18.513	19.000	19.164	19.247	19.296	19.396
3	10.128	9.552	9.277	9.117	9.013	8.786
4	7.709	6.944	6.591	6.388	6.256	5.964
5	6.608	5.786	5.409	5.192	5.050	4.735
10	4.965	4.103	3.708	3.478	3.326	2.978
20	4.351	3.493	3.098	2.866	2.711	2.348
30	4.171	3.316	2.922	2.690	2.534	2.164
60	4.001	3.150	2.758	2.525	2.369	1.998
120	3.920	3.071	2.680	2.447	2.291	1.917

A.5 Ki-Kare Dağılımı Kritik Değerleri

Aşağıdaki tabloda seçilmiş serbestlik dereceleri için ki-kare kritik değerleri verilmiştir. Sütunlar üst kuyruk olasılıklarını göstermektedir; yani $P(\chi^2 > c) = \alpha$.

sd	$\alpha = 0.995$	$\alpha = 0.99$	$\alpha = 0.975$	$\alpha = 0.95$	$\alpha = 0.90$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.025$	$\alpha = 0.01$	$\alpha = 0.005$
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	40.000
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
30	13.787	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672

Ek B

Cevap Anahtarı

B.1 Bölüm 1: Olasılığın Temelleri ve Aksiyomatik Yapı

Temel Düzey

- B1. $\Omega = \{YY, YT, TY, TT\}$; en az bir yazı olayı $\{YY, YT, TY\}$ biçimindedir.
B2. $A \cup B = \{1, 3, 4, 5, 6\}$, $A \cap B = \{5\}$ ve $A^c = \{2, 4, 6\}$.
B3. İçerme-dışlama ile $P(A \cup B) = 0.35 + 0.50 - 0.20 = 0.65$.
B4. $B = A \cup (B \cap A^c)$ ayrık birleşim olduğundan $P(B \cap A^c) = 0.62 - 0.28 = 0.34$.
B5. En az bir dersi alma olasılığı $0.40 + 0.25 - 0.10 = 0.55$.

Orta Düzey

- I1. $P(\text{iki beyaz}) = (4/10)(3/9) = 2/15$.
I2. $P(A \cap B) = 0.6 + 0.5 - 0.8 = 0.3$ ve $P(A)P(B) = 0.30$; dolayısıyla olaylar bağımsızdır.
I3. Bayes teoremiyle

$$P(H | +) = \frac{0.92(0.03)}{0.92(0.03) + 0.06(0.97)} \approx 0.322.$$

- I4. İçerme-dışlama ile $0.40 + 0.35 + 0.30 - 0.12 - 0.10 - 0.08 + 0.03 = 0.78$.
I5. Paralel sistemin çalışmama olasılığı $(0.10)(0.20) = 0.02$ olduğundan güvenirlilik 0.98 'dir.

İleri Düzey

- A1. $B_1 = A_1$ ve $B_n = A_n \setminus A_{n-1}$ tanımlanırsa B_n 'ler ayrıktır ve $A = \bigcup_{n \geq 1} B_n$ olur. Sayılabilir toplamsallık $P(A) = \sum_{n \geq 1} P(B_n) = \lim_n P(A_n)$ sonucunu verir.
- A2. Ayrıklık nedeniyle $P(A \cap B) = 0$ olur. Bağımsızlık olsaydı $P(A \cap B) = P(A)P(B) > 0$ gerekirdi; bu çelişkidir.
- A3. $P(\text{çakışma}) = 1 - \prod_{j=0}^{23} (1 - j/365) \approx 0.538$.
- A4. Önsel oran $0.01/0.99$ 'dur. Sonsal oran $20(0.01/0.99) = 20/99$ ve sonsal olasılık $(20/99)/(1 + 20/99) = 20/119 \approx 0.168$.
- A5. Boole eşitsizliği $P(\bigcup_{i=1}^{250} A_i) \leq 250(0.002) = 0.5$ verir. Bu üst sınır bağımsızlık varsayımı gerektirmez.

B.2 Bölüm 2: Kombinatorik Analiz ve Sayma Yöntemleri

Temel Düzey

- B1. Sıra önemli olduğundan $P(7, 3) = 7 \cdot 6 \cdot 5 = 210$.
- B2. Sıra önemli değildir: $\binom{12}{5} = 792$.
- B3. Her basamak için 10 seçenek vardır; toplam $10^4 = 10000$ şifre oluşur.
- B4. A harfi 3, N harfi 2 kez geçtiğinden $6!/(3!2!) = 60$.
- B5. Çarpma kuralıyla $3 \cdot 5 \cdot 4 = 60$ farklı menü oluşturulur.

Orta Düzey

- I1. $\frac{\binom{13}{3}\binom{39}{2}}{\binom{52}{5}}$.
- I2. Önce ekip seçilip sonra kaptan seçilir: $\binom{9}{4} \cdot 4 = 504$.
- I3. $x_1 + x_2 + x_3 = 8$ ve $x_i \geq 1$ için çözüm sayısı $\binom{7}{2} = 21$.
- I4. 5 sağ ve 4 yukarı adımın sıralanması gerekir: $\binom{9}{5} = 126$.
- I5. $\frac{\binom{2}{1}\binom{8}{2}}{\binom{10}{3}} = \frac{56}{120} = \frac{7}{15}$.

İleri Düzey

- A1. İki belirli kişi blok kabul edilirse 7 nesne dairesel dizilir ve blok içi 2 sıralama vardır: $(7 - 1)! \cdot 2 = 1440$.

- A2. Örtten fonksiyon sayısı $3^7 - \binom{3}{1}2^7 + \binom{3}{2}1^7 = 2187 - 384 + 3 = 1806$.
- A3. Multinom katsayısı $\frac{6!}{2!3!1!} = 60$.
- A4. İki belirli kişi komitedeyse kalan 4 kişi diğer 18 kişiden seçilir. Olasılık $\binom{18}{4} / \binom{20}{6} = 1/19$.
- A5. n elemanlı bir kümede işaretli bir eleman içeren alt kümeleri iki şekilde sayalım. Sol taraf, alt küme büyüklüğü k iken işaretli elemanı seçer; sağ taraf önce işaretli elemanı n yolla seçer, kalan $n - 1$ elemanın alt kümeye girip girmemesini 2^{n-1} yolla belirler.

B.3 Bölüm 3: Rassal Değişkenler ve Dağılımlar

Temel Düzey

- B1. $P(X = 0) = 0.3$, $E[X] = 0.7$ ve $\text{Var}(X) = 0.7(0.3) = 0.21$.
- B2. $E[X] = 0(0.2) + 1(0.5) + 2(0.3) = 1.1$.
- B3. $P(1 < X \leq 4) = F(4) - F(1) = 0.80 - 0.25 = 0.55$.
- B4. $P(X = 0) = e^{-2}$ ve $P(X = 1) = 2e^{-2}$.
- B5. $E[Y] = 2E[X] + 1 = 9$ ve $\text{Var}(Y) = 2^2 \text{Var}(X) = 36$.

Orta Düzey

- I1. $P(X = 3) = \binom{10}{3}(0.4)^3(0.6)^7 \approx 0.215$.
- I2. $P(X > 3) = (0.75)^3 \approx 0.422$ ve $E[X] = 1/0.25 = 4$.
- I3. $\frac{\binom{5}{1}\binom{15}{3}}{\binom{20}{4}}$.
- I4. $E[X] = 1(0.2) + 3(0.3) + 5(0.5) = 3.6$ ve $E[X^2] = 1(0.2) + 9(0.3) + 25(0.5) = 15.4$ olduğundan $\text{Var}(X) = 15.4 - 3.6^2 = 2.44$.
- I5. Toplam $X + Y \sim \text{Poisson}(4)$ olur; $P(X + Y = 2) = e^{-4}4^2/2! = 8e^{-4} \approx 0.147$.

İleri Düzey

- A1. Kuyruk toplamıyla $E[X] = \sum_{k \geq 1} P(X \geq k) = \sum_{k \geq 1} (0.6)^k = 0.6/0.4 = 1.5$.
- A2. $X = I_1 + \dots + I_n$ yazılır; burada I_j başarı göstergesidir. $E[I_j] = p$ olduğundan doğrusallıkla $E[X] = \sum_j E[I_j] = np$.
- A3. $P(X = 2) = 0.3$ ve $E[X] = G'(1) = 0.5 + 0.6 = 1.1$.

- A4. $\frac{P(X+1)}{P(X)} = \frac{\lambda}{X+1}$. Oran 1'den küçük olduğunda, yani $X+1 > \lambda$ iken olasılıklar azalmaya başlar.
- A5. $P(X > m+n \mid X > m) = \frac{(1-p)^{m+n}}{(1-p)^m} = (1-p)^n = P(X > n)$.

B.4 Bölüm 4: Sürekli Dağılımlar ve Dönüşümler

Temel Düzey

- B1. Düzgün dağılımda uzunluk oranı kullanılır: $P(3 < X < 6) = (6-3)/(8-2) = 1/2$.
- B2. $P(X > 4) = e^{-0.5 \cdot 4} = e^{-2} \approx 0.135$.
- B3. $P(X > 0.5) = \int_{0.5}^1 2x \, dx = 1 - 0.25 = 0.75$.
- B4. Standart normal tablodan $P(Z \leq 1.96) \approx 0.975$.
- B5. Ters dönüşüm $x = (y-1)/2$ ve Jacobian $1/2$ olduğundan $f_Y(y) = \frac{1}{2}f_X((y-1)/2)$.

Orta Düzey

- I1. $F(q) = 1 - e^{-\lambda q} = 0.90$ olduğundan $q_{0.90} = \log(10)/\lambda$.
- I2. $F_Y(y) = P(X^3 \leq y) = P(X \leq y^{1/3}) = (y^{1/3})^3 = y$; dolayısıyla $Y \sim U(0, 1)$.
- I3. $P(Y \leq y) = P(-\log X/\lambda \leq y) = P(X \geq e^{-\lambda y}) = 1 - e^{-\lambda y}$, $y \geq 0$; yani $Y \sim \text{Exp}(\lambda)$.
- I4. Evrişimle $f_S(s) = \int_0^s 1 \, dx = s$; $0 < s < 1$.
- I5. Standartlaştırma sonucu $Z \sim N(0, 1)$.

İleri Düzey

- A1. $y \in (0, 1)$ için ters kollar $\pm\sqrt{y}$ ve $|dx/dy| = 1/(2\sqrt{y})$ 'dir. Bu nedenle $f_Y(y) = \frac{1}{2} \frac{1}{2\sqrt{y}} + \frac{1}{2} \frac{1}{2\sqrt{y}} = \frac{1}{2\sqrt{y}}$.
- A2. $h(x) = f(x)/S(x) = -S'(x)/S(x) = 3x^2$.
- A3. $M(t) = \lambda/(\lambda - t)$, $t < \lambda$. $M'(0) = 1/\lambda$, $M''(0) = 2/\lambda^2$ ve $\text{Var}(X) = M''(0) - M'(0)^2 = 1/\lambda^2$.
- A4. $f_T(t) = \int_0^t \lambda e^{-\lambda x} \lambda e^{-\lambda(t-x)} \, dx = \lambda^2 t e^{-\lambda t}$, $t > 0$.
- A5. $x = e^y$ ve $dx/dy = e^y$ olduğundan $f_Y(y) = f_X(e^y)e^y$. Ters yönde $X = e^Y$ ve $Y \sim N(\mu, \sigma^2)$ alınırsa lognormal yoğunluğu Jacobian $1/x$ çarpanıyla elde edilir.

B.5 Bölüm 5: Ortak Dağılımlar ve Bağımlılık

Temel Düzey

- B1. $p_X(0) = 0.2 + 0.3 = 0.5$ ve $p_Y(0) = 0.2 + 0.1 = 0.3$.
 B2. $P(X = 1, Y = 1) = 0.4$, $P(X = 1) = 0.5$ ve $P(Y = 1) = 0.7$ olduğundan $P(X = 1)P(Y = 1) = 0.35$. Eşitlik sağlanmadığı için bağımsız değildir.
 B3. $\text{Cov}(X, Y) = E[XY] - E[X]E[Y] = 12 - 2 \cdot 5 = 2$.
 B4. $\text{Var}(X + Y) = 4 + 9 + 2(3) = 19$.
 B5. $\rho = \text{Cov}(X, Y) / (\sigma_X \sigma_Y) = -4 / (2 \cdot 5) = -0.4$.

Orta Düzey

- I1. Bölge alanı 2 olduğu için $c = 1/2$. Olasılık $\int_0^{1/2} \int_1^2 (1/2) dy dx = 1/4$.
 I2. Marjinaller $f_X(x) = \int_0^1 6xy^2 dy = 2x$ ve $f_Y(y) = \int_0^1 6xy^2 dx = 3y^2$ olur. Çarpım $f_X(x)f_Y(y) = 6xy^2$ olduğundan bağımsızdır.
 I3. $P(Y = 1 | X = 1) = p(1, 1) / p_X(1) = 0.4 / (0.3 + 0.4) = 4/7$.
 I4. $\text{Cov}(X, Y) = \rho \sigma_X \sigma_Y = (-0.5)(1)(2) = -1$. Bu nedenle $\text{Var}(2X - Y) = 4 \text{Var}(X) + \text{Var}(Y) - 4 \text{Cov}(X, Y) = 4 + 4 + 4 = 12$.
 I5. Bağımsızlıktan $E[XY] = E[X]E[Y] = 3$ ve $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) = 7$.

İleri Düzey

- A1. Üçgenin alanı $1/2$ olduğundan yoğunluk 2'dir. Sabit y için $0 < x < y$ olduğundan $f_Y(y) = \int_0^y 2 dx = 2y$, $0 < y < 1$.
 A2. $Y = 1$ her zaman olduğundan sabit değişkendir ve X ile bağımsızdır. Ayrıca $\text{Cov}(X, Y) = \text{Cov}(X, 1) = 0$; burada sıfır kovaryans bağımsızlıkla uyumludur çünkü Y sabittir.
 A3. $E[Y | X = 12] = 20 + 0.6(5/2)(12 - 10) = 23$ ve $\text{Var}(Y | X = 12) = 25(1 - 0.36) = 16$.
 A4. $\text{Cov}(N_1, N_3) = -100(0.2)(0.3) = -6$. $\text{Var}(N_1) = 100(0.2)(0.8) = 16$, $\text{Var}(N_3) = 100(0.3)(0.7) = 21$; dolayısıyla $\text{Corr}(N_1, N_3) = -6 / \sqrt{16 \cdot 21} \approx -0.327$.
 A5. $\text{Cov}(U, V) = \text{Cov}(X + Y, X - Y) = \text{Var}(X) - \text{Var}(Y) = 0$. U ve V ortak normal ve kovaryansları sıfır olduğundan bağımsızdır.

B.6 Bölüm 6: Örneklem Teorisi ve Örneklem Dağılımları

Bölüm Sonu Alıştırmaları

- E1. Örneklem dağılımı, $\bar{X}$ veya S^2 gibi bir istatistiğin tekrarlı örneklemeden örnekleme nasıl değiştiğini gösterir. Tek tek ham gözlemlerin anakütle dağılımı değildir.
- E2. $SE(\bar{X}) = \sigma/\sqrt{n} = 5/10 = 0.5$.
- E3. Pratik bir koşul, gözlemlerin bağımsız olması, varyansın sonlu olması ve örneklem büyüklüğünün yeterince büyük olmasıdır. Kuvvetli çarpık veya ağır kuyruklu dağılımlarda daha büyük örneklem gerekir.
- E4. Yanlılık, tahmin edicinin ortalama olarak hedeften sistematik sapmasını; varyans ise örneklemden örnekleme kararsızlığını ölçer. İyi bir yöntem çoğu zaman ikisini birlikte, örneğin MSE ile, kontrol eder.

B.7 Bölüm 7: Asimptotik Teori ve Limit Teoremleri

Bölüm Sonu Alıştırmaları

- E1. Zayıf büyük sayılar yasası, uygun koşullarda $\bar{X}_n \xrightarrow{p} \mu$ olduğunu söyler. Pratik anlamı, örneklem ortalamalarının n büyüdükçe yüksek olasılıkla anakütle ortalamasına yaklaşmasıdır.
- E2. Örneğin sabit bir rassal değişken Z için $X_n = Z/n$ alınırsa $X_n \xrightarrow{p} 0$ olur; çünkü $P(|Z|/n > \varepsilon) = P(|Z| > n\varepsilon) \rightarrow 0$.
- E3. X_i 'lerin ortalaması μ , varyansı σ^2 ise büyük n için $Z_n = \sqrt{n}(\bar{X}_n - \mu)/\sigma$ yaklaşık standart normaldir.
- E4. Delta yöntemi, T_n 'in asimptotik dağılımından $g(T_n)$ 'in dağılımını türev yardımıyla yaklaşık; özellikle doğrusal olmayan dönüşümlerde kullanılır.

B.8 Bölüm 8: Özel Dağılımlar ve Dağılım Aileleri

Bölüm Sonu Alıştırmaları

- E1. Sabit sayıda üründe kusurlu sayısı, her ürün için aynı kusur olasılığı ve yaklaşık bağımsızlık varsa binom modeliyle modellenebilir.
- E2. Geometrik dağılım ilk başarıya kadar deneme sayısını; negatif binom dağılımı ise belirli sayıda başarıya kadar deneme sayısını modeller.
- E3. Üstel hafızasızlık $P(X > s + t \mid X > s) = P(X > t)$ demektir; geçen süre kalan bekleme dağılımını değiştirmez.

- E4. Normal dağılım, çok sayıda küçük ve yaklaşık bağımsız etkinin toplamı/ortalaması merkezi limit sezgisiyle yaklaşık normal olduğu için sık görülür.

B.9 Bölüm 9: Koşullu Dağılımlar ve Beklenti

Bölüm Sonu Alıştırmaları

- E1. Koşullu beklenti, mevcut bilgi verildiğinde bir rassal değişkenin koşullu ortalamasıdır. Koşullanan bilgi rassal olduğunda koşullu beklenti de rassal değişken olur.
- E2. $A_1, \dots, A_k$ bir ayrışım ise, tanımlı olduğu durumda $E(X) = \sum_{i=1}^k E(X | A_i)P(A_i)$.
- E3. Tanı testi örneğinde hastalığın önsel olasılığı, pozitif test sonucu ve testin duyarlılık/yanlış pozitiflik bilgisiyle Bayes teoremi kullanılarak güncellenir.
- E4. Sıradan bağımsızlık $P(A \cap B) = P(A)P(B)$ der; koşullu bağımsızlık ise bu çarpınlaşmanın üçüncü bir olay veya değişken üzerine koşullandıktan sonra geçerli olmasıdır.

B.10 Bölüm 10: Olasılık Eşitsizlikleri ve Yakınsama

Temel Düzey

- B1. Markov eşitsizliği $P(X \geq 30) \leq E[X]/30 = 12/30 = 0.4$ verir.
 B2. Çebişev eşitsizliği $P(|X - 5| \geq 6) \leq 9/36 = 0.25$ verir.
 B3. Jensen ile $(E[X])^2 \leq E[X^2]$ olur; dolayısıyla $\text{Var}(X) = E[X^2] - (E[X])^2 \geq 0$.
 B4. Boole eşitsizliği $P(A \cup B \cup C) \leq P(A) + P(B) + P(C)$ der; burada üst sınır 0.35 olur.
 B5. $X_n \xrightarrow{p} X$, her $\varepsilon > 0$ için $P(|X_n - X| > \varepsilon) \rightarrow 0$ demektir.

Orta Düzey

- I1. Markov eşitsizliğini negatif olmayan $|X|^p$ değişkenine uyguluyoruz: $P(|X| \geq a) = P(|X|^p \geq a^p) \leq E|X|^p/a^p$.
 I2. Cauchy-Schwarz ile $E|X_n - X| \leq \sqrt{E[(X_n - X)^2]}$; bu nedenle L^2 yakınsama L^1 yakınsamayı gerektirir.
 I3. $\text{Var}(\bar{X}_n) = \sigma^2/n$ olduğundan Çebişev $P(|\bar{X}_n - \mu| > \varepsilon) \leq \sigma^2/(n\varepsilon^2) \rightarrow 0$ verir.
 I4. $\sum_n n^{-3} < \infty$ olduğundan birinci Borel-Cantelli lemması A_n olaylarının olasılık 1 ile yalnız sonlu sayıda gerçekleşeceğini söyler.
 I5. $Z \sim N(0, 1)$ için $P(Z \geq a) = P(e^{tZ} \geq e^{ta}) \leq e^{-ta+t^2/2}$; $t = a$ seçimiyle $P(Z \geq a) \leq e^{-a^2/2}$.

İleri Düzey

- A1. Markov'u $(X - \mu + t)^2$ değişkenine uygulayıp $(\sigma^2 + t^2)/(a + t)^2$ ifadesini $t > 0$ üzerinde en iyileyerek $P(X - \mu \geq a) \leq \sigma^2/(\sigma^2 + a^2)$ elde edilir.
 A2. $X_n \xrightarrow{p} X$ ise F_X 'in süreklilik noktası x için $P(X \leq x - \varepsilon) - P(|X_n - X| > \varepsilon) \leq P(X_n \leq x) \leq P(X \leq x + \varepsilon) + P(|X_n - X| > \varepsilon)$ yazılır. Önce $n \rightarrow \infty$, sonra $\varepsilon \downarrow 0$ alınır ve $X_n \xrightarrow{d} X$ elde edilir.
 A3. $X_n = Z$ ve X de $Z \sim N(0, 1)$ ile aynı dağılıma sahip bağımsız bir kopya olsun. Bu durumda $X_n \xrightarrow{d} X$ olur; fakat bu bağımsız kopyaya olasılıkta yakınsama yoktur.
 A4. Ortak yakınsama lemmasına göre $X_n \xrightarrow{d} X$ ve $Y_n \xrightarrow{p} c$ ise $(X_n, Y_n) \xrightarrow{d} (X, c)$. Sürekli dönüşüm teoremi $h(x, y) = x + y$ için uygulanırsa $X_n + Y_n \xrightarrow{d} X + c$ elde edilir.

A5. Hoeffding $P(|\hat{p} - p| \geq 0.01) \leq 2e^{-2n(0.01)^2} \leq 0.05$ verir; bu nedenle $n \geq \log(40)/0.0002 \approx 18445$ yeterlidir.

B.11 Bölüm 11: Simülasyon ve Monte Carlo Yöntemleri

Bölüm Sonu Alıştırmaları

- E1. Sabit tohumla çok sayıda standart normal üretilip $\hat{p} = n^{-1} \sum I\{Z_i > 1.96\}$ hesaplanır; Monte Carlo standart hatası $\sqrt{\hat{p}(1 - \hat{p})/n}$ olur.
- E2. $U_i \sim U(0, 1)$ için $\hat{I} = n^{-1} \sum_{i=1}^n \sqrt{U_i}$ kullanılır. Kesin değer $\int_0^1 \sqrt{x} dx = 2/3$ 'tür.
- E3. Monte Carlo standart hatası $1/\sqrt{n}$ ile orantılıdır; n dört katına çıkınca standart hata ikiye bölünür.
- E4. Kontrol değişkeni, beklenen değeri bilinen ve hedef değişkenle güçlü korelasyona sahip bir değişken varsa yararlıdır; bilinen beklenti kullanılarak varyans azaltılır.

Ek C
Hızlı Başvuru Kartı

Temel Dağılımlar İçin Hızlı Başvuru

Bu ek, en sık kullanılan dağılımların temel formüllerini ve özet ölçülerini tek bakışta görebilmeniz için hazırlanmıştır. Amaç, konu çalışırken ya da soru çözerken gerekli bilgiyi hızlıca bulmayı kolaylaştırmaktır.

Dağılım	PMF / PDF	$\mathbb{E}[X]$	$\text{Var}(X)$
Bernoulli(p)	$P(X = x) = p^x(1-p)^{1-x}, x \in \{0, 1\}$	p	$p(1-p)$
Binom(n, p)	$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$
Poisson(λ)	$P(X = k) = e^{-\lambda} \lambda^k / k!$	λ	λ
Geometrik(p)	$P(X = k) = (1-p)^{k-1} p, k \geq 1$	$1/p$	$(1-p)/p^2$
Düzgün(a, b)	$f(x) = 1/(b-a), a \leq x \leq b$	$(a+b)/2$	$(b-a)^2/12$
Üstel(λ)	$f(x) = \lambda e^{-\lambda x}, x \geq 0$	$1/\lambda$	$1/\lambda^2$
Normal(μ, σ^2)	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)}$	μ	σ^2

Kavram Haritası:

Kesikli dağılımlarda tek tek sonuçların olasılığı PMF ile, sürekli dağılımlarda ise aralıkların olasılığı PDF ve integral mantığıyla değerlendirilir. Beklenen değer merkezi, varyans ise yayılımı özetler.

Hızlı Notlar

- Binom dağılımı, sabit sayıda bağımsız Bernoulli denemesinin toplamıdır.
- Poisson dağılımı, nadir olay sayımlarında temel modeldir.
- Üstel dağılım, Poisson sürecindeki bekleme sürelerini modeller.
- Normal dağılım, standartlaştırma ile Z tablosu üzerinden hesaplamayı mümkün kılar.

Ek D
Modelleme Rehberi

Olasılık Modeli Kurma Sanatı

Bu ek, bir problemi doğrudan formüle taşımadan önce nasıl bir olasılık modeline dönüştüreceğimizi kısa ve sistematik biçimde özetler. Amaç, yalnız sonuca ulaşmak değil; hangi durumda hangi modeli neden seçtiğini açıkça görebilmektir.

Sezgisel Katman

Gerçek hayattaki soru çoğu zaman “kaç tane?”, “ne kadar sürer?” ya da “hangi aralıkta değişir?” biçiminde gelir. Doğru model seçiminin ilk adımı, bu sorunun hangi rassal değişkene karşılık geldiğini açıkça yazmaktır.

Matematiksel Katman:

Model kurma süreci dört temel adıma indirgenebilir: rassal değişkeni tanımla, destek kümesini belirle, uygun dağılım ailesini seç ve parametreleri bağlamdan çıkar. Son adım ise, elde edilen modelin yorum gücünü problem bağlamında sınamaktır.

Mini Karar Ağacı

Model seçerken aşağıdaki kısa akışı izleyin. Her adım, bir sonraki kararı daraltmak için tasarlanmıştır.

- Önce soruyu isimlendirin.** Sorulan şey olay *sayısı* mı, bir *bekleme süresi* mi, yoksa bir *ölçüm* mü?
- Sonra destek kümesini kontrol edin.** Değerler yalnız tam sayılar alıyorsa çoğu zaman kesikli; zaman, uzunluk, ağırlık gibi ara değerler alabiliyorsa çoğu zaman sürekli model gerekir.
- Ardından uygun aileyi düşünün.**
 - Sabit sayıda denemede başarı sayısı → **Binom**
 - Belirli sürede nadir olay sayısı → **Poisson**
 - Bir sonraki olaya kadar geçen süre → **Üstel**
 - Merkez etrafında yaklaşık simetrik ölçüm → **Normal**

4. **Parametreyi bağlamdan çekin.** Başarı olasılığı, olay hızı, ortalama veya standart sapma çoğu zaman problem metninde doğrudan ya da dolaylı olarak verilir.
5. **Sonuçta bir gerçeklik kontrolü yapın.** Seçtiğiniz model, değişkenin doğasıyla ve sorunun bağlamıyla gerçekten uyumlu mu?



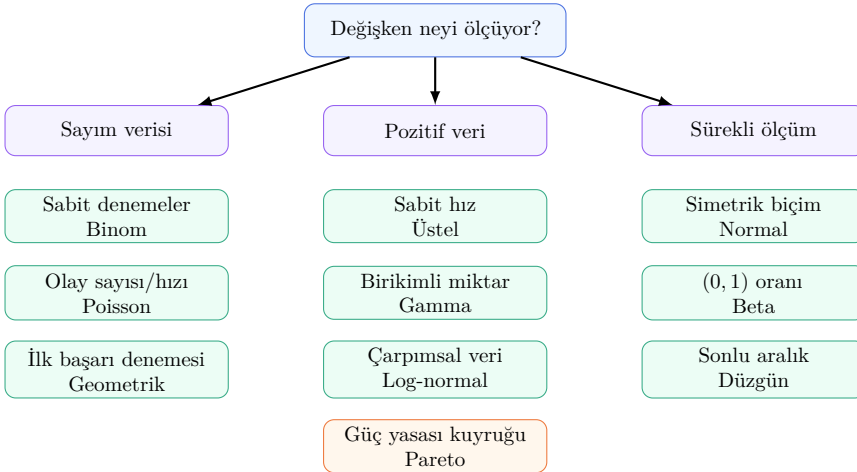
Kavram Haritası:

Soru türü → destek kümesi → dağılım ailesi → parametre → yorum. Bu zincir bozulduğunda model seçimi de zayıflar.

Bir Sayfalık Dağılım Seçim Posterı

Araştırmacı Notu

Bu akış şeması, “elimdeki veri hangi dağılıma uyuyor?” sorusuna hızlı bir ilk yanıt vermek için tasarlanmıştır. Nihai karar için alan bilgisi, grafikler, tanısal kontroller ve gerekirse simülasyonla duyarlılık analizi birlikte kullanılmalıdır.



Yaygın Hata Kalıbı:

Karar ağacı, otomatik model seçici değildir. Özellikle bağımlı gözlemler, sansürlü veriler, sıfır yığılması, aykırı değerler ve ölçüm hataları varsa dağılım seçimi ek kontroller gerektirir.

Vaka Analizi:

Çağrı merkezi örneği: Bir saatte gelen çağrı sayısı soruluyorsa rassal değişken olay sayısidir; bu nedenle kesikli model gerekir. Eğer çağrılar bağımsız ve seyrek kabul ediliyorsa Poisson doğal adaydır. Buna karşılık “bir sonraki çağrıya kadar geçen süre” soruluyorsa aynı sistem bu kez üstel dağılım ile modellenenebilir.

Yanlış Çözüm:

Bir ürün ömrünü modellemek isteyen öğrenci, negatif değer alamayan bir değişken için doğrudan normal dağılım seçiyor.

Doğru Çözüm:

Ömür değişkenleri için önce destek kümesi düşünülmelidir. Negatif değer alamayan süre değişkenlerinde üstel veya gamma ailesi daha doğal olabilir; normal model ancak simetrik ve ortalamaya göre yeterince büyük destek durumlarında yaklaşık olarak savunulabilir.

Neden Hata Yapıldı?

Hata, formülden önce değişkenin doğasını incelememekten kaynaklanır. Model seçimi yalnız hesap kolaylığına göre değil, değişkenin yapısına göre yapılmalıdır.

Kendini Dene:

Aşağıdaki üç problemi uygun dağılım ailesi ile eşleştirmeyi deneyin: (i) 20 denemede başarı sayısı, (ii) bir saatlik aralıktaki gelen e-posta sayısı, (iii) bir makinenin bozulmasına kadar geçen süre.

Bölüm Sonu Hızlı Kart

- Sayı sayıyorsan önce kesikli model düşün.
- Süre veya ölçüm ise sürekli model düşün.
- Destek kümesi, model seçiminde ilk filtredir.
- Parametre yorumu yapılmayan model, eksik modeldir.

Kaynakça

- Bertsekas, D. P., & Tsitsiklis, J. N. (2008). *Introduction to probability* (2nd ed.). Athena Scientific.
- Billingsley, P. (1995). *Probability and measure* (3rd ed.). Wiley.
- Blitzstein, J. K., & Hwang, J. (2019). *Introduction to probability* (2nd ed.). Chapman and Hall/CRC.
- Brualdi, R. A. (2010). *Introductory combinatorics* (5th ed.). Pearson.
- Bryc, W. (1996). *Applied probability and stochastic processes*. Lecture notes, University of Cincinnati.
- Capiński, M., & Kopp, E. (2004). *Measure, integral and probability* (2nd ed.). Springer.
- Casella, G., & Berger, R. L. (2002). *Statistical inference* (2nd ed.). Duxbury.
- Çınlar, E. (2011). *Probability and stochastics*. Springer.
- David, H. A., & Nagaraja, H. N. (2003). *Order statistics* (3rd ed.). Wiley.
- DeGroot, M. H., & Schervish, M. J. (2012). *Probability and statistics* (4th ed.). Pearson.
- Dekking, F. M., Kraaikamp, C., Lopuhaä, H. P., & Meester, L. E. (2005). *A modern introduction to probability and statistics: Understanding why and how*. Springer.
- Devroye, L. (1986). *Non-uniform random variate generation*. Springer-Verlag.
- Feller, W. (1968). *An introduction to probability theory and its applications* (Vol. 1, 3rd ed.). Wiley.
- Feller, W. (1971). *An introduction to probability theory and its applications* (Vol. 2, 2nd ed.). Wiley.
- Fishman, G. S. (1996). *Monte Carlo: Concepts, algorithms, and applications*. Springer.

- Geiss, C., & Geiss, S. (2004). *An introduction to probability theory*. Lecture notes.
- Gentle, J. E. (2003). *Random number generation and Monte Carlo methods* (2nd ed.). Springer.
- Grimaldi, R. P. (2003). *Discrete and combinatorial mathematics: An applied introduction* (5th ed.). Addison-Wesley.
- Grimmett, G., & Stirzaker, D. (2020). *Probability and random processes* (4th ed.). Oxford University Press.
- Grinstead, C. M., & Snell, J. L. (1997). *Introduction to probability* (2nd rev. ed.). American Mathematical Society.
- Gut, A. (2005). *Probability: A graduate course*. Springer.
- Hogg, R. V., McKean, J. W., & Craig, A. T. (2019). *Introduction to mathematical statistics* (8th ed.). Pearson.
- Hsu, H. P. (1997). *Schaum's outline of theory and problems of probability, random variables, and random processes*. McGraw-Hill.
- Jaynes, E. T. (2003). *Probability theory: The logic of science*. Cambridge University Press.
- Kolmogorov, A. N. (1956). *Foundations of the theory of probability* (2nd English ed., N. Morrison, Trans. Ed.). Chelsea Publishing Company.
- Lipschutz, S., & Lipson, M. (2011). *Schaum's outline of probability* (2nd ed.). McGraw-Hill.
- Montgomery, D. C., & Runger, G. C. (2018). *Applied statistics and probability for engineers* (7th ed.). Wiley.
- Mood, A. M., Graybill, F. A., & Boes, D. C. (1974). *Introduction to the theory of statistics* (3rd ed.). McGraw-Hill.
- Mukhopadhyay, N. (2000). *Probability and statistical inference*. Marcel Dekker.
- Nelson, E. (1987). *Radically elementary probability theory*. Annals of Mathematics Studies, Princeton University Press.
- Ramachandran, K. M., & Tsokos, C. P. (2009). *Mathematical statistics with applications*. Academic Press.
- Robert, C. P., & Casella, G. (2004). *Monte Carlo statistical methods* (2nd ed.). Springer.
- Rosen, K. H. (2018). *Discrete mathematics and its applications* (8th ed.). McGraw-Hill.
- Ross, S. M. (2014a). *A first course in probability* (9th ed.). Pearson.
- Ross, S. M. (2014b). *Introduction to probability models* (11th ed.). Academic Press.
- Ross, S. M. (2019). *A first course in probability* (10th ed.). Pearson.

- Shao, J. (2003). *Mathematical statistics* (2nd ed.). Springer.
- Soong, T. T. (2004). *Fundamentals of probability and statistics for engineers*. Wiley.
- Tarantola, A. (2001). *Probability and measurements*. Manuscript, Université de Paris, Institut de Physique du Globe.
- van der Vaart, A. W. (1998). *Asymptotic statistics*. Cambridge University Press.
- Vershynin, R. (2018). *High-dimensional probability: An introduction with applications in data science*. Cambridge University Press.
- Walpole, R. E., Myers, R. H., Myers, S. L., & Ye, K. (2012). *Probability & statistics for engineers & scientists* (9th ed.). Pearson.
- Wasserman, L. (2004). *All of statistics: A concise course in statistical inference*. Springer.
- Williams, D. (1991). *Probability with martingales*. Cambridge University Press.

Dizin

- Ampirik Dağılım Fonksiyonu, 94
Anakütle, Örneklem ve İstatistik, 94
Anakütle ve Parametre, 94
Aralık Olasılığı, 41, 60
Artan ve Azalan Olay Dizileri, 3
Asimptotik Düşünce, 112
Asimptotik Normallik, 122
Ayrık Olaylar, 3
Ayrışım, 11
- Bağımsız Olaylar, 13
Bağımsız Rassal Değişkenler, 81
Bağımsız Toplamlarda MGF Çarpımı, 67
Bağımsız Toplamların Karakteristik Fonksiyonu, 68
Bağımsız Ürün Olasılığı, 15
Bağımsızlığın Mühendislik Uygulaması: Sistem Güvenilirliği, 15
Bağımsızlık, 13
Bağımsızlık ve Koşullama, 151
Basit Rastgele Yürüyüş, 191
Bayes Kuralı ve Güncelleme, 155
Bayes Teoremi, 12
Bayes Teoreminin Olasılık Oranı Biçimi, 12
Beklenen Değer, 45
Beklenen Değer, Varyans ve Momentler, 45
Beklentinin Bilgiye Koşullu Yorumu, 53
Beklentinin Doğrusallığı ve Varyansın Ölçeklenmesi, 46
Bernoulli Dağılımı, 50, 131
Bernoulli ve Binom Dağılımları, 130
Beta Dağılımı, 70, 138
Beta Dağılımının Ortalaması, 70
Bilinçsiz İstatistikçi Yasası (LOTUS), 46
Binom Dağılımı, 50, 131
Binom Dağılımının Ortalaması ve Varyansı, 50
- Binom Katsayıları ve Binom Teoremi, 30
Birleşim, Kesişim ve Tümlen, 2
Bölge Olasılığının Koordinattan Bağımsızlığı, 16
Bonferroni Eşitsizlikleri, 6
Boole Eşitsizliği (Alt Toplamsallık), 6
Boole Eşitsizliği ve Birleşim Sınırı, 171
Borel–Cantelli Lemması, 173
Boş Kümenin Olasılığı, 4
Büyük Sayılar Yasası, 115
- Çarpıklık ve Basıklık, 46, 65
Çarpım Kuralı, 9
Çarpma İlkesi, 8
Cauchy Dağılımı, 71
Cauchy–Schwarz Eşitsizliği, 170
Çebişev Eşitsizliği, 167
Çok Değişkenli Jacobian Kuralı, 64
- Dağılım Ailesi, 130
Dağılım Ailesi ve Parametre Mantiği, 130
Dağılım Fonksiyonu, 40
Dağılım Fonksiyonunun Temel Özellikleri, 40
Dağılım Seçim Kontrol Listesi, 142
Dağılımda Yakınsama, 113, 172
Dağılımın Bilgiye Bağlı Yorumu, 53
Dağılımlar Arası İlişkiler, 141
Dağılımlara Hazırlık: Binom ve Hipergeometrik Yapı, 33
De Morgan Kuralları, 3
Değişim Katsayısı, 66
Delta Yöntemi, 121
Doğrusal Birleşimlerin ve Toplamların Varyansı, 85
Dönüşümle Yeni Dağılım Elde Etme, 63

Eş Olasılıklı Sonlu Model, 24
Eşitsizliklerden Büyük Sayılar Yasasına, 175

F Dağılımı, 140
Faktöriyel, 25
Faktöriyel ve Permütasyon, 25

Gamma Dağılımı, 69, 137
Geometrik Dağılım, 50, 132
Geometrik Dağılımın Beklenen Değeri, 51
Geometrik Dağılımın Hafızasızlığı, 51
Geometrik ve Negatif Binom Dağılımları, 132
Gösterge Değişkenleri ve Sayma, 47
Gösterge Rassal Değişkeni, 47
Göstergenin Beklenen Değeri, 47

Hemen Hemen Kesin Yakınsama, 114, 172
Hipergeometrik Dağılım, 51, 133
Hipergeometrik Dağılımın Ortalaması, 51

İçerme-Dışlama ile Sayma, 31
İçerme-Dışlama İlkesi, 6
İki Değişkenli Beklenti, 82
İki Değişkenli Normal Dağılım, 86
İki Olay İçin Toplama Kuralı, 5

Jensen Eşitsizliği, 66, 169

Kabul-Red Yöntemi, 185
Kanıt Altında Dağılım, 53
Kantil Fonksiyonu, 53, 62
Kantiller ve Olasılık İntegral Dönüşümü, 52
Kantiller ve Ters Dönüşüm, 62
Karakteristik Fonksiyon, 68
Karakteristik Fonksiyonların Tekillliği, 68
Karakteristik Fonksiyonun Temel Özellikleri, 69
Karesel Kayıp Altında Ortalama, 65
Karma Tip Dağılım, 44
Karma Tip Dağılımlar ve Model Seçimi, 44
Karşılıklı Bağımsızlık, 14
Kesikli Dağılım Simülasyonu, 184
Kesikli Değişkende Olay Olasılığı, 42
Kesikli Düzgün Dağılım, 52, 136
Kesikli Koşullu Dağılımlar, 149
Kesikli Ortak Dağılımlar ve Marjinaller, 79
Kesikli Rassal Değişkenler, 42
Ki-Kare Dağılımı, 140
Klasik Olasılık, 8
Kolmogorov Aksiyomları, 4
Kombinasyon, 27
Konveks Fonksiyon, 169

Korelasyon Katsayısı, 84
Koşullama Fikrinin Temeli, 148
Koşullama ile Parçalama Stratejisi, 158
Koşullu Bağımsızlık, 14
Koşullu Beklenti, 152
Koşullu Beklenti Bir Rassal Değişkendir, 152
Koşullu Beklenti ve En İyi Tahmin, 156
Koşullu Dağılım Fonksiyonu, 151
Koşullu Olasılık, 9, 148
Koşullu Olasılık Bir Olasılık Ölçüsüdür, 10
Koşullu Olasılık ve Bağımsızlık, 13
Koşullu PMF, 149
Koşullu Varyans, 154
Koşullu Varyans ve Toplam Varyans Yasası, 154
Koşullu Yoğunluk, 150
Kovaryans, 82
Kule Özelliği ve Bilgi Düzeyleri, 158
Kümülatif Aralık Yöntemi, 184
Kuyruk İntegrali ile Beklenti, 61

L^1 ve L^2 Yakınsama, 172
Laplace Dağılımı, 71
Lognormal Dağılım, 71
 L^p Uzayı, 169

Marjinal PMF, 79
Marjinal Yoğunluk, 80
Markov Eşitsizliği, 166
Markov Eşitsizliği (Ön İzleme), 47
Medyan ve Çeyrekler, 62
Merkezi Limit Teoremi, 117
Merkezi Limit Teoremi ve Büyük Örneklem Yaklaşımı, 103
MGF Tekillliği, 68
MGF'den Moment Üretimi, 49
MGF'nin Tekliği ve Toplam Özelliği, 50
Moment, 45
Moment Sınırları ve Genelleştirme, 169
Moment Üreten Fonksiyon, 49, 66, 67
Moment Üretim Teoremi, 67
Moment ve Merkezi Moment, 65
Momentler ve Dağılımın Şekli, 65
Monoton Olmayan Dönüşüm, 64
Monotonluk, 5
Monte Carlo Düşüncesi, 182
Monte Carlo Güven Aralığı, 187
Monte Carlo İntegrasyonu, 188
Monte Carlo Standart Hatası, 187
Monte Carlo Tahmini, 182
Multinom Dağılım, 87
Multinom Katsayıları, 31
Multinom Katsayısı, 31

- Neden Olasılık Eşitsizlikleri?, 166
 Negatif Binom Dağılımı, 132
 Negatif Binom (Pascal) Dağılımı, 52
 Negatif Olmayan Değişkende Kuyruk
 Toplamı Formülü, 41
 Normal Anakütlerde Kesin Örneklem
 Dağılımları, 100
 Normal Dağılım, 69, 139
 Normal Dağılımdan Türeyen Dağılımlar,
 140
 Normal Yaklaşım ve Süreklilik Düzeltmesi,
 119
 Normalden Ki-Kareye Geçiş, 69

 Olasılığın Sürekliliği, 7
 Olasılığın Yorumu ve Modelleme Disiplini,
 16
 Olasılık İntegral Dönüşümü, 53, 63
 Olasılık Kütle Fonksiyonu, 42
 Olasılık Ölçüsü ve Kolmogorov Aksiyomları,
 4
 Olasılık Problemlerinde Sayma, 32
 Olasılık Üreten Fonksiyon, 48
 Olasılık Uzaı, 4
 Olasılık Yoğunluğu ve Hacim Elemanı, 16
 Olasılık Yoğunluk Fonksiyonu, 43
 Olasılıkta Yakınsama, 113, 172
 Olay Cebiri ve Küme İşlemleri, 2
 Olay Dizisinin Üst ve Alt Limiti, 3
 Olay Olasılığı Tahmini, 186
 Örnek Nokta ve Olay, 2
 Örnek Uzaı, 2
 Örneklem Oranı ve Binom Modeli, 101
 Örneklem Ortalamasının Dağılımı, 97
 Örneklem Varyansı, 98
 Örneklem ve İstatistik, 94
 Örneklem Dağılımı, 96
 Ortak Dağılım Fonksiyonu, 78
 Ortak Olasılık Kütle Fonksiyonu, 79
 Ortak Yoğunluk Fonksiyonu, 80
 Ortalama Karede Yakınsama, 113

 Permütasyon, 25
 Permütasyon ve Kombinasyon, 8
 PGF ile Moment ve Toplam, 49
 PGF'den Olasılıkların Elde Edilmesi, 48
 Poisson, Binomun Limiti Olarak, 52
 Poisson Dağılımı, 51, 134
 Poisson Dağılımının Ortalaması ve Varyansı,
 52
 Poisson Süreci, 137

 Rassal Değişken, 40
 Rassal Değişken Dizisi, 112
 Rassal Değişken Kavramı, 40
 Rassal Değişkenin Ürettiği Olay, 40
 Rassal Değişkenlerin Bağımsızlığı, 81
 Rassal Deney, 2
 Rassal Deney, Örnek Uzaı ve Olay, 2
 Rassal Örneklem, 95
 Rassal Sürec, 191

 Sağ Kalım Fonksiyonu, 41
 Sayma Probleminin Olasılıktaki Yeri, 24
 Seri ve Paralel Sistem Güvenilirliği, 15
 Sıçrama Büyüklüğü, 41
 σ -cebiri, 3
 Simülasyon Sonucunu Raporlama, 192
 Sıralı İstatistik, 102
 Sıralı İstatistikler, 102
 Sıralı-Sırasız ve Tekrarlı-Tekrarsız Seçim,
 29
 Sonlu Anakütleden Örneklem, 99
 Sonlu Toplamsallık, 6
 Sonlu ve Eş Olasılıklı Modeller, 8
 Sözde Rassal Sayı, 183
 Sözde Rassal Sayılar ve Tekrarlanabilirlik,
 183
 Standart Hata, 97
 Stokastik Mertebe: O_p ve o_p , 113
 Student t Dağılımı, 140
 Sürekli Değişkende Aralık Olasılığı, 43
 Sürekli Değişkende Tek Nokta Olasılığı, 43
 Sürekli Dönüşüm ve Slutsky Teoremi, 120
 Sürekli Koşullu Yoğunluk, 150
 Sürekli Ortak Dağılımlar ve Marjinaller, 80
 Sürekli Rassal Değişken, 60
 Sürekli Rassal Değişkenler, 43
 Sürekli Düzgün Dağılım, 136
 Sürekli Yoğunluk ve Dağılım Fonksiyonu,
 60

 Tehlike (Hazard) Fonksiyonu, 61
 Tek Değişkenli Yoğunluk Dönüşümü, 63
 Tekrarlı Kombinasyon, 29
 Tekrarlı Permütasyon, 26
 Temel Kesikli Dağılımlar, 50
 Temel Sürekli Dağılım Aileleri, 69
 Ters Dönüşüm Yöntemi, 62, 184
 Tohum Değeri, 183
 Toplam Beklenti Yasası, 153
 Toplam Olasılık Yasası, 11
 Toplam Olasılık Yasası ve Bayes Teoremi,
 11
 Toplama ve Çarpma Kuralları, 24
 Toplamın Yoğunluğu: Evrişim, 65
 Tümleyen Kuralı, 5
 Tümleyenlerin Bağımsızlığı, 14

Türetilmiş Dağılım, 63
 Tutarlı Tahmin Edici, 122
 Tutarlılık ve Asimptotik Normallik, 122

Ürün σ -cebiri, 15
 Ürün Uzayları ve Bağımsız Deneyler, 14
 Üstel Dağılım, 69, 137
 Üstel Dağılımın Hafızasızlığı, 70
 Üstel ve Gamma Dağılımları, 136

Varyans Azaltma Fikirleri, 189
 Varyans ve Standart Sapma, 45

Varyansın Hesaplama Formülü, 45

Weibull Dağılımı, 70

Yakınsama İlişkileri, 172
 Yakınsama Türleri, 113, 172
 Yerine Koymalı Seçim, 32
 Yerine Koymalı ve Yerine Koymasız
 Örneklem, 32
 Yerine Koymasız Seçim, 32
 Yoğunlaşma: Hoeffding Eşitsizliği, 174
 Yoğunluk Fonksiyonu, 60

Olasılık Teorisi

Temeller, Dağılımlar ve Uygulamalı Modelleme

Bu kitap, olasılık teorisini aksiyomatik temellerden başlayarak rassal değişkenler, dağılımlar, koşullu olasılık, beklenen değer, yakınsama, limit teoremleri ve simülasyon yöntemlerine uzanan bütünlüklü bir yapı içinde sunar.

Matematiksel tanımlar ve teoremler; sezgisel açıklamalar, çözümlü örnekler, gerçek hayat uygulamaları ve bölüm sonu problemleriyle desteklenmiştir. Python ve R kodları, okuyucunun hesaplamaları yeniden üretmesine, rassal süreçleri simüle etmesine ve kuramsal sonuçları sayısal deneylerle incelemesine olanak sağlar.

Lisans ve lisansüstü derslerde temel ya da yardımcı ders kitabı olarak kullanılabilecek eser; istatistik, veri bilimi, matematik, mühendislik, ekonomi ve eğitim alanlarındaki araştırmacılar için de düzenli bir başvuru kaynağı olarak tasarlanmıştır.

Kitabın kapsamı

- Olasılık aksiyomları ve sayma yöntemleri
- Kesikli ve sürekli dağılımlar
- Ortak dağılımlar, bağımlılık ve koşullu beklenti
- Örneklem, asimptotik teori ve limit teoremleri
- Python ve R ile simülasyon ve uygulamalı modelleme

Prof. Dr. İbrahim Güney

Istanbul Sabahattin Zaim University

TEMELLER, DAĞILIMLAR VE UYGULAMALI MODELLEME



ISBN 978-625-7558-54-9



9 786257 558549