

Olasılık ve İstatistik Ders Notları

Arzu Erdem

© 2013/2014 Güz dönemi Matematik Bölümü notları¹

Kaynaklar

- Probabilty and Statistic, Morris H. DeGroot, 1986.
- Applied Probabilty and Statistic, Mario Lefebvre, 2006.
- A Modern Introduction to Probability and Statistics, Frederik Michel Dekking, Cornelis Kraaikamp, Hendrik Paul
Lopuhaä, Ludolf Erwin Meester, 2005.
- A course in Probability and Statistics, Charles J. Stone, 1996.
- INTRODUCTION TO PROBABILITY AND STATISTICS FOR ENGINEERS AND SCIENTISTS, Sheldon M. Ross,
2004.
- Probability & Statistics for Engineers & Scientists, Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying
Ye, 2012.
- Theory and Problems of Probability and Statistics, Murray R. Spiegel, 1998.
- Teori ve Problemlerle Olasılık, Seymour Lipschutz, Schaum Serisi, 1974.

¹email : erdem.arzu @ gmail.com ; December 23, 2014

Contents

List of Figures	v
List of Tables	vii
Chapter 1. Olasılık	1
Olasılığın Tarihi	1
Olasılığa Giriş	1
Sayma Yöntemleri	4
Olasılık Aksiyomları	16
Koşullu Olasılık	22
Bağımsız Olaylar (Independent Events)	28
Rasgele Değişkenler (Random Variables)	31
Birleşik Olasılık Dağılımı (Joint Probability Distributions)	44
Rasgele değişkenlerinin beklenen değeri (Mathematical Expectation of random variable)	64
Rasgele değişkenlerinin varyansı ve kovaryansı (Variance and Covariance of Random Variables)	72
Rasgele değişkenlerinin lineer kombinasyonlarının varyansı ve ortalaması (Means and Variances of Linear Combination)	94
Moment ve Moment üreten fonksiyonlar (Moments and Moment-Generating Functions)	94
Chebysev Eşitsizliği ve Büyük Sayılar Yasası (CHEBYSHEV'S INEQUALITY AND THE LAW OF LARGE NUMBERS)	101
Önemli Kesikli Dağılım Fonksiyonları	101
13.1 Binom Dağılımı (Binomial Distributions)	101
13.2 Hipergeometrik Dağılım (Hypergeometric Distribution)	106
13.3 Poisson Dağılımı (Poisson Distribution)	108
Önemli Sürekli Dağılım Fonksiyonları	110
14.1 Düzgün Sürekli Dağılım (Continuous Uniform Distribution)	110
14.2 Normal Dağılım (Normal Distribution)	111
Chapter 2. İstatistik	121
Giriş - Örneklem Teorisi (Sampling Theory)	121
Veri Düzenlemesi ve Analizi	126
2.1 Giriş ve Önbilgiler	126
2.2 Veri Düzenlemesi	127
2.3 Frenkans (Sıklık) Dağılımı	128

Grafiksel Gösterimler	131
3.1 Çizgisel Grafik	131
3.2 Dairesel Grafik	132
3.3 Poligon (Frekans Çokgeni)	132
3.4 Eklemeli Frekans Çokgeni	133
Merkezi Eğilim Ölçüleri	134
4.1 Aritmetik Ortalama	135
4.2 Medyan (Orta değer = Ortanca)	137
4.3 Mod (Tepe Değer)	140
4.4 Aritmetik Ortalama, Orta Değer ve Mod'un özellikleri	142
4.4.1 Aritmetik Ortalama	142
4.4.2 Orta Değer	142
4.4.3 Mod	142
4.5 Geometrik Ortalama	142
4.6 Harmonik Ortalama	143
Dağılım Ölçüleri	143
5.1 Çarpıklık Ölçütü	149
5.2 Sivrilik (basıklık) ölçütü	150
Örnekleme Dağılımları ve Tahmin Etme	150
Örneklem Ortalaması ve Varyansın Bazı Özellikleri	151

List of Figures

2.1	Denklemin grafiđi	7
14.1(a)	Normal dađılım eđrisi, $\mu = 50, \sigma = 5$, (b) $\sigma_1 = \sigma_2, \mu_1 < \mu_2$ olduđunda normal dađılım eđrisi	112
14.2(a)	$\mu_1 = \mu_2, \sigma_1 < \sigma_2$ olduđunda normal dađılım eđrisi, (b) $\mu_1 < \mu_2, \sigma_1 < \sigma_2$ olduđunda normal dađılım eđrisi	112
14.3	$P(x_1 < X < x_2)$	113
14.4	Normal Dagilim Tablosu	119
1.1	Rasgele sayılar tablosu	124

List of Tables

Chapter 1

Olasılık

0 Olasılığın Tarihi

Belirsizlikler ve değişim kavramı insalık tarihi kadar eskidir. İnsanoğlu her zaman, hava durumundaki belirsizliklerle, yiyecek ihtiyaçlarının belirsizlikleriyle veya diğer çevresel faktörlerdeki belirsizliklerle ilgilenmiş ve burdaki belirsizlikleri ve onların etkilerini indirgemek için uğraşmıştır. Kumar oyunları da bu belirsizliğe dahil olmaktadır. M.Ö. 3500 yılında kemikler ile oynanan oyunlar varken, M.Ö. 200 yıllarında Mısırda zar oyunları bulunmuştur. Zar ile oynanan kumar oyunları çok popüler olmasının yanı sıra olasılık teorisinin gelişmesinde de önemli bir rol oynamaktadır.

Genel anlamda olasılık teorisi, Fransız matematikçiler Blaise Pascal (1623-1662) ve Pierre Fermat (1601-1665) tarafından ortaya çıkmıştır ve zar ile oynanan kumar oyunlarındaki olasılık teorisini inşa etmişlerdir. Onların çözdüğü problemlerin geçerliliği 300 yıl kadar olsa da, zar kombinasyonlarının nümeriksel olasılıkları ilk defa Girolamo Cardano (1501-1576) ile Galileo Galilei (1564-1642) tarafından verilmiştir.

O zamandan bu yana oldukça hızlı bir gelişme göstermiş olup mühendislik, fen bilimleri ve sosyal bilimler alanlarında çok fazla uygulaması vardır.

1 Olasılığa Giriş

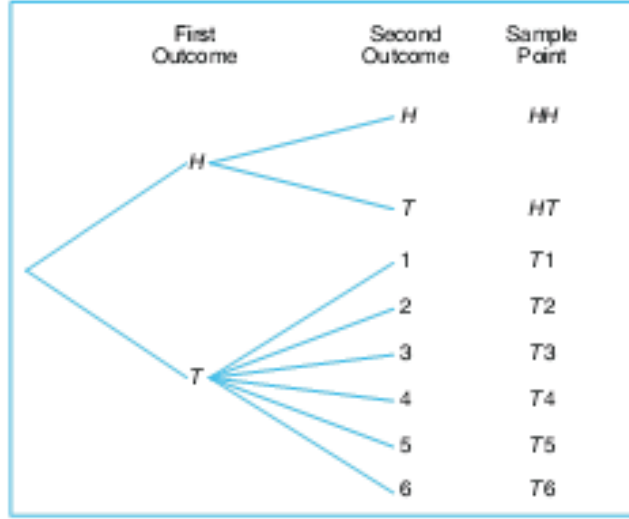
Tanım 1.1. *Sonuçları kesin olarak bilinemeyen yöntemlere deney denir.*

Örnek 1.2. *Bir deneyde, gözlemci, 10 kez sırayla atılan bir paranın en az 4 kez yazı gelmesini isteyebilir. Bir deneyde, 90 gün boyunca hava sıcaklığı ortalama olarak belirtilen bir sıcaklığın üstünde olması istenebilir. Bilimsel bir projede, projenin sonuçlarının belirtilen bir zamanda başarı olasılığının hesaplanması istenebilir.*

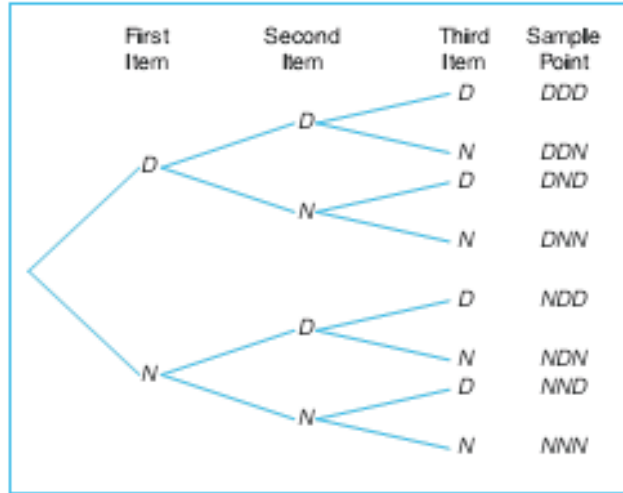
Tanım 1.3. *Rastgele bir deneyin, olası tüm sonuçlarını içeren kümeye örneklem uzayı (sample space) denir ve S ile gösterilir. Örneklem uzayındaki herbir elemana örneklem noktası veya örnek (sample) denir.*

Örnek 1.4. *Bir zarın atılmasını düşünelim. Zarın üstündeki rakamlar ile ilgileniyorsak örneklem uzayını $S = \{1, 2, 3, 4, 5, 6\}$ olarak verebiliriz. Ancak sayıların tek veya çift gelmesi ile ilgileniyorsak örneklem uzayını $S = \{\text{tek}, \text{çift}\}$ olarak yazarız. Ancak görüldüğü üzere 2. yazdığımız örneklem uzayı elemanların belirlenmesinde bize çok kesin bilgi verememektedir.*

Örnek 1.5. Bir başka deneyde, bir para atılıyor. Eğer para yazı geliyorsa bir daha atılıyor. Ancak tura geliyorsa bu seferde zar atılıyor. Bu durumda örneklem uzayını aşağıdaki gibi verebiliriz. $S = \{YY, YT, T1, T2, T3, T4, T5, T6\}$



Örnek 1.6. Diğer bir deney için, bir fabrikada üretilen ürünlerden arızalı olanları D (defective) arızalı olmayanları da N (nondefective) ile gösterelim. Bu ürünlerden 3 adet seçileceğine göre örneklem uzayını aşağıdaki gibi



Örnek 1.7. Bir deneyde, popülasyonu 1 milyonun üstünde şehirler istenebilir ki bu durumda örneklem uzayını $S = \{x : x, 1 \text{ milyon nüfusu aşan bir şehir}\}$

Tanım 1.8. Bir örneklem uzayının alt kümelerine olay denir.

Örnek 1.9. Yolda rastgele karşılaştığımız birinin hangi ayda doğduğunu merak ediyorsak, bu deneyin örneklem uzayı $S = \{\text{Ocak, Subat, Mart, Nisan, Mayıs, Haziran, Temmuz, Ağustos, Eylül, Ekim, Kasım, Aralık}\}$ dır. Bu deneyde, kişinin uzun ayda (31 günlük aylardan) doğma ihtimalini belirten olay

$L = \{Ocak, Mart, Mayıs, Temmuz, Ağustos, Ekim, Aralık\}$. Veya başka bir olay olarak , içinde "i" harfi olan aylardan birinde doğma ihtimali olarak belirtirsek, $I = \{Nisan, Haziran, Ekim\}$.

Tanım 1.10. Bir olayın gerçekleşmemesi durumuna olanaksız (impossible) denir ve \emptyset ile gösterilir. S örneklem uzayın kendisine de kesin (certain) olay denir.

Örnek 1.11. $B = \{x : x, 7 \text{ nin çift bölenleri}\}$ olayı imkansızdır yani $B = \emptyset$.

Tanım 1.12. S örneklem uzayında bir A olayının gerçekleşmediği kümeye A nın tümleyeni (complement) denir ve A' ile gösterilir. $A' = \{x \in S : x \notin A\}$.

Örnek 1.13. Bir zar atma deneyinde, asal sayıların geldiği olay $A = \{2, 3, 5\}$ olmak üzere asal sayıların gelmediği olay $A' = \{1, 4, 6\}$.

Tanım 1.14. A ve B olaylarında hem A nın hemde B nin gerçekleştiği olaya A ve B nin arakesiti (intersection) denir ve $A \cap B = \{x \in S : x \in A \text{ ve } x \in B\}$.

Örnek 1.15. E olayı bir sınavda mühendisliği seçen insanları göstermek üzere, F olayda sınavda tercihte bulunan bayanları göstermek üzere sınavda bayan ve mühendisliği seçen kişileri $E \cap F$ ile gösteririz.

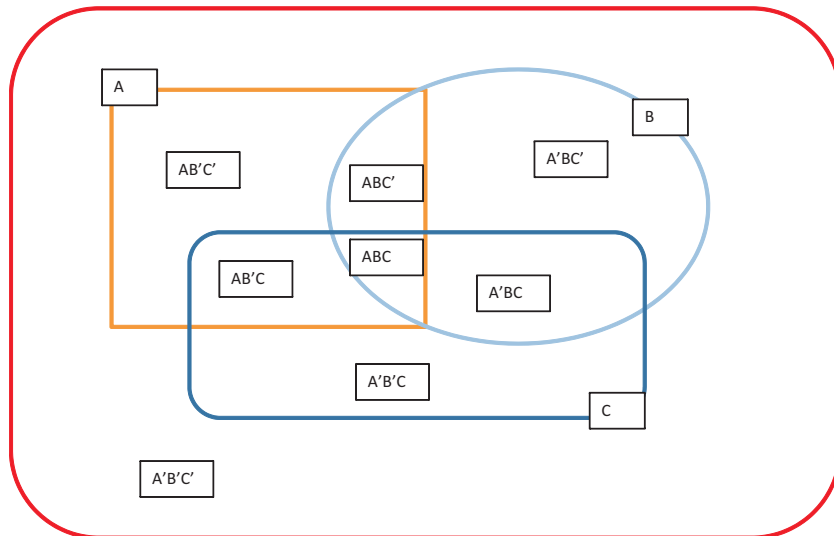
Tanım 1.16. A ve B olayları kesikli ise yani $A \cap B = \emptyset$ ise A ve B olayları birbirinin dışındadır denir.

Örnek 1.17. Üç kez peşpeşe zar atma deneyinde, A olayını 2.de yazı gelmesi, B olayını da hiç yazı gelmemesi olarak belirtirsek: $A = \{YYY, YYT, TYY, TYT\}$, $B = \{TTT\}$ olma üzere $A \cap B = \emptyset$ olduğundan A ve B olayları birbirinin dışındadır.

Tanım 1.18. A ve B olaylarında A nın veya B nin gerçekleştiği olaya A ve B nin birleşimi (union) denir ve $A \cup B = \{x \in S : x \in A \text{ veya } x \in B\}$.

Örnek 1.19. Bir zar atma deneyinde, asal sayıların geldiği olay $A = \{2, 3, 5\}$ olmak üzere çift sayıların geldiği olay $B = \{2, 4, 6\}$ olmak üzere çift veya asal sayıların gelmesi olayı $A \cup B = \{2, 3, 4, 5, 6\}$

Örnek 1.20. A, B, C olaylarının diyargaramını aşağıdaki gibi verebiliriz.



Alıřtırmalar

- (1) Ařağıdaki olaylardan hangileri eřittir?
 - (i) $A = \{1, 3\}$
 - (ii) $B = \{x : x, \text{zarın üstündeki rakamlar}\}$
 - (iii) $C = \{x : x^2 - 4x + 3 = 0\}$
- (2) 2 adet zarın atılmasındaki örneklem uzayını veriniz.
 - (i) $A = \{x : x, \text{zarların sayılarının toplamı 8 den büyük ve eřit}\}$
 - (ii) $B = \{x : x, \text{zarlardan en az birinde 2 gelmesi}\}$
 - (iii) $C = \{x : x, \text{zarların sayıları 4 ten büyük ve eřit}\}$
 - (iv) Venn řeması ile bu olayları gösteriniz.
- (3) Bir deneyde önce bir zar atılıyor. Zarın üstündeki rakam çift olduğıunda 1 kez para atılıyor, tek olduğıunda ise para 2 kez atılıyor. Buna göre örneklem uzayını yazınız.
 - (i) $A = \{\text{Zarın üstündeki rakamlar 3 ten küçük ve eřit}\}$
 - (ii) $B = \{\text{Paranın 2 kez atılması}\}$
 - (iii) A' karşılık gelen olay
 - (iv) $A' \cap B$ olayı
 - (v) $A \cup B$ olayı
- (4) 3 nehir sularından alınan örneklerde, nehirde balık avlanan durumları E ile avlanamayan durumları H ile gösterilecektir. Buna göre örneklem uzayı belirtiniz.
 - (i) $A = \{\text{en az 2 nehirde balık avlanabilir}\}$
 - (ii) $\{EEE, HEE, EEH, HEH\}$ olayını sözlü olarak ifade ediniz. (2. nehirin avlanabilir olduğı durum)
- (5) Bkz: D:\Work\lectures\2013_2014\probabilty&statistic\literaturs\Probability & Statistics for Engineers & Scientists (9th Edition) - Walpole.pdf (page 43)

2 Sayma Yöntemleri

Pek çok durumda bir olayın olasılığını hesaplamak için örneklem uzayını listelemeye gerek duymayız. Bunun temel ilkesi çarpım kuralı diye adlandırılır. (multiplication rule)

Kural 1 (1). *Eğer bir etkinlik n_1 yol ile, 2. bir etkinlik ise n_2 yol ile hazırlanıyorsa toplamda bu etkinliğin hazırlanması $n_1.n_2$ kadar olur.*

Örnek 2.1. *2 adet zarın atılmasında örneklem uzayında toplam olarak kaç adet örnek vardır.*

Çözüm: 1. zarın atılması $n_1 = 6$ çeřit yolla yapılırken 1.zarın atılması da $n_2 = 6$ yolla gerçekleşmektedir. Buna göre 2 adet zar toplam olarak $n_1.n_2 = 36$ yolla gerçekleşir. \square

Örnek 2.2. *22 üyesi olan bir klüp, bir başkan bir de veznedar seçimi yapacaktır. Kaç farklı yollu seçim gerçekleşir.*

Çözüm: 22 kişiden 1 başkan seçilmesi $n_1 = 22$ çeşit yolla yapılırken, başkan seçilminden sonra geriye kalan 21 kişinin veznedar seçimi $n_2 = 21$ yolla gerçekleşmektedir. Buna göre bir başkan bir de veznedar seçimi toplam olarak $n_1.n_2 = 21 * 22 = 462$ yolla gerçekleşir. \square

Kural 2 (2). (Genelleşmiş çarpma kuralı) Eğer bir etkinlik n_1 yol ile, 2. bir etkinlik ise n_2 yol ile, 3. bir etkinlik ise n_k yol ile ve bu şekilde devam ederek k . bir etkinlik n_k yol ile hazırlanıyorsa toplamda bu etkinliğin hazırlanması $n_1.n_2.n_3 \dots n_k$ kadar olur.

Örnek 2.3. Bir sürücü kartında 2 farklı harf ve birincisi 0 olmayan 3 rakamdan oluşmaktadır. Buna göre kaç farklı şekilde bir kart belirlemesi gerçekleşir.

Çözüm: 29 harften 1 harfin seçilmesi $n_1 = 29$ çeşit yolla yapılırken, bir harfin seçilminden sonra geriye kalan 28 harfin seçimi $n_2 = 28$ yolla gerçekleşmektedir. 1. rakamın 0 olmaması 3. seçimin $n_3 = 9$, diğer 2 rakamın seçimleri ise $n_4 = 10, n_5 = 10$ şekilde gerçekleşir. Buna göre 2 farklı harf ve birincisi 0 olmayan 3 rakamdan oluşan sürücü kartının seçimi toplam olarak $n_1.n_2.n_3.n_4.n_5 = 29 * 28 * 9 * 10 * 10 = 730800$ yolla gerçekleşir. \square

Örnek 2.4. Sedat bilgisayar parçalarını kendisi toplayarak bir bilgisayar almayı düşünmektedir. Mikrodevre seçimini 2 farklı markadan, hard drive seçimini 4 farklı markadan, aksesuarları ise 5 farklı markadan seçebilir. Buna göre Sedat kaç farklı şekilde bir bilgisayar belirler.

Çözüm: Mikrodevre seçimi $n_1 = 2$ çeşit yolla yapılırken, hard drive seçimi $n_2 = 4$ yolla, aksesuar seçimi ise $n_3 = 5$ şekilde gerçekleşir. Buna göre Sedat $n_1.n_2.n_3 = 2 * 4 * 5 = 40$ yolla bir bilgisayar belirler. \square

Örnek 2.5. $\{0, 1, 2, 5, 6, 9\}$ rakamları kullanılarak 4 basamaklı çift sayı kaç farklı yolla oluşturulur.

Çözüm: Sayıyı $abcd$ olara gösterelim. d rakamı, sayı çift olacağından dolayı $\{0, 2, 6\}$ sayılarından biri olmalıdır. Ancak 0 rakamı a da kullanılamayacağından d nin 0 olması veya olmaması durumlarını 2 farklı yolla incelemeliyiz.

1.yol:

$$d = 0 \Rightarrow n_4 = 1$$

$$a : 5 \text{ farklı sayı olabilir} \Rightarrow n_1 = 5$$

$$b : 4 \text{ farklı sayı olabilir} \Rightarrow n_2 = 4$$

$$c : 3 \text{ farklı sayı olabilir} \Rightarrow n_3 = 3$$

Buna göre saymanın temel ilkesine göre

$$n_1.n_2.n_3.n_4 = 5 * 4 * 3 * 1 = 60$$

2.yol:

$$d \in \{2, 6\} \Rightarrow n_4 = 2$$

$$a \in \{1, 2 \vee 6, 5, 9\} \Rightarrow 4 \text{ farklı sayı olabilir} \Rightarrow n_1 = 4$$

$$b : 4 \text{ farklı sayı olabilir} \Rightarrow n_2 = 4$$

$$c : 3 \text{ farklı sayı olabilir} \Rightarrow n_3 = 3$$

Buna göre saymanın temel ilkesine göre

$$n_1.n_2.n_3.n_4 = 4 * 4 * 3 * 2 = 96$$

Buna göre toplamda $60 + 96 = 156$ farklı yolla 4 basamaklı çift sayı oluşturulur. \square

Tanım 2.6. Negatif olmayan n sayısı için $n!$ gösterimine n -faktoriyel denir. $0! = 1, n! = n \cdot (n-1) \cdot (n-2) \dots 3 \cdot 2 \cdot 1$

Tanım 2.7. n elemanın birbirinden farklı düzenlenmesine permütasyon denir.

Örnek 2.8. a, b, c harflerinden oluşan 3 farklı kombinasyon $abc, acb, bac, bca, cab, cba$ olmak üzere 6 farklı şekildedir. Veya kısaca $3! = 6$ olarak hesaplanır.

Tanım 2.9. n elemandan r tane alınarak yapılan permütasyonu

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

olarak tanımlarız.

Örnek 2.10. a, b, c, d harfleri kullanılarak 2 harfli kombinasyonlar $ab, ac, ad, ba, bc, bd, ca, cb, cd, da, db, dc$ olmak üzere 12 şekildedir. Veya saymanın ilkesine göre

$$n_1 n_2 = 4 * 3 = 12 \text{ veya permütasyon kullanılarak}$$

$$P_{4,2} = \frac{4!}{(4-2)!} = 12$$

Örnek 2.11. Matematik bölümünde, 25 kişilik bir sınıfta, araştırma, öğrenim ve geliştirme adlarında 3 adet farklı ödül verilmektedir. Her bir öğrenci en fazla 1 ödül alabildiğine göre, kaç olası seçim gerçekleşir?

Çözüm: Ödüller birbirinden bağımsız olduğundan, permütasyon problemidir:

$$P_{25,3} = \frac{25!}{(25-3)!} = 25 * 24 * 23 = 13\,800$$

\square

Örnek 2.12. 50 kişilik bir öğrenci klübünde bir başkan ve birde basın sorumlusu seçilecektir.

(i) Hiç bir kısırlama olmadan,

(ii) Atakan başkan seçilirse, seçimde oy kullanacaktır,

(iii) Sercan ve Ece birlikte seçilmesi veya seçilmemesi durumunda

(iv) Eslem veya Kubilayın birlikte seçilmemesi durumunda

kaç farklı seçim gerçekleşir?

Çözüm: (i)

$$P_{50,2} = \frac{50!}{(50-2)!} = 50 * 49 = 2450$$

(ii) Atakan başkan seçilirse, geriye 49 kişi kalmaktadır. Böylece 49 farklı olasılık vardır. Atakan başkan seçilmezse oy kullanmayacağından geriye 49 kişi kalmaktadır. Ve bu 49 kişiden 2 adet kişi

$$P_{49,2} = \frac{49!}{47!} = 2352$$

farklı yolla seçilir.

(iii) Sercan ve Ece birlikte seçilmesi durumu 2 şekildedir. Seçilmemesi durumunda ise seçim

$$P_{48,2} = \frac{48!}{46!} = 2256$$

farklı şekilde yapılır.

(iv) 1.Yol : Eslem başkan olduğunda, basından sorumlu aday seçimi Kubilay hariç 48 farklı yolla yapılır. Benzer şekilde Eslem basın sorumlusu olduğunda, başkanlık aday seçimi Kubilay hariç 48 farklı yolla yapılır. Buna göre Eslemin aday olması ve Kubilayın aday olmaması durumu $2 * 48 = 96$ farklı şekilde ifade edilirken aynı yolla Kubilayın aday olması ve Eslemin aday olmaması da $2 * 48 = 96$ farklı şekilde ifade edilir. Bununla birlikte ikisinin birlikte aday olmaması durumunda kalan 48 kişiden adaylık seçimleri

$$P_{48,2} = \frac{48!}{46!} = 2256$$

farklı şekilde yapılır. Buna göre toplamda $96 + 96 + 2256 = 2448$ farklı seçimle yapılır.

2.yol : Eslem ve :Kubilayın seçilmesi durumu 2 türdür. (Eslem:başkan, Kubilay:basın sorumlusu veya Eslem:basın sorumlusu, Kubilay:başkan). Toplam seçim $2450 - 2 = 2448$ türde olur. \square

Teorem 2.13. n elemanın dairesel olarak birbirinden farklı düzenlenmesi $(n - 1)!$ şekilde olur.

İspat: A,B,C,D objelerinin yuvarlak bir masada aşağıdaki gibi sıralandığını söyleyebiliriz.

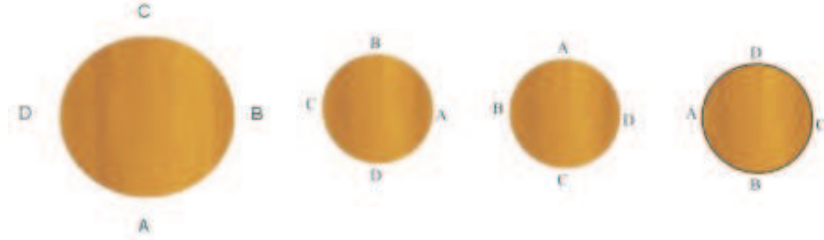


FIGURE 2.1. Denklemin grafiği

fakat her bir objenin birbirine göre göreceli pozisyonları aynıdır! Böylece n elemanın, n şekilde pozisyonu ayarlanabilir. K permütasyonu göstermek üzere, n farklı objenin dairesel olarak permütasyonunu

$$K.n = n!$$

olarak verebiliriz. Buna göre

$$K = \frac{n!}{n} = (n - 1)!$$

Örnek 2.14. 5 çocuk, yuvarlak bir masa etrafında kaç farklı şekilde oturur.

Çözüm: Dairesel permütasyona göre $(5 - 1)! = 24$ farklı şekilde. \square

Örnek 2.15. 3 Amerikalı, 4 Fransız, 4 Danimarkalı ve 2 İtalyan aynı uyruklular yan yana gelmemek koşulu ile yuvarlak bir masa etrafında kaç türlü oturabilir.

Çözüm: 4 farklı uyruk yuvarlak masa etrafında $(4 - 1)! = 6$ farklı şekilde sıralanırken, 3 Amerikalı 3! biçimde, 4 Fransız 4! biçimde, 4 Danimarkalı 4! biçimde ve 2 İtalyan 2! biçimde sıralanırlar. Buna göre çarpmanın temel ilkesine göre $6 * 3! * 4! * 4! * 2! = 41\,472$ şekilde farklı düzen oluşur. \square

Teorem 2.16. n_1 benzer, n_2 benzer, \dots, n_k benzer nesneleri olan n toplam nesnenin permütasyonu

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

olarak hesaplanır.

İspat: n_1 benzer nesnenin permütasyonu $n_1!$, n_2 benzer nesnenin permütasyonu $n_2!$, \dots, n_k benzer nesnenin permütasyonu $n_k!$ olarak hesaplanırken n nesnenin permütasyonu ise $n!$ olarak verilir. O halde sonuç permutasyonunu x ile gösterirsek:

$$n_1!n_2!\dots n_k!x = n! \Rightarrow x = \frac{n!}{n_1!n_2!\dots n_k!}$$

Örnek 2.17. DADDY harflerinden yararlanarak 5 harfli kaç türlü kelime türetilir?

Çözüm: 3 adet D harfi olduğundan D harfleri tekrarlanacaktır. Bu durumda permütasyon

$$\frac{5!}{3!} = 20$$

şekilde kelime oluşturulabilir. \square

Örnek 2.18. 28 farklı kitabı 6 öğrenci kendi arasında şu şekilde paylaşıyor. 2 öğrenci, herbirine 4'er kitap alacak şekilde seçerken geri kalan 4 öğrenci herbirine 5'er kitap seçmektedir. Buna göre bu 28 kitap kaç farklı yolda dağıtılır?

Çözüm: 1. ve 2. öğrenci için $4!, 3., 4., 5.$ öğrenciler için ise $5!$ farklı yolla kitap dağıtılabilirken toplamda

$$\frac{28!}{4! * 4! * 5! * 5! * 5!} = 306\,319\,469\,171\,683\,968\,000$$

değişik yolla kitap dağıtılabilir. \square

Örnek 2.19. Bir futbol takımında, teknik direktörün 10 oyuncuya istiyacı bulunmaktadır. Bunlardan 1' i 1.sınıftan, 2'si 2.sınıftan, 4'ü 3.sınıftan, 3'ü de 4.sınıftan seçilecektir. Kaç farklı şekilde 10 oyuncu seçilebilir?

Çözüm:

$$\frac{10!}{1! * 2! * 4! * 3!} = 12\,600$$

farklı yolda 10 oyuncu seçilebilir. \square

Örnek 2.20. 7 yüksek lisans öğrencisi, konferans boyunca 1 üç kişilik odada ve 2 adet çift kişilik odalarda kaç farklı şekilde kalabilir?

Çözüm: 7 yüksek lisans öğrencisi

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

farklı şekilde kalabilir \square

Örnek 2.21. Bir bridge masasında 52 kart 4 masaya eşit olarak kaç farklı durumda dağıtılır?

Çözüm: Her bir masaya 13 kart düşmektedir. Buna göre yinelemeli permütasyonu

$$\frac{52!}{13! * 13! * 13! * 13!} = 53\,644\,737\,765\,488\,792\,839\,237\,440\,000$$

olarak hesaplarız. □

Tanım 2.22. Birbirinden farklı n elemanın, r tanesi alınarak sıra gözetmeksizin yapılan seçimine kombinasyon denir ve $C_{n,r} = \binom{n}{r}$ ile gösterilir. Örneğin pokerde elinizde $A\spadesuit$ (karo – as), $5\heartsuit$ (kupa – 5), $7\clubsuit$ (sinek – 7), $10\spadesuit$ (maca – 10), $K\spadesuit$ (papaz – maca) olması ile $5\heartsuit$ (kupa – 5), $10\spadesuit$ (maca – 10), $7\clubsuit$ (sinek – 7), $K\spadesuit$ (papaz – maca), $A\spadesuit$ (karo – as) olması aynı anlama gelir.

Örnek 2.23. a, b, c, d 4 harften 3 ü alınarak elde edilen permütasyonlar

$$P_{4,3} = \frac{4!}{(4-3)!} = 24$$

adettir ve sırasıyla şöyledir: $\{abc, acb, bac, bca, cab, cba, abd, adb, bad, bda, dab, dba, acd, adc, cad, cda, dac, dca, bcd, bdc, cbd, cdb, dbc, dcb, dc b\}$. Halbuki bunlardan $\{abc, acb, bac, bca, cab, cba\}$ a, b, c elemanlarının farklı sıralanmasıyla, $\{abd, adb, bad, bda, dab, dba\}$ a, b, d elemanlarının farklı sıralanmasıyla, $\{acd, adc, cad, cda, dac, dca\}$ a, c, d elemanlarının farklı sıralanmasıyla, $\{bcd, bdc, cbd, cdb, dbc, dcb\}$ ise b, c, d elemanlarının farklı sıralanmasıyla elde edilmiştir. Kombinasyon ise bunların eşit olduğu ve bunlardan birinin alınmasının yeterli olduğu düşünülmelidir. Buna göre 4 harften 3 lü kombinasyonların kümesini $\{abc, abd, acd, bcd\}$ olarak düşünebiliriz.

Kombinasyonlar	Permütasyonlar
abc	$abc, acb, bac, bca, cab, cba$
abd	$abd, adb, bad, bda, dab, dba$
acd	$acd, adc, cad, cda, dac, dca$
bcd	$bcd, bdc, cbd, cdb, dbc, dcb$

Buna göre kombinasyonu $C_{4,3} * 3! = P_{4,3}$ olarak da verebiliriz.

Teorem 2.24.

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

İspat: n elemanlı S örneklem uzayında r farklı adet seçimi $C_{n,r}$ kombinasyonu ile yapılır. Seçilen bu elemanların sıralaması ise $r!$ biçimde gerçekleşir. Buna göre $r!C_{n,r}$ çarpımı, n elemanın r adet seçimine yani $P_{n,r}$ ye eşittir.

$$P_{n,r} = r!C_{n,r} \Rightarrow C_{n,r} = \frac{P_{n,r}}{r!} \Rightarrow C_{n,r} = \frac{\frac{n!}{(n-r)!}}{r!} = \frac{n!}{(n-r)!r!}$$

Örnek 2.25. Bir komitede, 9 kişilik bir gruptan 5 kişi kaç değişik yolla seçilir?

Çözüm:

$$C_{9,5} = \frac{9!}{4! * 5!} = 126$$

□

Örnek 2.26. 3 erkek ve 2 kızı içeren bir kuruk, 7 erkek ve 5 kız arasından kaç türlü seçilebilir?

Çözüm:

$$C_{7,5} * C_{5,2} = \frac{7!}{2! * 5!} * \frac{5!}{3! * 2!} = 210$$

□

Örnek 2.27. Bir okulda öğrenci birliği için her sene 4 üye seçiliyor.

(i) 12 öğrenci arasından birliğe kaç türlü öğrenci seçilebilir.

(ii) Uygun öğrencilerden 2 si birlikte seçilmediği takdirde kaç türlü üye seçilebilir.

Çözüm: (i)

$$C_{12,4} = \frac{12!}{8! * 4!} = 495$$

(ii) 1.yol: 2 öğrenciden birine A, diğerine B diyelim. A ve B nin seçime katılmadığı durumda 10 öğrenciden 4 seçim yapılacak demektir:

$$C_{10,4} = \frac{10!}{6! * 4!} = 210$$

A nı seçildiği B ninde seçime katılmadığı durumda 10 kişiden 3 kişi seçim yapılacaktır:

$$C_{10,3} = \frac{10!}{7! * 3!} = 120$$

Benzer şekilde B nı seçildiği A ninde seçime katılmadığı durumda 10 kişiden 3 kişi seçim yapılacaktır:

$$C_{10,3} = \frac{10!}{7! * 3!} = 120$$

Toplamda $210 + 120 * 2 = 450$ farklı yolla seçim yapılır.

(ii) 2.yol : 12 öğrenciden 4 kişilik seçim 495 türlü yapılabilirken. Hem A nın hemde B nin seçildiği durumda

$$C_{10,2} = \frac{10!}{8! * 2!} = 45$$

türlü gerçekleşir. Buna göre A veya B nin aynı anda seçilmediği durum $495 - 45 = 450$ dir.

□

Örnek 2.28. Bir öğrenci sınavda 10 sorudan 8 ini yanıtlayacaktır. (i) Kaç türlü seçeneği vardır. (ii) ilk 3 soruyu yanıtlama koşulu ile kaç türlü seçeneği gerçekleşir? (iii) ilk 5 sorudan en az 4 ünü yanıtlama koşulu ile kaç türlü seçeneği vardır?

Çözüm: (i)

$$C_{10,8} = \frac{10!}{2! * 8!} = 45$$

(ii) Eğer ilk 3 soruya yanıt vermişse geriye kalan 7 sorudan 5 soruyu cevaplamalıdır:

$$C_{7,5} = \frac{7!}{2! * 5!} = 21$$

(iii) İlk 5 sorudan 4 ünü yanıtlarsa bu soruları yanıtlama türü

$$C_{5,4} = \frac{5!}{1! * 4!} = 5$$

şekildedir. Geriye kalan 5 sorudan da 4 ünü yine $C_{5,4} = \frac{5!}{1! \cdot 4!} = 5$ farklı yolla yapar böylece $5 \cdot 5 = 25$ yolla seçimini yapabilir. Diğer yandan en az 4 dediği için ilk 5 sorudan 5 ini cevaplayabilir. Geriye kalan 5 sorudan 3 ünü cevaplaması ise

$$C_{5,3} = \frac{5!}{2! \cdot 3!} = 10$$

şekilde gerçekleşir. Buna göre toplamda $10 + 25 = 35$ seçenekte soruları cevaplayabilir. \square

Örnek 2.29. Bir öğretmen 6 öğrenci arasından en az 1 öğrenciyi kaç biçimde seçebilir?

Çözüm: 1.yol: 6 öğrenciden 1 i öğrenci, $C_{6,1} = \frac{6!}{5! \cdot 1!} = 6$ şekilde seçilir. 6 öğrenciden 2 i öğrenci, $C_{6,2} = \frac{6!}{4! \cdot 2!} = 15$ şekilde seçilir. 6 öğrenciden 3 i öğrenci, $C_{6,3} = \frac{6!}{3! \cdot 3!} = 20$ şekilde seçilir. 6 öğrenciden 4 i öğrenci, $C_{6,4} = \frac{6!}{2! \cdot 4!} = 15$ şekilde seçilir. 6 öğrenciden 5 i öğrenci, $C_{6,5} = \frac{6!}{1! \cdot 5!} = 6$ şekilde seçilir. 6 öğrenciden 6 i öğrenci, $C_{6,6} = \frac{6!}{0! \cdot 6!} = 1$ şekilde seçilir. Buna göre toplamda $6 + 15 + 20 + 15 + 6 + 1 = 63$ şekilde seçilir. 2.yol: 6 öğrencinin içeren kümenin toplam $2^6 = 64$ alt kümesi vardır. Buna göre hiç seçilmemesi yani \emptyset seçenği 1 adet olduğundan $64 - 1 = 63$ adet seçim yapılır. \square

Lemma 2.30. Her $1 \leq r \leq n$ için

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

İspat:

$$\begin{aligned} \binom{n}{r} + \binom{n}{r-1} &= \frac{n!}{(n-r)!r!} + \frac{n!}{(n-r+1)!(r-1)!} \\ &= \frac{n!(n-r+1)!(r-1)! + n!(n-r)!r!}{(n-r)!r!(n-r+1)!(r-1)!} \\ &= \frac{n!(n-r)!(r-1)![(n-r+1)+r]}{(n-r)!r!(n-r+1)!(r-1)!} \\ &= \frac{n![(n-r+1)+r]}{r!(n-r+1)!} \\ &= \frac{n![n+1]}{r!(n-r+1)!} = \frac{(n+1)!}{r!(n-r+1)!} \\ &= \frac{(n+1)!}{r!(n+1-r)!} = \binom{n+1}{r} \end{aligned}$$

Teorem 2.31. (Binom Teoremi) $\forall n \in \mathbb{N}, x, a \in \mathbb{R}$ için

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

İspat: Tümevarım yöntemi ile ispatlayalım. Öncelikle $n = 0$ ve $n = 1$ için doğruluğunu gösterelim:

(i) $n = 0$ olsun.

$$\begin{aligned}
 (x+a)^0 &= 1 \\
 \sum_{k=0}^0 \binom{n}{k} a^{n-k} x^k &= \binom{0}{0} a^0 x^0 \\
 \frac{0!}{0! * 0!} &= 1 \Rightarrow \\
 (x+a)^n &= \sum_{k=0}^n \binom{n}{k} a^{n-k} x^k, n=0
 \end{aligned}$$

(ii) $n = 1$ olsun.

$$\begin{aligned}
 (x+a)^1 &= x+a \\
 \sum_{k=0}^1 \binom{n}{k} a^{n-k} x^k &= \binom{1}{0} a^1 x^0 + \binom{1}{1} a^0 x^1 \\
 &= \frac{1!}{1! * 0!} a + \frac{1!}{0! * 1!} x \\
 &= a + x \Rightarrow \\
 (x+a)^n &= \sum_{k=0}^n \binom{n}{k} a^{n-k} x^k, n=1
 \end{aligned}$$

(iii) keyfi n için sağlansın. $(n+1)$ içinde doğruluğunu göstermeliyiz.

$$\begin{aligned}
 (x+a)^{n+1} &= (x+a)^n (x+a) \\
 &= \left[\sum_{k=0}^n \binom{n}{k} a^{n-k} x^k \right] (x+a) \\
 &= \sum_{k=0}^n \binom{n}{k} a^{n-k} x^{k+1} + \sum_{k=0}^n \binom{n}{k} a^{n-k+1} x^k
 \end{aligned}$$

$\sum_{k=0}^n \binom{n}{k} a^{n-k} x^{k+1}$ ifadesinde k yerine $k-1$ yazarsak

$$\begin{aligned}
 (x+a)^{n+1} &= \sum_{k=1}^{n+1} \binom{n}{k-1} a^{n-k+1} x^k + \sum_{k=0}^n \binom{n}{k} a^{n-k+1} x^k \\
 &= \sum_{k=1}^{n+1} \binom{n}{k-1} a^{n-k+1} x^k + \sum_{k=0}^n \binom{n}{k} a^{n-k+1} x^k \\
 &= \binom{n}{n} a^0 x^{n+1} + \sum_{k=1}^n \left[\binom{n}{k-1} + \binom{n}{k} \right] a^{n-k+1} x^k + \binom{n}{0} a^{n+1} x^0 \\
 &= \binom{n+1}{n+1} a^0 x^{n+1} + \sum_{k=1}^n \left[\binom{n}{k-1} + \binom{n}{k} \right] a^{n-k+1} x^k + \binom{n+1}{0} a^{n+1} x^0
 \end{aligned}$$

Yukarıdaki lemmayı kullanırsak

$$\begin{aligned}(x+a)^{n+1} &= \binom{n+1}{n+1} a^0 x^{n+1} + \sum_{k=1}^n \binom{n+1}{k} a^{n-k+1} x^k + \binom{n+1}{0} a^{n+1} x^0 \\ &= \sum_{k=0}^{n+1} \binom{n+1}{k} a^{n-k+1} x^k\end{aligned}$$

Böylece $n+1$ içinde sağlanmış olur.

Sonuç 2.32. $\forall n \in \mathbb{N}, a_1, a_2, \dots, a_k \in \mathbb{R}$ için

$$(a_1 + a_2 + \dots + a_k)^n = \sum_{n_1+n_2+\dots+n_k=n} \frac{n!}{n_1! * n_2! * \dots * n_k!} a_1^{n_1} a_2^{n_2} \dots a_k^{n_k}$$

Örnek 2.33. $(2x - y)^4$ ifadesinin açılımı nedir?

Çözüm:

$$\begin{aligned}(2x - y)^4 &= (2x + (-y))^4 \\ &= \binom{4}{0} (2x)^4 + \binom{4}{1} (2x)^3 (-y) + \binom{4}{2} (2x)^2 (-y)^2 + \binom{4}{3} (2x) (-y)^3 + \binom{4}{4} (-y)^4 \\ &= 16x^4 - 32x^3y + 24x^2y^2 - 8xy^3 + y^4\end{aligned}$$

□

Örnek 2.34. $(2x - 3y)^6$ ifadesinin açılımındaki x^3y^3 teriminin katsayısı nedir?

Çözüm: x^3y^3 teriminin katsayısı $k = 3$ durumudur.

$$\begin{aligned}\binom{6}{3} (2x)^3 (-3y)^3 &= \frac{6!}{3! * 3!} * 8x^3 * (-27) y^3 \\ &= -4320x^3y^3\end{aligned}$$

□

Alıştırma 2.35. $\forall n \in \mathbb{N}$, için

$$\binom{2n}{2} = 2 \binom{n}{2} + n^2$$

olduğunu gösteriniz.

Çözüm:

$$\begin{aligned}
 \binom{2n}{2} &= \frac{(2n)!}{(2n-2)!2!} = \frac{(2n)(2n-1)}{2!} \\
 &= 2n^2 - n = n(n-1) + n^2 \\
 &= 2 \frac{n(n-1)}{2!} + n^2 \\
 &= 2 \binom{n}{2} + n^2
 \end{aligned}$$

□

Alıştırma 2.36. $\forall n \in \mathbb{N}$, için

$$\binom{2n}{3} = 2 \binom{n}{3} + 2n \binom{n}{2}$$

olduğunu gösteriniz.

Çözüm:

$$\begin{aligned}
 2 \binom{n}{3} + 2n \binom{n}{2} &= 2 \frac{n(n-1)(n-2)}{3!} + 2n \frac{n(n-1)}{2!} \\
 &= \frac{4}{3}n^3 - 2n^2 + \frac{2}{3}n \\
 &= \frac{2n(2n^2 - 3n + 1)}{3} \\
 &= \frac{2n(2n-1)(n-1)}{3} \\
 &= \frac{2n(2n-1)(2n-2)}{3!} \\
 &= \binom{2n}{3}
 \end{aligned}$$

□

Alıştırma 2.37. $\forall n, r, k \in \mathbb{N}$, için

$$\binom{n}{r} \binom{r}{k} = \binom{n}{k} \binom{n-k}{r-k}$$

olduğunu gösteriniz.

Çözüm:

$$\begin{aligned}
\binom{n}{r} \binom{r}{k} &= \frac{n!}{(n-r)!r!} \frac{r!}{(r-k)!k!} \\
&= \frac{n!}{(n-r)!(n-k)!} \frac{(n-k)!}{(r-k)!k!} \\
&= \frac{n!}{(n-k)!k!} \frac{(n-k)!}{(n-r)!(r-k)!} \\
&= \binom{n}{k} \binom{n-k}{r-k}
\end{aligned}$$

□

Alıştırma 2.38. $\forall n, k \in \mathbb{N}$, için

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

olduğunu gösteriniz.

Çözüm:

$$\begin{aligned}
\binom{n-1}{k-1} + \binom{n-1}{k} &= \frac{(n-1)!}{(n-k)!(k-1)!} + \frac{(n-1)!}{(n-1-k)!k!} \\
&= \frac{(n-1)!}{(n-k)(n-k-1)!(k-1)!} + \frac{(n-1)!}{(n-1-k)!k(k-1)!} \\
&= \frac{(n-1)!}{(n-k-1)!(k-1)!} \left[\frac{1}{n-k} + \frac{1}{k} \right] \\
&= \frac{(n-1)!}{(n-k-1)!(k-1)!} \left[\frac{n}{(n-k)k} \right] \\
&= \frac{n!}{(n-k)!k!} \\
&= \binom{n}{k}
\end{aligned}$$

□

Alıştırma 2.39. $\forall n, k \in \mathbb{N}$, için

$$\sum_{k=0}^n \binom{n}{k} = 2^n$$

olduğunu gösteriniz.

Çözüm: Binom teoreminden

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

olduđuna göre $x = a = 1$ olarak alırsak

$$2^n = \sum_{k=0}^n \binom{n}{k}$$

elde ederiz. □

Alıştırma 2.40. Binom teoremını kullanarak 3^n ifadesini 2^n in kuvvetleri cinsinden yazınız.

Çözüm:

$$\begin{aligned} 3^n &= (2+1)^n = \sum_{k=0}^n \binom{n}{k} 1^{n-k} 2^k \\ &= \sum_{k=0}^n \binom{n}{k} 2^k \end{aligned}$$

□

3 Olasılık Aksiyomları

Tanım 3.1. Olasılık rastlantı veya kesin olmayan olaylarla uğraşır. Bir zar atma deneyinde zarın yere düşeceği kesindir ancak zarın yüzeyinde görünen rakamın 6 gelmesi kesin değildir. Olasılığı altında bir fonksiyon gibi tanımlayabiliriz. S örneklem uzayı ve $A \subset S$ örneklem uzayında bir olay olsun. Aşağıdaki koşulları sağlayan $P : A \rightarrow [0, 1]$ fonksiyonuna A olayının olasılığı denir, $P(A)$ ile gösterilir ve $P(A) = \frac{n(A)}{n(S)}$ ile tanımlanır.. Burada $n(A) = A$ 'da gerçekleşen olayların sayısı, $n(S) = S$ 'de gerçekleşen olayların sayısı

(i) $\forall A$ olayı için $0 \leq P(A) \leq 1$

(ii) $P(S) = 1$

(iii) Eğer A ve B kesikli olaylar ise $(A \cap B = \emptyset)$ $P(A \cup B) = P(A) + P(B)$

(iv) Eğer A_1, A_2, A_3, \dots kesikli olaylar ise $(\cap_{i=1}^{\infty} A_i = \emptyset)$ $P(A_1 \cup A_2 \cup A_3 \cup \dots) = P(A_1) + P(A_2) + \dots$

Örnek 3.2. Bir zar para iki kez atılıyor. En az bir yazı gelme olasılığı nedir?

Çözüm: $S = \{YY, YT, TY, TT\}$, $A = \{YY, YT, TY\}$ olmak üzere $P(A) = \frac{3}{4} = 75\%$ □

Örnek 3.3. Bir zar atma deneyinde çift rakamların gelme olasılığı, tek rakamlarınkinin 2 katıdır olarak tanımlarsak 4 ten düşük rakamların gelem olasılığı nedir?

Çözüm: $S = \{1, 2, 3, 4, 5, 6\}$ olmak üzere olasılığı x ile gösterelim. Buna göre tek rakamlar $\{1, 3, 5\}$ in toplam olasılığı $3x$ iken çift rakamların $\{2, 4, 6\}$ olasılığı $3 * 2x = 6x$ dir. Toplam olasılık 1 olması gerektiğine göre

$$9x = 1 \Rightarrow x = \frac{1}{9}$$

$A = \{1, 2, 3\} \Rightarrow 1ve3$ gelme olasılıkları $\frac{1}{9}$ iken 2 gelme olasılığı $\frac{2}{9}$. Olaslığın (iii) kuralına göre

$$\begin{aligned} P(A) &= P(\{1\}) + P(\{2\}) + P(\{3\}) \\ &= \frac{1}{9} + \frac{2}{9} + \frac{1}{9} = \frac{4}{9} \end{aligned}$$

□

Örnek 3.4. Bir zar atma deneyinde, A olayını çift rakamların gelmesi, B olayını da 3 e tam bölünebilen sayıların gelmesi olarak tanımlarsak $P(A), P(B)$ nedir. Çift rakam veya 3 e bölünebilme olasılığı ile çift ve 3 e bölünebilme olasılıkları nedir?

Çözüm: $S = \{1, 2, 3, 4, 5, 6\}$ olmak üzere, $A = \{2, 4, 6\}$, $B = \{3, 6\}$ olmak üzere

$$P(A) = \frac{3}{6} = 50\%, P(B) = \frac{2}{6} = \frac{1}{3}$$

Çift rakam veya 3 e bölünebilme olayı $A \cup B = \{2, 3, 4, 6\}$, çift ve 3 e bölünebilme olayı $A \cap B = \{6\}$ olmak üzere

$$P(A \cup B) = \frac{4}{6}, P(A \cap B) = \frac{1}{6}$$

□

Örnek 3.5. 25 endüstri mühendisi, 10 makine, 10 elektrik ve 8 inşaat mühendisliği öğrencilerinin bulunduğu istatistik sınıfında rastgele endüstri mühendisliğinde birinin seçilmesi olasılığı nedir? Elektrik veya inşaat mühendisliği seçilmesi olasılığı nedir?

Çözüm: $n(S) = 53, E = \{\text{endüstri} - \text{müh.} - \text{öğrencisi}\} \Rightarrow n(E) = 25$

$$P(E) = \frac{25}{53}$$

$B = \{\text{inşaat veya elektrik mühendisliği}\} \Rightarrow n(B) = 18$

$$P(B) = \frac{18}{53}$$

□

Örnek 3.6. Bir poker oyununda 5 karttan 2 sinin as 3 ünün vale gelme olasılığı nedir?

Çözüm: 4 renk astan 2 as 1

$$C_{4,2} = \frac{4!}{2! * 2!} = 6$$

değişik yolla belirleriz. Benzer şekilde 4 valeden 3 ünü

$$C_{4,3} = \frac{4!}{1! * 3!} = 4$$

yolla seçeriz. 52 karttan 5 ini ise

$$C_{52,5} = \frac{52!}{47! * 5!} = 2598960$$

farklı yolla seçilir. Buna göre 5 karttan 2 sinin as 3 ünün vale gelmesi çarpma ilkesine göre $6 * 4 = 24$ farklı yolla seçilir. $C = \{2\text{-as}, 3\text{-vale}\}$ olmak üzere

$$P(C) = \frac{6 * 4}{2598960} = \frac{24}{2598960} = 9.2345 \times 10^{-6}$$

□

Önerme 3.7. $P(A') = 1 - P(A)$

İspat: $S = A \cup A'$ ve $A \cap A' = \emptyset$ olduğundan olasılığın (ii) & (iii) özelliklerinden

$$\begin{aligned} P(S) &= P(A) + P(A') \\ 1 &= P(A) + P(A') \Rightarrow \\ P(A') &= 1 - P(A) \end{aligned}$$

Önerme 3.8. $P(\emptyset) = 0$

İspat: $S' = \emptyset$, yukarıdaki önermeden

$$\begin{aligned} P(S') &= P(\emptyset) = 1 - P(S) \\ &= 1 - 1 \text{ ((ii) den)} \\ &= 0 \end{aligned}$$

Önerme 3.9.

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

İspat:

$$\begin{aligned} A \cup B &= (A \cap B') \cup (A \cap B) \cup (A' \cap B) \\ \emptyset &= (A \cap B') \cap (A \cap B) \cap (A' \cap B) \Rightarrow (iii)' \text{ ten} \\ P(A \cup B) &= P(A \cap B') + P(A \cap B) + P(A' \cap B) \\ A &= (A \cap B) \cup (A \cap B'), \end{aligned} \tag{3.1}$$

$$\begin{aligned} \emptyset &= (A \cap B) \cap (A \cap B') \Rightarrow \\ P(A) &= P(A \cap B) + P(A \cap B') \Rightarrow \\ P(A \cap B') &= P(A) - P(A \cap B) \end{aligned} \tag{3.2}$$

$$\begin{aligned} B &= (A' \cap B) \cup (A \cap B) \\ \emptyset &= (A' \cap B) \cap (A \cap B) \Rightarrow \\ P(B) &= P(A' \cap B) + P(A \cap B) \Rightarrow \\ P(A' \cap B) &= P(B) - P(A \cap B) \end{aligned} \tag{3.3}$$

(3.1) - (3.2) - (3.3)'ten

$$\begin{aligned} P(A \cup B) &= P(A \cap B') + P(A \cap B) + P(A' \cap B) \\ &= P(A) - P(A \cap B) + P(A \cap B) + P(B) - P(A \cap B) \\ &= P(A) + P(B) - P(A \cap B) \end{aligned}$$

Önerme 3.10.

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) \\ &\quad - P(A \cap B) - P(A \cap C) - P(B \cap C) \\ &\quad + P(A \cap B \cap C) \end{aligned}$$

İspat: $D = B \cup C$ olarak yazarsak yukarıdaki önerme'den

$$\begin{aligned}
 P(A \cup D) &= P(A) + P(D) - P(A \cap D) \\
 P(D) &= P(B \cup C) = P(B) + P(C) - P(B \cap C) \\
 A \cap D &= A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \\
 P(A \cap D) &= P((A \cap B) \cup (A \cap C)) \\
 &= P(A \cap B) + P(A \cap C) - P((A \cap B) \cap (A \cap C)) \\
 &= P(A \cap B) + P(A \cap C) - P(A \cap B \cap C) \Rightarrow \\
 P(A \cup D) &= P(A) + P(D) - P(A \cap D) \\
 &= P(A) + (P(B) + P(C) - P(B \cap C)) \\
 &\quad - (P(A \cap B) + P(A \cap C) - P(A \cap B \cap C)) \\
 &= P(A) + P(B) + P(C) \\
 &\quad - P(A \cap B) - P(A \cap C) - P(B \cap C) \\
 &\quad + P(A \cap B \cap C)
 \end{aligned}$$

Örnek 3.11. Serhat endüstri mühendisliğinden bu dönem sonu mezun olacaktır. Beğendiği 2 şirketle görüştüğünden sonra, A şirketinden tekrar görüşmeye çağırılma oranını %80, B şirketinden çağırılma oranını ise %60 olarak düşünülmektedir. Her iki şirketten görüşmeye çağırılma olasılığını ise %50 olarak değerlendirmektedir. Buna göre en az bir şirketten çağırılma olasılığı nedir?

Çözüm:

$$\begin{aligned}
 A &= \{A \text{ şirketinden çağırılması}\} \Rightarrow P(A) = 0.8 \\
 B &= \{B \text{ şirketinden çağırılması}\} \Rightarrow P(B) = 0.6 \\
 P(A \cap B) &= 0.5 \Rightarrow \\
 P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\
 &= 0.8 + 0.6 - 0.5 = 0.9
 \end{aligned}$$

□

Örnek 3.12. Peşpeşe atılan bir zarda gelen rakamların toplamının 7 veya 11 gelme olasılığı nedir?

Çözüm: $n(S) = 6 * 6 = 36$,

$$\begin{aligned}
 A &= \{\text{rakamlar toplamı } 7\} = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\} \Rightarrow n(A) = 6 \Rightarrow P(A) = \frac{6}{36} \\
 B &= \{\text{rakamlar toplamı } 11\} = \{(5, 6), (6, 5)\} \Rightarrow n(B) = 2 \Rightarrow P(B) = \frac{2}{36}, A \cap B = \emptyset
 \end{aligned}$$

$$P(A \cup B) = \frac{6}{36} + \frac{2}{36} = \frac{8}{36} = \frac{2}{9}$$

□

Örnek 3.13. Bir kişinin yeşil, beyaz, kırmızı ve mavi araba seçme olasılığı sırasıyla 0.09, 0.15, 0.21, 0.23 olduğuna göre, galericinin bu renklerden birinden araba satma olasılığı nedir?

İspat: $Y = \{\text{yeşil araba satılması}\} \Rightarrow P(Y) = 0.09$

$B = \{\text{beyaz araba satılması}\} \Rightarrow P(B) = 0.15$

$K = \{\text{kırmızı araba satılması}\} \Rightarrow P(K) = 0.21$

$M = \{\text{mavi araba satılması}\} \Rightarrow P(M) = 0.23$

$$\begin{aligned} P(Y \cup B \cup K \cup M) &= P(Y) + P(B) + P(K) + P(M) \\ &= 0.09 + 0.15 + 0.21 + 0.23 \\ &= 0.68 \end{aligned}$$

Örnek 3.14. Bir araba teknisyeninin günde 3,4,5,6,7, 8 ve daha fazlası araba tamir etme olasılığı sırasıyla 0.12, 0.19, 0.28, 0.24, 0.10 ve 0.07 olmak üzere günde en az 5 araba tamir etme olasılığı nedir?

Çözüm:

$$P(3) + P(4) = 0.12 + 0.19 = 0.31$$

$$P(\{5 \text{ ten az araba tamir etme}\}) = 1 - 0.31 = 0.69$$

□

Örnek 3.15. 5 i yanmış olan 15 ampülden 3 ü rasgele seçiliyor. (i) Hiçbirinin yanmamış olma (ii) yalnız birinin yanmış olma (iii) en az birinin yanmış olma olasılıklarını bulunuz.

Çözüm: (i) 15 ampülden 3 ampül $C_{15,3} = \frac{15!}{12! \cdot 3!} = 455$ yolla seçilirken 10 sağlam ampülden 3 ünün seçilmesi $C_{10,3} = \frac{10!}{7! \cdot 3!} = 120$ yolla belirlenir. Buna göre

$$P(\{ \text{Hiçbirinin yanmamış olması} \}) = \frac{120}{455} = \frac{24}{91}$$

(ii) 5 i yanmış olan ampülden 1 ini 5 farklı yolla seçebiliriz. Geriye kalan 10 sağlam ampülden 2 si $C_{10,2} = \frac{10!}{8! \cdot 2!} = 45$ türlü seçilir. Buna göre

$$P(\{ \text{yalnız birinin yanmış olması} \}) = \frac{45}{455} = \frac{9}{91}$$

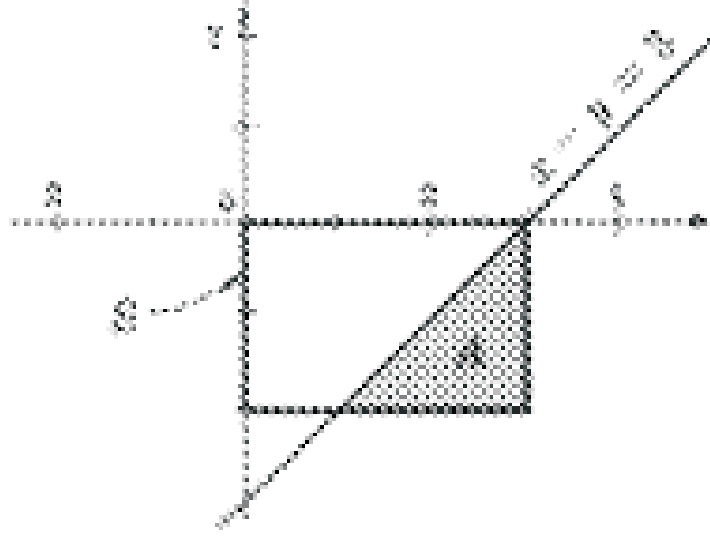
(iii) en az birinin yanmış olması tüm olasılıktan hiç yanmamış olmasının çıkarılmasıdır.

$$P(\{ \text{en az birinin yanmış olma} \}) = 1 - P(\{ \text{Hiçbirinin yanmamış olması} \}) = 1 - \frac{24}{91} = \frac{67}{91}$$

□

Örnek 3.16. \mathbb{R} reel sayılar düzelmünde a ve b noktaları için $-2 \leq b \leq 0, 0 \leq a \leq 3$ koşulu sağlanmaktadır. Buna göre a ve b arasındaki uzaklığın 3 ten büyük olma olasılığı nedir?

Çözüm: a ve b noktalarının sağladığı alanı S ile gösterirsek $n(S) = \text{alan}(S) = 3 \cdot 2 = 6$, $a - b \geq 3$ koşulunu sağlayan bölge ise A taralı üçgeni ve $x=1, y=2$ noktasında kesişir böylece $n(A) = \text{alan}(A) = \frac{2 \cdot 2}{2} = 2$. O halde $P(\{ a - b \geq 3, -2 \leq b \leq 0, 0 \leq a \leq 3 \}) = \frac{2}{6} = \frac{1}{3}$



□

Örnek 3.17. Bir zar 100 kez atılıyor. Aşağıdaki çizelgede 6 zar sayısının ve her gelen sayının sıklığı gösteriliyor. Buna göre (i) 3'ün gelme (ii) 5'in gelme (iii) çift sayı gelme (iv) asal sayı gelme (v) çift ve asal sayı gelmemesi olasılıklarını bulunuz.

SAyı	1	2	3	4	5	6
Sıklık	14	17	20	18	15	16

Çözüm: (i)

$$P(\{3 \text{ ün gelmesi}\}) = \frac{20}{100} = 0.2$$

(ii)

$$P(\{5 \text{ in gelmesi}\}) = \frac{15}{100} = 0.15$$

(iii)

$$\begin{aligned} P(\{ \text{çift sayı gelmesi} \}) &= P(\{2 \text{ ün gelmesi}\}) + P(\{4 \text{ ün gelmesi}\}) + P(\{6 \text{ ün gelmesi}\}) \\ &= \frac{17 + 18 + 16}{100} = \frac{51}{100} = 0.51 \end{aligned}$$

(iv)

$$\begin{aligned} P(\{ \text{asal sayı gelme} \}) &= P(\{2 \text{ ün gelmesi}\}) + P(\{3 \text{ ün gelmesi}\}) + P(\{5 \text{ ün gelmesi}\}) \\ &= \frac{17 + 20 + 15}{100} = \frac{52}{100} = 0.52 \end{aligned}$$

(v)

$$\begin{aligned} P(\{ \text{çift ve asal sayı gelmemesi} \}) &= P(\{2 \text{ ün gelmemesi}\}) \\ &= 1 - P(\{2 \text{ ün gelmesi}\}) \\ &= 1 - \frac{17}{100} = \frac{83}{100} = 0.83 \end{aligned}$$

□

Örnek 3.18. 25 kişilik bir sınıfta (i) en az 2 kişinin aynı doğum gününe sahip olma olasılığı (ii) en az bir kişinin seninle aynı doğum gününe sahip olma olasılığı nedir?

Çözüm: Saymanın temel ilkesine göre 25 kişilik bir sınıftaki kişilerin doğum günlerini olduğu örenklem uzayı için

$$n(S) = \underbrace{365 * 365 * 365 * \dots * 365}_{25-tane} = 365^{25}$$

25 kişinin her birinin farklı günlerde doğma olayını A ile gösterirsek

$$n(A) = 365 * 364 * 363 * \dots * (365 - (25 - 1)) = 365 * 364 * 363 * \dots * 341$$

Buna göre en az 2 kişinin aynı doğum gününe sahip olma olasılığı

$$\begin{aligned} P(A') &= 1 - P(A) = 1 - \frac{365 * 364 * 363 * \dots * 341}{365^{25}} \\ &= 1 - \frac{365!}{(365 - 25)! * 365^{25}} \\ &= 1 - 0.4313 = 0.5687 \end{aligned}$$

(ii) Sen hariç geri kalan 24 kişinin seninle aynı günde doğmama olayını B ile gösterirsek

$$n(B) = \underbrace{364 * 364 * 364 * \dots * 364}_{24-tane} = 364^{24}$$

buna göre en az bir kişinin seninle aynı doğum gününe sahip olma olasılığı

$$P(B') = 1 - P(B) = 1 - \frac{364^{24}}{365^{25}} = 0.99743$$

□

4 Koşullu Olasılık

Tanım 4.1. Bir B olayı, A olayından sonra gerçekleşiyorsa B olayının gerçekleşmesi olasılığına koşullu olasılık denir ve $P(B|A)$ ile gösterilir, A dan sonra B nin gerçekleşmesi olasılığı diye adlandırılır ve

$$P(B|A) = \frac{P(B \cap A)}{P(A)}, P(A) > 0$$

ile tanımlanır. Buna göre

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

ile de ifade edilir.

Örnek 4.2. Hileli bir zar atma deneyinde çift rakamların gelme olasılığı, tek rakamlarının 2 katıdır. Rakamların 3 ten büyük geldiği bilindiğine göre gelen sayının karekökü tam sayı olması olasılığı nedir?

Çözüm: $S = \{1, 2, 3, 4, 5, 6\}$ olmak üzere olasılığı x ile gösterelim. Buna göre tek rakamlar $\{1, 3, 5\}$ in toplam olasılığı $3x$ iken çift rakamların $\{2, 4, 6\}$ olasılığı $3 * 2x = 6x$ dir. Toplam olasılık 1 olması gerektiğine göre

$$9x = 1 \Rightarrow x = \frac{1}{9}$$

Rakamların 3 ten büyük gelmesi olayı $A = \{4, 5, 6\}$,

$$\begin{aligned} P(A) &= P(\{4\}) + P(\{5\}) + P(\{6\}) \\ &= \frac{2}{9} + \frac{1}{9} + \frac{2}{9} = \frac{5}{9} \end{aligned}$$

sayının karekökü tam sayı olması olayı ise $B = \{4\}$ olmak üzere $A \cap B = \{4\} \Rightarrow$

$$\begin{aligned} P(A \cap B) &= P(\{4\}) = \frac{2}{9} \Rightarrow \\ P(B|A) &= \frac{P(B \cap A)}{P(A)} = \frac{2/9}{5/9} = \frac{2}{5} \end{aligned}$$

□

Örnek 4.3.

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

İspat:

$$\begin{aligned} P(B|A) &= \frac{P(B \cap A)}{P(A)} \\ &= \frac{\frac{n(B \cap A)}{n(S)}}{\frac{n(A)}{n(S)}} = \\ &= \frac{n(B \cap A)}{n(A)} \end{aligned}$$

Örnek 4.4. İki hilesiz zar atılma deneyinde, eğer gelen iki sayının toplamı 6 ise, bir zarda gelen sayının 2 olma olasılığı nedir?

Çözüm: $A = \{ \text{iki sayının toplamı 6 olması} \} = \{(1, 5), (2, 4), (3, 3), (4, 2), (5, 1)\}$, $A \cap B = \{(2, 4), (4, 2)\}$

$$P(B|A) = \frac{n(B \cap A)}{n(A)} = \frac{2}{5}$$

□

Örnek 4.5. Düzenli sefer yapan bir uçuş için uçağın zamanında kalkma olasılığı 0.83, zamanında varma olasılığı ise 0.82 olup hem zamanında varması hemde zamanında kalkma olasılığı 0.78 dir. Buna göre (i) uçağın kalkış saati verildiğine göre zamanında varma (ii) varış saati bilindiğine göre zamanında kalkma olasılıklarını bulunuz.

Çözüm: $K = \{ \text{uçağın zamanında kalkması} \} \Rightarrow P(K) = 0.83$, $V = \{ \text{uçağın zamanında varması} \} \Rightarrow P(V) = 0.82$, $P(K \cap V) = 0.78$

$$\begin{aligned} (i) \quad P(V|K) &= \frac{P(V \cap K)}{P(K)} = \frac{0.78}{0.83} = 0.93976 \\ (ii) \quad P(K|V) &= \frac{P(V \cap K)}{P(V)} = \frac{0.78}{0.82} = 0.95122 \end{aligned}$$

□

Örnek 4.6. Şeydanın çalıştığı şirketin Apple distribütörlüğünü yapma olasılığı %30 olup, bunun gerçekleşmesi durumunda Şeydanın şirketin yöneticisi olması ise %60 olasılıktadır. Buna göre Şeydanın Apple distribütörlüğü yapan şirketin yöneticisi olması olasılığı nedir?

Çözüm: $A = \{ \text{şirketin Apple distribütörlüğünü yapma} \} \Rightarrow P(A) = 0.30$, $B = \{ \text{Şeydanın şirket yöneticisi olması} \} \Rightarrow P(B|A) = 0.60 \Rightarrow P(A \cap B) = ?$

$$\begin{aligned} P(A \cap B) &= P(A) P(B|A) \\ &= 0.30 * 0.60 \\ &= 0.18 \end{aligned}$$

□

Önerme 4.7. Herhangi A_1, A_2, \dots, A_n olayları için

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1) P(A_2|A_1) P(A_3|A_1 \cap A_2) \dots P(A_n|A_1 \cap A_2 \cap \dots \cap A_{n-1})$$

İspat:

$$\begin{aligned} P(A_1 \cap A_2 \cap \dots \cap A_n) &= P(A_1) \frac{P(A_1 \cap A_2)}{P(A_1)} \frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cap A_2)} \dots \frac{P(A_1 \cap A_2 \cap \dots \cap A_n)}{P(A_1 \cap A_2 \cap \dots \cap A_{n-1})} \\ &= P(A_1) P(A_2|A_1) P(A_3|A_1 \cap A_2) \dots P(A_n|A_1 \cap A_2 \cap \dots \cap A_{n-1}) \end{aligned}$$

Örnek 4.8. Bir kutuda 4 ü bozuk 12 nesne vardır. 3 nesne kutudan ard arda çekiliyor. 3 nesnenin de bozuk olmama olasılığı nedir?

Çözüm: $A_1 = \{1.\text{nesnenin seçilmesi}\}$, 12 nesneden 8 i bozuk olmadığı için $P(A_1) = \frac{8}{12}$,
 $A_2 = \{2.\text{nesnenin seçilmesi}\}$, kalan 11 nesneden 7 i bozuk olmadığı için $P(A_2|A_1) = \frac{7}{11}$,
 $A_3 = \{3.\text{nesnenin seçilmesi}\}$, kalan 10 nesneden 6 i bozuk olmadığı için $P(A_3|A_1 \cap A_2) = \frac{6}{10}$

$$\begin{aligned} P(A_1 \cap A_2 \cap A_3) &= P(A_1) P(A_2|A_1) P(A_3|A_1 \cap A_2) \\ &= \frac{8}{12} * \frac{7}{11} * \frac{6}{10} = \frac{14}{55} \end{aligned}$$

□

Önerme 4.9. A ve B $P(A) \cdot P(B) > 0$ koşulunu sağlayan olaylar olmak üzere

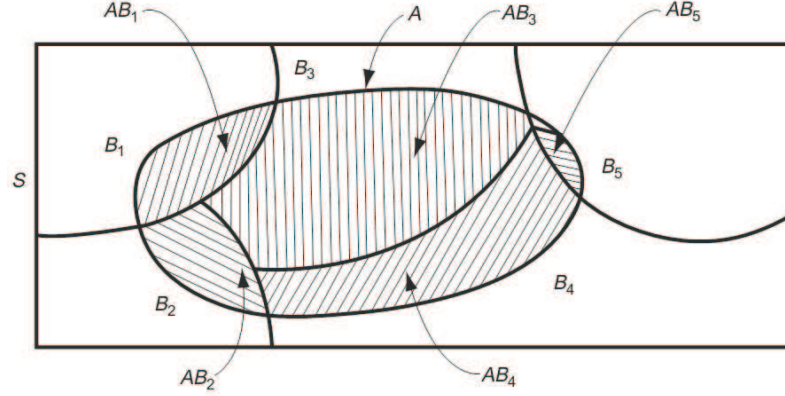
$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}$$

İspat:

$$\begin{aligned} P(A|B) &= \frac{P(A \cap B)}{P(B)}, \\ P(A \cap B) &= P(B|A) P(A) \end{aligned}$$

ifadesini yazdığımızda istenilen sonucu elde ederiz.

Tanım 4.10. A_1, A_2, \dots, A_n olayları için $S = \cup_{i=1}^n A_i$ ve $A_i \cap A_j = \emptyset, i \neq j$ koşulunu sağlıyorsa A_i olaylarına S örneklem uzayının parçalanması denir.



Önerme 4.11. A_1, A_2, \dots, A_n olayları S nin parçalanması ve $B \subset S$ olmak üzere

$$P(B) = \sum_{i=1}^n P(B|A_i) P(A_i)$$

İspat:

$$B = (B \cap A_1) \cup (B \cap A_2) \cup \dots \cup (B \cap A_n)$$

ve A_1, A_2, \dots, A_n olayları S nin parçalanması olduğundan $A_i \cap A_j = \emptyset, i \neq j$ koşulunu sağlar. O halde

$$(B \cap A_1) \cap (B \cap A_2) \cap \dots \cap (B \cap A_n) = \emptyset$$

olduğundan olasılık aksiyomundan

$$P(B) = P(B \cap A_1) + P(B \cap A_2) + \dots + P(B \cap A_n)$$

koşullu olasılık önermesinden:

$$P(B \cap A_i) = P(B|A_i) P(A_i), 1 \leq i \leq n$$

bu ifadeleri kullandığımızda

$$P(B) = \sum_{i=1}^n P(B|A_i) P(A_i)$$

sonucunu elde ederiz.

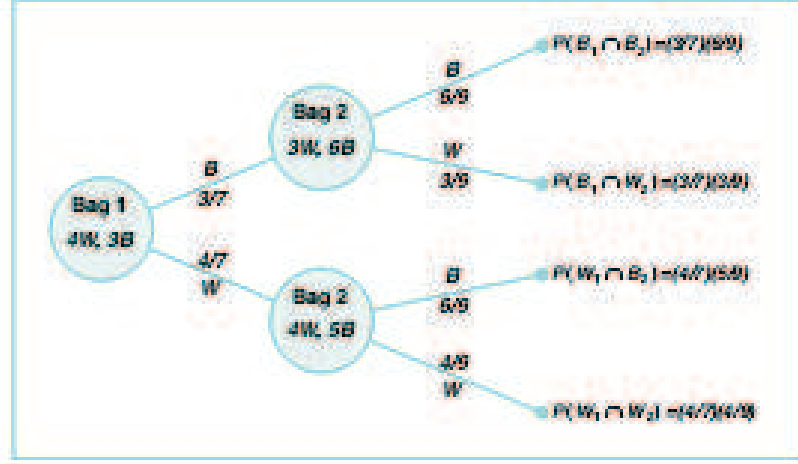
Örnek 4.12. Bir torbada 4 beyaz 3 siyah top bulunmakta, 2. torbada ise 3 beyaz 5 siyah top bulunmaktadır. 1. torbadan bir top seçildiğinde rengi görülmeden 2. torbaya atılıyor. Buna göre 2. torbadan siyah top çekilme olasılığı nedir?

Çözüm: $A_1 = \{1.\text{torbadan beyaz çekilmesi}\}, A_2 = \{1.\text{torbadan siyah çekilmesi}\}, A_1 \cap A_2 = \emptyset, B_1 = \{2.\text{torbadan siyah çekilmesi}\}$

$$B_1 = (A_1 \cap B_1) \cup (A_2 \cap B_1)$$

1.torbadan beyaz .çekilme olasılığı 4/7 ve çekilen beyaz 2. torbaya konulduğunda toplam 9 top olmaktadır.

buna göre 5 siyahtan birinin seçilmesi $5/9$ oranındadır. Benzer şekilde 1.torbadan siyah çekilme olasılığı $3/7$ olup çekilen siyah 2.tobaya atıldığında siyahların sayısı 6 olup toplam top sayısı 9 dur. buna göre 2. torbadan siyah seçilme olasılığı $6/9$ dur.



$$\begin{aligned}
 P(B_1) &= P(A_1) P(B_1|A_1) + P(A_2) P(B_1|A_2) \\
 &= \frac{4}{7} * \frac{5}{9} + \frac{3}{7} * \frac{6}{9} = \frac{38}{63}
 \end{aligned}$$

□

Örnek 4.13. A, B ve C gibi üç makine bir fabrikadaki ürünleri sırasıyla %50, %30 ve %20 sini üretmektedir. Bu makineler, %3, %4 ve %5 oranlarında bozuk ürün vermektedir. Buna göre rasgele seçilen bir ürünün bozuk olma olasılığı nedir?

Çözüm: $X = \{ \text{seçilen bir ürünün bozuk olması} \}$, $A = \{A \text{ makinesinin seçilmesi} \}$, $B = \{B \text{ makinesinin seçilmesi} \}$, $C = \{C \text{ makinesinin seçilmesi} \}$

$$\begin{aligned}
 P(X) &= P(A) P(X|A) + P(B) P(X|B) + P(C) P(X|C) \\
 &= 0.50 * 0.03 + 0.30 * 0.04 + 0.20 * 0.05 \\
 &= 0.037
 \end{aligned}$$

□

Örnek 4.14. Bir metalurji mühendisi A bölgesinde, şiddetli bir fırtına esnasında B_1 kasabasının %60 oranında, B_2 kasabasının %30, B_3 kasabasının ise %10 oranında zarar görebileceğini saptamıştır. Fırtına esnasında şiddetli akım oranı B_1 kasabasında yokken B_2 kasabasında %20 oranında B_3 kasabasında ise %60 oranında olması tahmin ediliyor. Buna göre A bölgesindeki şiddetli akım oranı tahmini olarak kaçtır?

Çözüm: $X = \{A \text{ bölgesindeki şiddetli akımı}\}$

$$\begin{aligned} P(X) &= P(B_1)P(X|B_1) + P(B_2)P(X|B_2) + P(B_3)P(X|B_3) \\ &= 0.60 * 0 + 0.30 * 0.20 + 0.10 * 0.60 \\ &= 0.12 \end{aligned}$$

□

Teorem 4.15. (Bayes Teoremi) A_1, A_2, \dots, A_n olayları S nin parçalanması, $P(A_i) > 0$, $1 \leq i \leq n$ ve $B \subset S$ olmak üzere

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{i=1}^n P(A_i)P(B|A_i)}$$

İspat:

$$\begin{aligned} P(A_i|B) &= \frac{P(A_i \cap B)}{P(B)} \\ P(A_i \cap B) &= P(A_i)P(B|A_i) \end{aligned}$$

ve yukarıdaki önermeden

$$P(B) = \sum_{i=1}^n P(B|A_i)P(A_i)$$

ifadelerini yerine yazdığımızda

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{i=1}^n P(A_i)P(B|A_i)}$$

sonucunu elde ederiz.

Örnek 4.16. A, B ve C gibi üç makine bir fabrikadaki ürünleri sırasıyla %50, %30 ve %20 sini üretmektedir. Bu makineler, %3, %4 ve %5 oranlarında bozuk ürün vermektedir. Buna göre rasgele seçilen bir ürünün zouk olduğu bilindiğine göre A makinesinde üretilme olasılığı nedir?

Çözüm: $X = \{\text{seçilen bir ürünün bozuk olması}\}$, $A = \{A \text{ makinesinin seçilmesi}\}$, $B = \{B \text{ makinesinin seçilmesi}\}$, $C = \{C \text{ makinesinin seçilmesi}\}$

$$\begin{aligned} P(A|X) &= ? \\ P(A) &= 0.50, P(X|A) = 0.03 \\ P(B) &= 0.30, P(X|B) = 0.04 \\ P(C) &= 0.20, P(X|C) = 0.05 \end{aligned}$$

Bayes formülüne göre

$$\begin{aligned} P(A|X) &= \frac{P(A)P(X|A)}{P(A)P(X|A) + P(B)P(X|B) + P(C)P(X|C)} \\ &= \frac{0.50 * 0.03}{0.50 * 0.03 + 0.30 * 0.04 + 0.20 * 0.05} \\ &= 0.40541 \end{aligned}$$

□

Örnek 4.17. Bir okuldaki öğrencilerden erkeklerin %4 ü, kızların %1 i 180cm den uzundur. Öğrencilerden %60 ı kızdır. Rasgele seçilen bir öğrenci 180 cm den uzun ise bu öğrencinin kız olması olasılığı nedir?

Çözüm: $A = \{180\text{cm'den uzun olması}\}$, $K = \{\text{kız öğrenci olması}\}$

$$\begin{aligned} P(K|A) &= ? \\ P(K|A) &= \frac{P(K) P(A|K)}{P(K) P(A|K) + P(E) P(A|E)} \\ &= \frac{0.60 * 0.01}{0.60 * 0.01 + 0.40 * 0.04} \\ &= 0.27273 \end{aligned}$$

□

Örnek 4.18. İletişim sistemlerinde bir sinyalin gönderilmesi 1 ile gönderilememesi ise 0 ile gösterilmektedir. E_1 sinyal yayıcısının seçilme olasılığı %60, E_2 sinyal yayıcısının ise seçilme olasılığı %40 tır. E_1 sinyal yayıcısının sinyali düzgün gönderme olasılığı %70 iken E_2 sinyal yayıcısının sinyali düzgün gönderme olasılığı %80 dir. İletişimde sinyalin varıldığı bilindiğine göre E_1 sinyal alıcısından olma olasılığı nedir?

Çözüm: $X = \{\text{Sinyalin başarılı gönderilmesi}\}$, $E_1 = \{E_1 \text{ sinyal yayıcısının sinyali}\}$, $E_2 = \{E_2 \text{ sinyal yayıcısının sinyali}\}$

$$\begin{aligned} P(E_1|X) &= ? \\ P(E_1|X) &= \frac{P(E_1) P(X|E_1)}{P(E_1) P(X|E_1) + P(E_2) P(X|E_2)} \\ &= \frac{0.60 * 0.70}{0.60 * 0.70 + 0.40 * 0.80} \\ &= 0.56757 \end{aligned}$$

□

5 Bağımsız Olaylar (Independent Events)

Tanım 5.1. A ve B olayları $P(B|A) = P(B)$ koşulunu sağlıyorsa bağımsız olaylardır denir.

Önerme 5.2. A ve B bağımsız olaylardır $\Leftrightarrow P(A \cap B) = P(A) P(B)$

İspat: A ve B bağımsız olaylar $\Leftrightarrow P(B|A) = P(B) \Leftrightarrow \frac{P(A \cap B)}{P(A)} = P(B) \Leftrightarrow P(A \cap B) = P(A) P(B)$

Önerme 5.3. A ve B bağımsız olaylardır $\Leftrightarrow P(A|B) = P(A)$

İspat: A ve B bağımsız olaylar $\Leftrightarrow P(A \cap B) = P(A) P(B) \Leftrightarrow \frac{P(A \cap B)}{P(B)} = P(A) \Leftrightarrow P(A|B) = P(A)$

Önerme 5.4. A ve B bağımsız olaylar ise

$$P(A \cup B) = P(B) + P(A) P(B') = P(A) + P(B) P(A')$$

olduğunu gösteriniz.

İspat:

$$\begin{aligned}
 P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\
 &= P(A) + P(B) - P(A)P(B) \\
 &= P(B) + P(A)(1 - P(B)) \\
 &= P(B) + P(A)P(B')
 \end{aligned}$$

Önerme 5.5. *A ve B olayları bağımsız ve $P(A) \neq 0, P(B) \neq 0$ koşullarını sağlıyorsa $A \cap B \neq \emptyset$ olduğunu gösteriniz.*

İspat: Varsayım: $A \cap B = \emptyset$ olsun. O halde $P(A \cap B) = 0$ dır. Ayrıca A ve B olayları bağımsız olduğundan $P(A \cap B) = P(A)P(B) = 0$. Bu ise $P(A) = 0$ veya $P(B) = 0$ demektir. Hipotezle çelişir.

Önerme 5.6. *A ve B bağımsız olaylar ise A' ve B' de bağımsız olaydır. Gösteriniz.*

İspat:

$$\begin{aligned}
 P(A' \cap B') &= P((A \cup B)') = 1 - P(A \cup B) \\
 &= 1 - (P(A) + P(B) - P(A \cap B)) \\
 &= 1 - P(A) - P(B) + P(A)P(B) \\
 &= (1 - P(A))(1 - P(B)) \\
 &= P(A')P(B')
 \end{aligned}$$

Örnek 5.7. $P(E) = 1/3, P(F|E) = 1/2, P(E|F) = 1/3$ olduğuna göre E ve F olaylarının bağımsız olaylar olduğunu gösteriniz.

Çözüm:

$$\begin{aligned}
 P(F|E) &= \frac{P(F \cap E)}{P(E)} \Rightarrow P(F \cap E) = \frac{1}{2} * \frac{1}{3} = \frac{1}{6} \\
 P(E|F) &= \frac{P(F \cap E)}{P(F)} \Rightarrow P(F) = \frac{1/6}{1/3} = \frac{1}{2} \\
 P(F \cap E) &= \frac{1}{6} = P(F) * P(E)
 \end{aligned}$$

olduğundan E ve F olaylarının bağımsız olaylardır. □

Örnek 5.8. *Bir kasabada itfaiye aracının ulaşılabilir olma olasılığı 0.98, ambulans aracının ulaşılabilir olması da 0.92 olasılıktadır. Bir yangın olayında ambulans ve itfaiyenin ulaşılabilir olması olasılığı nedir?*

Çözüm: $I = \{\text{itfaiye aracının ulaşılabilir olması}\}, A = \{\text{ambulans aracının ulaşılabilir olması}\}$

$$\begin{aligned}
 P(A \cap I) &= P(A)P(I) \\
 &= 0.98 * 0.92 \\
 &= 0.9016
 \end{aligned}$$
□

Örnek 5.9. Bir para 2 kez atılsın. İki kez de tura gelmesi olasılığı nedir?

Çözüm: $A = \{\text{İlk atışta tura gelmesi}\}, B = \{\text{İkinci atışta tura gelmesi}\}$ olmak üzere iki olay bağımsız olduğundan

$$P(A \cap B) = P(A) P(B) = \frac{1}{2} * \frac{1}{2} = \frac{1}{4}$$

Diğer yandan $S = \{YY, YT, TY, TT\}$ olmak üzere $P(\{TT\}) = 1/4$ aynı sonucu elde ederiz. \square

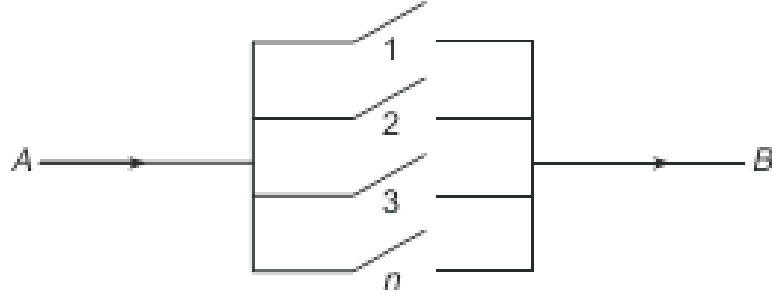
Örnek 5.10. İki zarın tek bir atılışında iki birli gelmesi olasılığı nedir?

Çözüm: $A = \{\text{İlk zar üzerinde 1 gelmesi}\}, B = \{\text{İkinci zar üzerinde 1 gelmesi}\}$ olmak üzere iki olay bağımsız olduğundan

$$P(A \cap B) = P(A) P(B) = \frac{1}{6} * \frac{1}{6} = \frac{1}{36}$$

Diğer yandan $S = \{(1, 1), (1, 2), \dots, (1, 6), (2, 1), (2, 2), \dots, (2, 6), \dots, (6, 1), (6, 2), \dots, (6, 6)\}$ olmak üzere $P(\{(1, 1)\}) = 1/36$ aynı sonucu elde ederiz. \square

Örnek 5.11. Şekildeki paralel devre sisteminde en az bir bileşenin kapalı olması durumunda sistem çalışmaktadır. Heri bileşenin çalışma olasılığı p_i olduğuna göre sistemin çalışması olasılığı nedir?



Çözüm:

$$\begin{aligned} P(\text{sistem çalışması}) &= 1 - P(\text{sistem çalışmıyor}) \\ &= 1 - P(\text{tüm bileşenler açık}) \\ &= 1 - P(A'_1 \cap A'_2 \cap \dots \cap A'_n) \\ &= 1 - P(A'_1) P(A'_2) \dots P(A'_n) \\ &= 1 - (1 - p_1) (1 - p_2) \dots (1 - p_n) \\ &= 1 - \prod_{i=1}^n (1 - p_i) \end{aligned}$$

\square

Örnek 5.12. A kutusunda 3 ü bozuk 8 nesne B kutusunda 2 si bozuk 5 nesne vardır. Her kutudan birer nesne seçiliyor. (i) Her iki nesneninde bozuk olma olasılığı (ii) birinin bozuk diğerinin sağlam olma olasılığı nedir?

Çözüm: $A = \{A \text{ kutusunda bozuk olması}\}$, $B = \{B \text{ kutusunda bozuk olması}\}$

(i)

$$\begin{aligned} P(A \cap B) &= P(A)P(B) \\ &= \frac{3}{8} * \frac{2}{5} = \frac{3}{20} \end{aligned}$$

(ii)

$$\begin{aligned} P(A \cap B') + P(A' \cap B) &= P(A)P(B') + P(A')P(B) \\ &= \frac{3}{8} * \frac{3}{5} + \frac{5}{8} * \frac{2}{5} \\ &= \frac{19}{40} \end{aligned}$$

□

Örnek 5.13. Aynı anda iki zar atalım. A olayı zarların üstünde gelen sayılar toplamının 11 olması, B olayı ise ilk zarda 5 gelmemesi olmak üzere A ve B olayları bağımsız olaylar mıdır?

Çözüm: $A = \{\text{zarların üstünde gelen sayılar toplamının 11}\} = \{(5, 6), (6, 5)\} \Rightarrow P(A) = 2/36$
 $B = \{\text{ilk zarda 5 gelmemesi}\} = \{(1, 1), \dots, (1, 6), (2, 1), \dots, (2, 6), (3, 1), \dots, (3, 6), (4, 1), \dots, (4, 6), (6, 1), \dots, (6, 6)\} \Rightarrow P(B) = 30/36$
 $A \cap B = \{(6, 5)\} \Rightarrow P(A \cap B) = 1/36$

$$\frac{1}{36} = P(A \cap B) \neq P(A)P(B) = \frac{2}{36} * \frac{30}{36} = \frac{5}{108}$$

□

Tanım 5.14. A_1, A_2, \dots, A_n olayları için

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1)P(A_2) \dots P(A_n)$$

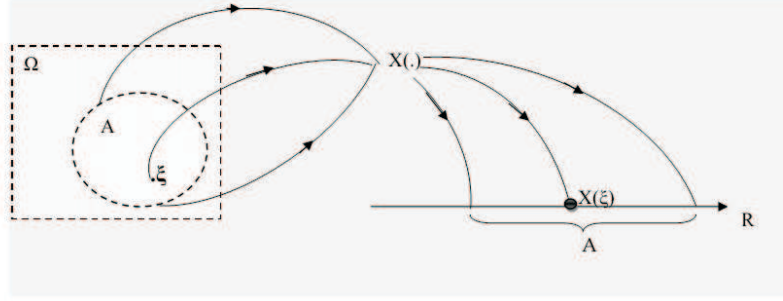
koşulları sağlanıyorsa bu olaylara bağımsız olaylar denir.

6 Rasgele Değişkenler (Random Variables)

Tanım 6.1. Örneklem uzayında her bir olaya karşılık bir sayının geldiğini varsayalım. Örneğin ard-arda atılan para deneyini ele alalım. Yazı gelmesi durumunu 1 ile tura gelmesi durumunu da 0 ile gösterirsek

Örneklem uzayı	YY	YT	TY	TT
	2	1	1	0

sayısal değerlerini elde edebiliriz. Böylece X rasgele değişkeninden söz ederken örneklem uzayındaki her bir elemana, reel sayı karşılık gelen bir fonksiyon anlaşılır. Eğer $X : S \rightarrow \mathbb{R}$ reel değerli fonksiyonu ve $x \in \mathbb{R}$ reel sayısı için $\{s : X(s) \leq x\}$ bir olay tanımlıyorsa $X(s)$ fonksiyonuna rasgele değişkeni denir.



Tüm rasgele değişkenleri X, Y, Z gibi büyük harfler ile gösterilirken, onlara karşılık gelen reel sayılar x, y, z gibi küçük harfler ile gösterilecektir.

Örnek 6.2. Bir depo görevlisi 3 kaskı sırasıyla üzerlerinde isimleri yazacak şekilde Uğur, Murat ve Tariğa vermektedir. Bu kişilerin kasklarını doğru almaları durumundaki rasgele değişkenleri belirleyiniz.

Çözüm: Doğru eşleme durumunu 1 aksi durumu 0 ile gösterelim.

Örneklem uzayı(s)	UMT	UTM	TUM	TMU	MUT	MTU
Doğru eşleme($X(s)$)	3	1	0	1	1	0

Örneğin $\{s : X(s) \leq 1\}$ bir olay berirler (en az bir doğru eşlenmesi olayı) □

Örnek 6.3. Bir kargo şirketinde ürünlerin hasarlı olup olmaması durumunu sırasıyla 1 ve 0 ile gösterirsek rasgele seçilen bir ürünün rasgele değişkenini belirleyiniz.

Çözüm: X ürünün rasgele seçilmesi olmak üzere

$$X = \begin{cases} 1, & \text{ürün hatalı ise} \\ 0, & \text{ürün hatalı değilse} \end{cases}$$

rasgele değişkeni 0 ve 1 olanlara Bernoulli rasgele değişkeni denir. □

Örnek 6.4. 12 si hasarlı 100 parça ürün içinden 10 ürün seçilecektir. X hasarlı ürün seçilmesi durumunun rasgele değişkeni 0,1,2,3,...,9,10 sayılarından biridir.

Örnek 6.5. Bir çağrı merkezi kişilerin taleplerini 0 ile 1 arasındaki sayılar ile sıralayarak cevaplandırmaktadır. X, kişinin talep oranı olmak üzere rasgele değişkeni $[0, 1]$ kapalı aralığıdır.

Tanım 6.6. Bir rasgele değişkeni sayılabilir sayıda reel sayıya karşılık geliyorsa kesikli rasgele değişkeni (discrete random variable) denir. Eğer bir rasgele değişkeni reel sayılarda bir aralığa karşılık geliyorsa sürekli rasgele değişkeni (continuous random variable) denir.

Örnek 6.7. 12 si hasarlı 100 parça ürün içinden 10 ürün seçilecektir. X hasarlı ürü seçilmesi durumunun rasgele değişkeni 0,1,2,3,...,9,10 sayılarından biridir. Buna göre X rasgele değişkeni kesikli rasgele değişkenidir. Bununla birlikte Bir çağrı merkezi kişilerin taleplerini 0 ile 1 arasındaki sayılar ile sıralayarak cevaplandırmaktadır. X, kişinin talep oranı olmak üzere rasgele değişkeni $[0, 1]$ kapalı aralığıdır. Buna göre X rasgele değişkeni sürekli rasgele değişkenidir.

Tanım 6.8. X kesikli rasgele değişkenine karşılık gelen sayısal değerler a_1, a_2, \dots olmak üzere, X kesikli rasgele değişkeninin olasılık yoğunluk fonksiyonu (bazen sadece olasılık fonksiyonu veya olasılık dağılım fonksiyonu da denir) (probability function- probability mass function - pmf- probability distribution function - pdf) aşağıdaki koşulları sağlayan $f : \mathbb{R} \rightarrow [0, 1]$, $f(a) = P(\{s \in S : X(s) = a\})$ fonksiyonudur.

(i) $f(a_i) \geq 0, i = 1, 2, 3, \dots$

(ii) $f(a_1) + f(a_2) + \dots = 1$

(iii) $a \neq a_i, f(a_i) = 0$

Örnek 6.9. Bir zarın olasılık yoğunluk fonksiyonunu bulunuz.

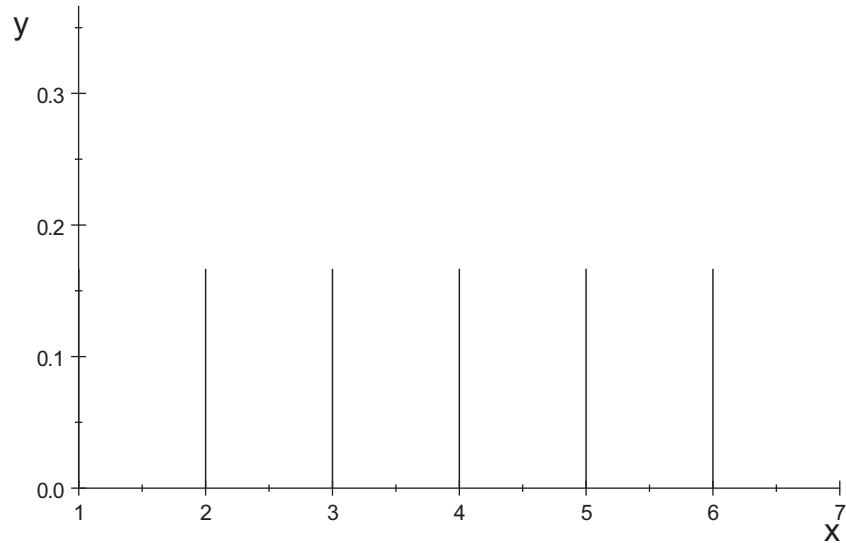
Çözüm:

Örneklem uzayı	1	2	3	4	5	6
	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

$$f(1) = P(X = 1) = \frac{1}{6}, f(2) = P(X = 2) = \frac{1}{6}$$

$$f(3) = P(X = 3) = \frac{1}{6}, f(4) = P(X = 4) = \frac{1}{6}$$

$$f(5) = P(X = 5) = \frac{1}{6}, f(6) = P(X = 6) = \frac{1}{6}$$



□

Örnek 6.10. 3 kez ardardına atılan bir para atma deneyinde yazı gelmesi durumunu 1 ile tura gelmesi durmunu da 0 ile gösterirsek X kesikli rasgele değişkenini ve bunların olasılık yoğunluk fonksiyonlarını bulunuz.

Çözüm:

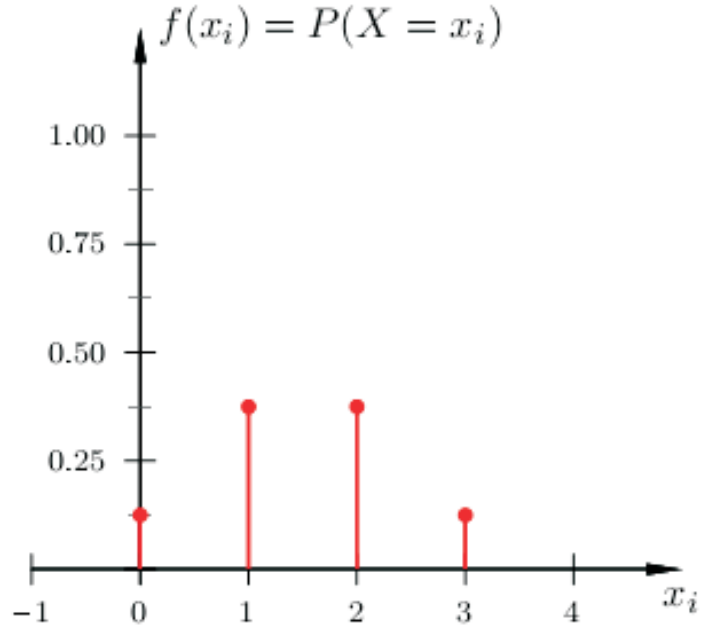
Örneklem uzayı	YYY	YYT	YTY	TTY	YTT	TYT	TTY	TTT
	3	2	2	2	1	1	1	0

$$f(0) = P(X=0) = P(TTT) = \frac{1}{2} * \frac{1}{2} * \frac{1}{2} = \frac{1}{8}$$

$$f(1) = P(X=1) = P(YTT) + P(TYT) + P(TTY) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8}$$

$$f(2) = P(X=2) = P(YYT) + P(YTY) + P(TYY) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8}$$

$$f(3) = P(X=3) = P(YYY) = \frac{1}{8}$$



□

Örnek 6.11. 3 ü bozuk 20 laptop içinden rastgele 2 bilgisayar seçilecektir. Bozuk olma durumu için olasılık yoğunluk fonksiyonlarını bulunuz.

Çözüm:

$$f(0) = P(X=0) = \frac{C_{3,0} * C_{17,2}}{C_{20,2}} = \frac{\frac{3!}{3!*0!} * \frac{17!}{15!*2!}}{\frac{20!}{18!*2!}} = \frac{68}{95}$$

$$f(1) = P(X=1) = \frac{C_{3,1} * C_{17,1}}{C_{20,2}} = \frac{\frac{3!}{2!*1!} * \frac{17!}{16!*1!}}{\frac{20!}{18!*2!}} = \frac{51}{190}$$

$$f(2) = P(X=2) = \frac{C_{3,2} * C_{17,0}}{C_{20,2}} = \frac{\frac{3!}{1!*2!} * \frac{17!}{17!*0!}}{\frac{20!}{18!*2!}} = \frac{3}{190}$$

□

Tanım 6.12. X kesikli rasgele değişkenine karşılık gelen sayısal değerler a_1, a_2, \dots olmak üzere, X kesikli rasgele değişkeninin birikimli olasılık dağılım fonksiyonu (cumulative distribution function) aşağıdaki koşulları sağlayan $F : \mathbb{R} \rightarrow [0, 1], F(a) = P(\{s \in S : X(s) \leq a\})$ fonksiyonudur.

$$(i) F(a_i) \geq 0, i = 1, 2, 3, \dots$$

$$(ii) F(a_1) + F(a_2) + \dots = 1$$

Olasılık dağılım fonksiyonu ile birikimli olasılık dağılım fonksiyonu arasında aşağıdaki eşitlik doğrudur:

$$F(a) = \sum_{a_i \leq a} f(a_i)$$

Uyarı 6.13.

$$\lim_{a \rightarrow \infty} F(a) = 1$$

$$\lim_{a \rightarrow -\infty} F(a) = 0$$

$$a \leq b \Rightarrow F(a) \leq F(b)$$

$$F(a) = F(a^+) = \lim_{\varepsilon \rightarrow 0} F(a + \varepsilon)$$

Örnek 6.14. Bir zarın birikimli olasılık dağılım fonksiyonunu bulunuz.

Çözüm:

Örneklem uzayı	1	2	3	4	5	6
	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

$$F(1) = P(X \leq 1) = \frac{1}{6},$$

$$F(2) = P(X \leq 2) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6}$$

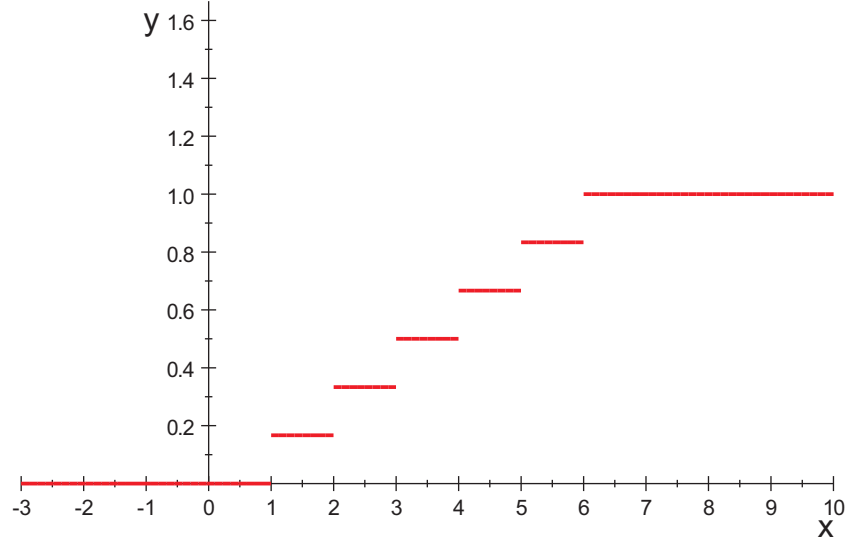
$$F(3) = P(X \leq 3) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6},$$

$$F(4) = P(X \leq 4) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{4}{6}$$

$$F(5) = P(X \leq 5) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{5}{6},$$

$$F(6) = P(X \leq 6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{6}{6}$$

$$F(x) = \begin{cases} 0 & \text{if } x < 1 \\ \frac{1}{6} & \text{if } 1 \leq x < 2 \\ \frac{2}{6} & \text{if } 2 \leq x < 3 \\ \frac{3}{6} & \text{if } 3 \leq x < 4 \\ \frac{4}{6} & \text{if } 4 \leq x < 5 \\ \frac{5}{6} & \text{if } 5 \leq x < 6 \\ \frac{6}{6} & \text{if } 6 \leq x \end{cases} = \begin{cases} 0 & \text{if } x < 1 \\ \frac{i}{6} & \text{if } i \leq x < i+1, i = 1, 2, 3, 4, 5 \\ 1 & \text{if } x \geq 6 \end{cases}$$



□

Örnek 6.15. 3 kez arıdardına atılan bir para atma deneyinde yazı gelmesi durumunu 1 ile tura gelmesi durumunu da 0 ile gösterirsek X kesikli rasgele değişkeninin birikimli olasılık dağılım fonksiyonlarını bulunuz.

Çözüm:

Örneklem uzayı	YYY	YYT	YTY	TTY	YTT	TYT	TTY	TTT
	3	2	2	2	1	1	1	0

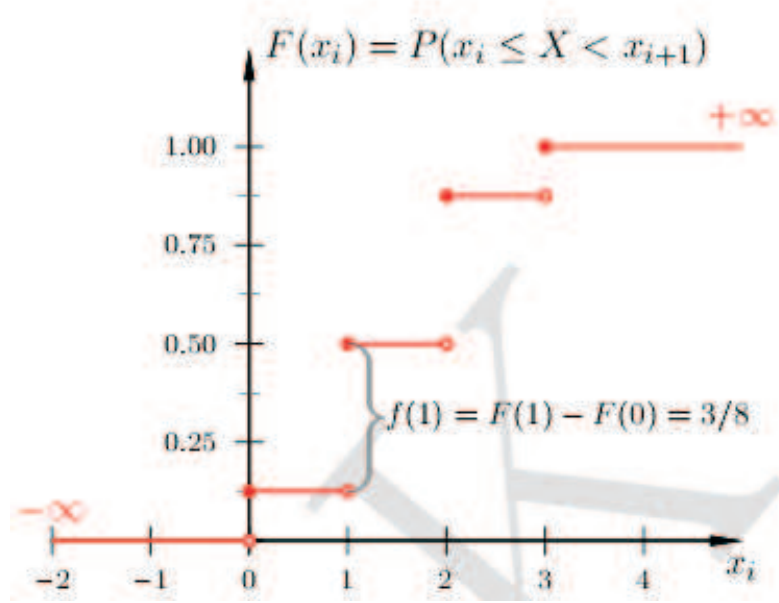
$$F(0) = P(X \leq 0) = P(TTT) = \frac{1}{2} * \frac{1}{2} * \frac{1}{2} = \frac{1}{8}$$

$$F(1) = P(X \leq 1) = P(YTT) + P(TYT) + P(TTY) + P(TTT) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{4}{8}$$

$$F(2) = P(X \leq 2) = P(YYT) + P(YTY) + P(TYY) + F(1) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{4}{8} = \frac{7}{8}$$

$$F(3) = P(X \leq 3) = P(YYY) + F(2) = \frac{1}{8} + \frac{7}{8} = 1$$

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{1}{8} & \text{if } 0 \leq x < 1 \\ \frac{4}{8} & \text{if } 1 \leq x < 2 \\ \frac{7}{8} & \text{if } 2 \leq x < 3 \\ 1 & \text{if } x \geq 3 \end{cases}$$



□

Örnek 6.16. X kesikli rasgele değişkeni için olasılık dağılımı aşağıdaki gibi verilmiştir.

$X=x$	-3	0	2	3
$P(X=x)$	0.2	0.1	0.4	c

Buna göre c değeri nedir? X in hangi değeri en büyük olasılığı alır? $P(X > 0)$ olasılığını bulunuz. $P(X = -2)$ olasılığını bulunuz. X kesikli rasgele değişkeninin olasılık yoğunluk fonksiyonunu ve birikimli olasılık dağılım fonksiyonunu bulunuz ve grafikte gösteriniz.

Çözüm:

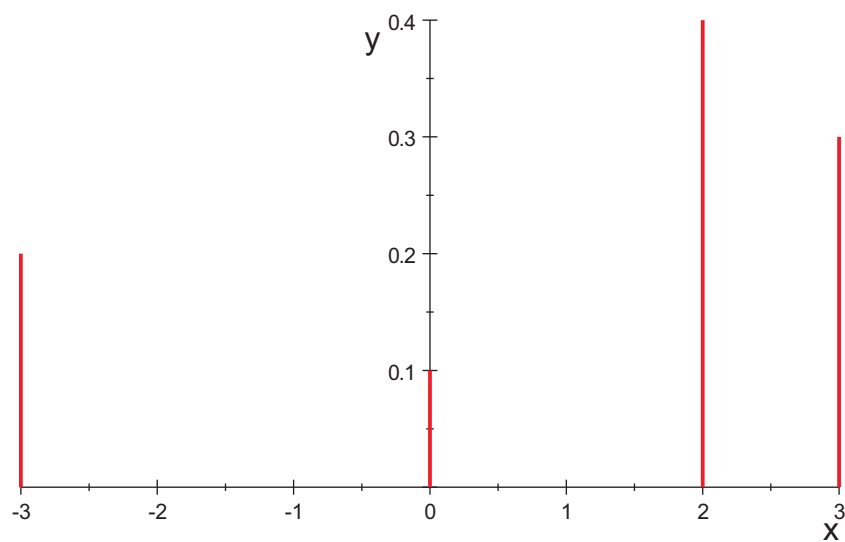
$$0.2 + 0.1 + 0.4 + c = 1 \Rightarrow c = 0.3$$

$$P(X = 0.2) = 0.4 \text{ olduğundan } X=0.2 \text{ de en büyük olasılık}$$

$$P(X > 0) = P(X = 2) + P(X = 3) = 0.4 + 0.3 = 0.7$$

$$P(X = -2) = 0$$

$$f(x) = \begin{cases} 0.2 & \text{if } x = -3 \\ 0.1 & \text{if } x = 0 \\ 0.4 & \text{if } x = 2 \\ 0.3 & \text{if } x = 3 \end{cases}$$



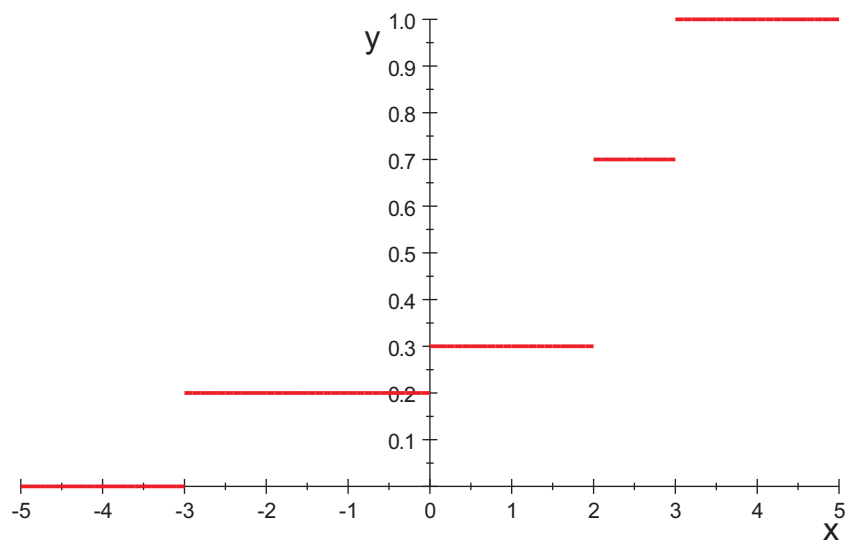
$$F(-3) = P(X \leq -3) = 0.2$$

$$F(0) = P(X \leq 0) = 0.1 + 0.2 = 0.3$$

$$F(2) = P(X \leq 2) = 0.4 + 0.1 + 0.2 = 0.7$$

$$F(3) = P(X \leq 3) = 0.3 + 0.4 + 0.1 + 0.2 = 1$$

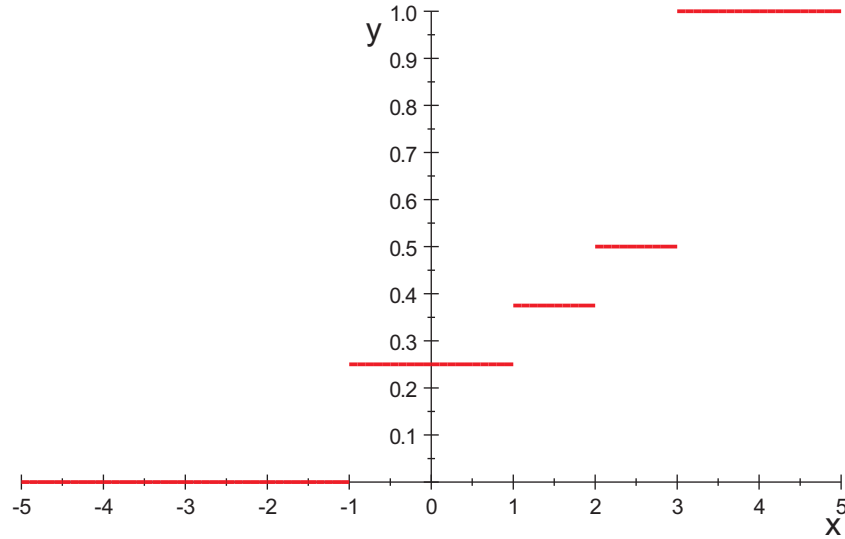
$$F(x) = \begin{cases} 0 & \text{if } x < -3 \\ 0.2 & \text{if } -3 \leq x < 0 \\ 0.3 & \text{if } 0 \leq x < 2 \\ 0.7 & \text{if } 2 \leq x < 3 \\ 1 & \text{if } x \geq 3 \end{cases}$$



Örnek 6.17. X rasgele değişkeni sırasıyla $-1, 1, 2, 3$ değerlerini almakta ve olasılıkları sırasıyla $1/4, 1/8, 1/8, 1/2$ değerleridir. Buna göre X kesikli rasgele değişkeninin birikimli olasılık dağılım fonksiyonunu bulunuz.

Çözüm:

$$\begin{aligned}
 f(-1) &= P(X = -1) = \frac{1}{4} \Rightarrow F(-1) = P(X \leq -1) = \frac{1}{4} \\
 f(1) &= P(X = 1) = \frac{1}{8} \Rightarrow F(1) = P(X \leq 1) = \frac{1}{4} + \frac{1}{8} = \frac{3}{8} \\
 f(2) &= P(X = 2) = \frac{1}{8} \Rightarrow F(2) = P(X \leq 2) = \frac{1}{4} + \frac{3}{8} = \frac{4}{8} \\
 f(3) &= P(X = 3) = \frac{1}{2} \Rightarrow F(3) = P(X \leq 3) = \frac{1}{2} + \frac{4}{8} = 1 \\
 F(x) &= \begin{cases} 0 & \text{if } x < -1 \\ \frac{1}{4} & \text{if } -1 \leq x < 1 \\ \frac{3}{8} & \text{if } 1 \leq x < 2 \\ \frac{4}{8} & \text{if } 2 \leq x < 3 \\ 1 & \text{if } x \geq 3 \end{cases}
 \end{aligned}$$



□

Önerme 6.18.

$$P(\{s \in S : a < X(s) \leq b\}) = F(b) - F(a)$$

İspat:

$$\begin{aligned}
 \{s \in S : X(s) \leq b\} &= \{s \in S : a < X(s) \leq b\} \cup \{s \in S : X(s) \leq a\} \\
 \emptyset &= \{s \in S : a < X(s) \leq b\} \cap \{s \in S : X(s) \leq a\} \Rightarrow \\
 P(\{s \in S : X(s) \leq b\}) &= P(\{s \in S : a < X(s) \leq b\}) + P(\{s \in S : X(s) \leq a\}) \\
 F(b) &= P(\{s \in S : a < X(s) \leq b\}) + F(a) \Rightarrow \\
 P(\{s \in S : a < X(s) \leq b\}) &= F(b) - F(a)
 \end{aligned}$$

Sonuç 6.19.

$$f(a) = F(a) - F(a^-), F(a^-) = \lim_{\varepsilon \rightarrow 0} F(a - \varepsilon)$$

İspat: Yukarıdaki ispatta a yerine $a - \varepsilon$ b yerine de a seçersek:

$$P(\{s \in S : a - \varepsilon < X(s) \leq a\}) = F(a) - F(a - \varepsilon),$$

$\varepsilon \rightarrow 0$ için limit alırsak

$$f(a) = P(\{s \in S : X(s) = a\}) = F(a) - F(a^-)$$

Tanım 6.20. X sürekli rasgele değişkeninin olasılık yoğunluk fonksiyonu (probability density function) aşağıdaki koşulları sağlayan $f : \mathbb{R} \rightarrow [0, 1], P(\{s \in S : a \leq X(s) \leq b\}) = \int_a^b f(x) dx$ fonksiyonudur.

$$(i) \quad f(x) \geq 0,$$

$$(ii) \quad \int_{-\infty}^{\infty} f(x) dx = 1$$

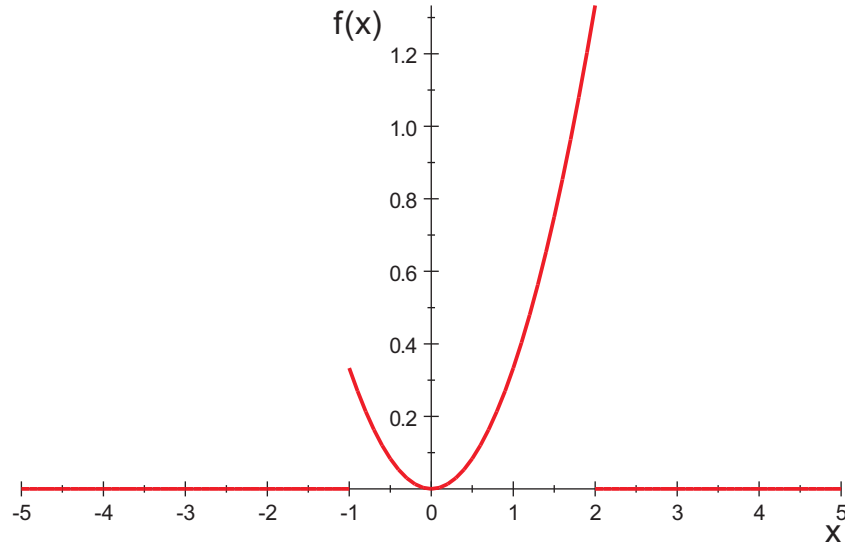
Örnek 6.21. Hava sıcaklığı ölçümündeki hataların sürekli rastgele değişkenler olarak belirtirsek bunların olasılık yoğunluk fonksiyonu

$$f(x) = \begin{cases} \frac{x^2}{3} & \text{if } -1 < x < 2 \\ 0 & \text{if } x \notin (-1, 2) \end{cases}$$

olarak veriliyor. Bu fonksiyonun olasılık yoğunluk fonksiyonu olduğunu kanıtlayınız ve $P(0 < X \leq 1)$ olasılığını hesaplayınız.

Çözüm: $f(x) \geq 0$ olduğu açıktır.

$$\begin{aligned}
 \int_{-\infty}^{\infty} f(x) dx &= \int_{-1}^2 \frac{x^2}{3} dx = \frac{1}{9} x^3 \Big|_{x=-1}^{x=2} = \frac{1}{9} (8 + 1) = 1 \\
 P(0 < X \leq 1) &= \int_0^1 f(x) dx = \int_0^1 \frac{x^2}{3} dx = \frac{1}{9} x^3 \Big|_{x=0}^{x=1} = \frac{1}{9}
 \end{aligned}$$



□

Örnek 6.22. X sürekli rasgele değişkeninin olasılık yoğunluk fonksiyonu

$$f(x) = \begin{cases} c(4x - 2x^2) & \text{if } 0 < x < 2 \\ 0 & \text{if } x \notin (-1, 2) \end{cases}$$

olarak veriliyor. c değerini hesaplayınız.

Çözüm:

$$1 = \int_{-\infty}^{\infty} f(x) dx = \int_0^2 c(4x - 2x^2) dx = c \left[2x^2 - \frac{2}{3}x^3 \right]_{x=0}^{x=2} = c \left(8 - \frac{16}{3} \right) = c \left(\frac{8}{3} \right) \Rightarrow c = \frac{3}{8}$$

□

Tanım 6.23. X sürekli rasgele değişkeninin birikimli olasılık dağılım fonksiyonu (cumulative distribution function) $F : \mathbb{R} \rightarrow [0, 1]$, $F(a) = P(\{s \in S : X(s) \leq a\}) = \int_{-\infty}^a f(t) dt$ fonksiyonudur. Olasılık yoğunluk fonksiyonu ile birikimli olasılık dağılım fonksiyonu arasında aşağıdaki eşitlik doğrudur:

$$\frac{dF(a)}{da} = f(a)$$

Uyarı 6.24. X sürekli rasgele değişkeninin birikimli olasılık dağılım fonksiyonu için de $P(\{s \in S : a < X(s) \leq b\}) = F(b) - F(a)$ özelliği doğrudur.

Örnek 6.25. Hava sıcaklığı ölçümündeki hataların sürekli rastgele değişkenler olarak belirtirsek bunların olasılık yoğunluk fonksiyonu

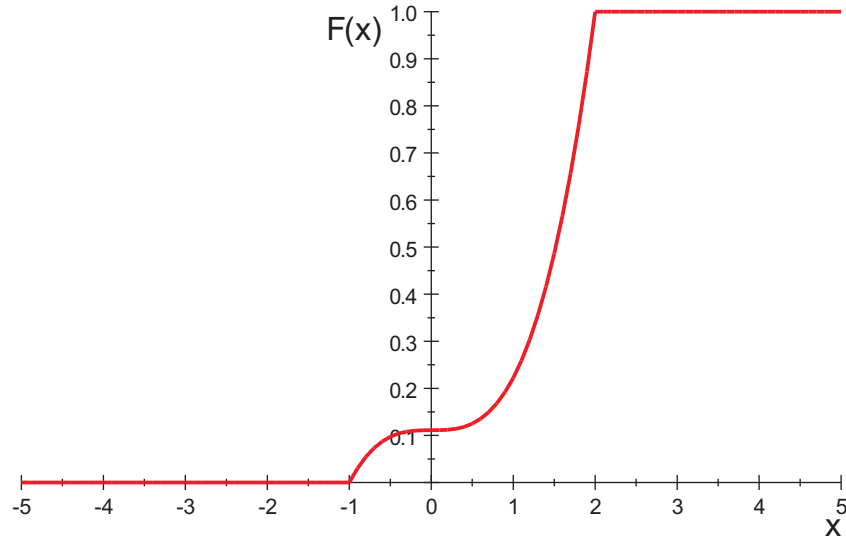
$$f(x) = \begin{cases} \frac{x^2}{3} & \text{if } -1 < x < 2 \\ 0 & \text{if } x \notin (-1, 2) \end{cases}$$

olarak veriliyor. Bu fonksiyonun birikimli olasılık dağılım fonksiyonunu hesaplayınız.

Çözüm:

$$\begin{aligned}
 (i) \quad a &\leq -1 \Rightarrow F(a) = \int_{-\infty}^a f(x) dx = \int_{-\infty}^a 0 dx = 0 \\
 (ii) \quad -1 &< a < 2 \Rightarrow F(a) = \int_{-\infty}^a f(x) dx = \int_{-1}^a \frac{x^2}{3} dx = \frac{1}{9} x^3 \Big|_{x=-1}^{x=a} = \frac{1}{9} (a^3 + 1) \\
 (iii) \quad a &\geq 2 \Rightarrow F(a) = \int_{-\infty}^a f(x) dx = \int_{-1}^2 \frac{x^2}{3} dx + \int_2^a 1 dx = \frac{1}{9} x^3 \Big|_{x=-1}^{x=2} + (a-2) = \frac{1}{9} (8+1) + (a-2) = 1 + (a-2) = a-1
 \end{aligned}$$

$$F(x) = \begin{cases} 0 & \text{if } x \leq -1 \\ \frac{1}{9} (x^3 + 1) & \text{if } -1 < x < 2 \\ 1 & \text{if } x \geq 2 \end{cases}$$



□

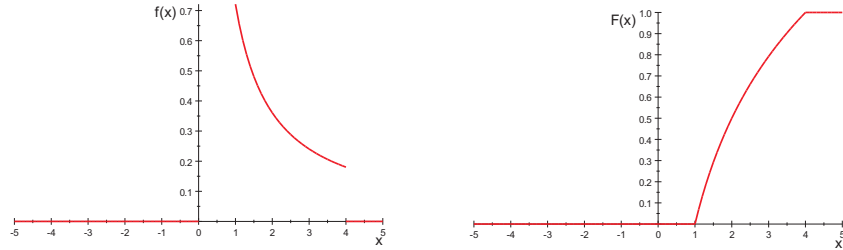
Örnek 6.26. X sürekli rasgele değişkeninin olasılık yoğunluk fonksiyonu

$$f(x) = \begin{cases} \frac{k}{x} & \text{if } 1 \leq x \leq 4 \\ 0 & \text{if } x \notin [1, 4] \end{cases}$$

olarak veriliyor. Birikimli olasılık dağılım fonksiyonunu hesaplayınız ve $P(X \leq 2.5)$ olasılığını hesaplayınız.

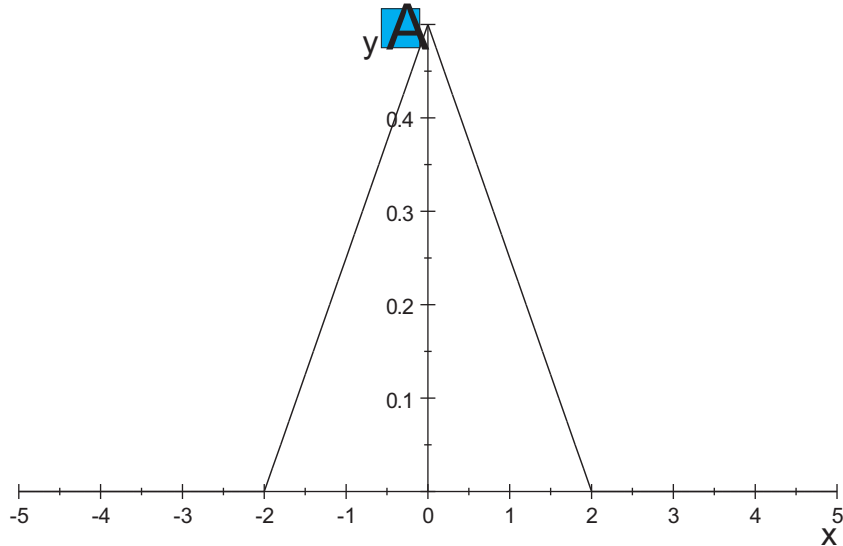
Çözüm:

$$\begin{aligned}
1 &= \int_{-\infty}^{\infty} f(x) dx = \int_1^4 \frac{k}{x} dx = k \ln x \Big|_{x=1}^{x=4} = k (\ln 4 - \ln 1) \Rightarrow k = \frac{1}{\ln 4} \\
F(x) &= \int_{-\infty}^x f(t) dt \\
(i) \ x < 1 &\Rightarrow F(x) = \int_{-\infty}^x f(t) dt = F(x) = \int_{-\infty}^x 0 dt = 0 \\
(ii) \ 1 \leq x \leq 4 &\Rightarrow F(x) = \int_{-\infty}^x f(t) dt = \int_1^x \frac{dt}{\ln 4t} = \frac{1}{\ln 4} \ln t \Big|_{t=1}^{t=x} = \frac{1}{\ln 4} \ln x \\
(iii) \ x > 4 &\Rightarrow F(x) = \int_{-\infty}^x f(t) dt = \int_1^4 \frac{dt}{\ln 4t} = 1 \\
F(x) &= \begin{cases} 0 & \text{if } x < 1 \\ \frac{1}{\ln 4} \ln x & \text{if } 1 \leq x \leq 4 \\ 1 & \text{if } x > 4 \end{cases} \\
P(X \leq 2.5) &= F(2.5) = \frac{1}{\ln 4} \ln(2.5) = 0.66096
\end{aligned}$$



□

Örnek 6.27. Sürekli bir X rasgele değişkeninin birikimli olasılık dağılım fonksiyonu aşağıdaki şekilde verilmiştir. $f(x)$ olasılık yoğunluk fonksiyonunu bulunuz. $P(X < 1/2)$ ve $P(X^2 > 1)$ olasılıklarını bulunuz.



Çözüm: Önce bilinmeyen A değerini bulalım: $-2 \leq x \leq 0$ için doğru

$$\frac{x}{-2} + \frac{y}{A} = 1 \Rightarrow y = A \left(1 + \frac{x}{2}\right)$$

$0 \leq x \leq 2$ için doğru

$$\frac{x}{2} + \frac{y}{A} = 1 \Rightarrow y = A \left(1 - \frac{x}{2}\right)$$

$$f(x) = \begin{cases} A \left(1 + \frac{x}{2}\right) & \text{if } -2 \leq x \leq 0 \\ A \left(1 - \frac{x}{2}\right) & \text{if } 0 \leq x \leq 2 \\ 0 & \text{if } x \notin [-2, 2] \end{cases}$$

$$\begin{aligned} \int_{-\infty}^{\infty} f(x) dx &= 1 \Rightarrow \int_{-2}^2 f(x) dx = 1 \Rightarrow \\ \int_{-2}^0 A \left(1 + \frac{x}{2}\right) dx + \int_0^2 A \left(1 - \frac{x}{2}\right) dx &= 1 \Rightarrow \\ A \left(x + \frac{x^2}{4}\right) \Big|_{x=-2}^{x=0} + A \left(x + \frac{x^2}{4}\right) \Big|_{x=0}^{x=2} &= 1 \Rightarrow \\ 2A &= 1 \Rightarrow A = \frac{1}{2} \Rightarrow \end{aligned}$$

$$f(x) = \begin{cases} \frac{1}{2} \left(1 + \frac{x}{2}\right) & \text{if } -2 \leq x \leq 0 \\ \frac{1}{2} \left(1 - \frac{x}{2}\right) & \text{if } 0 \leq x \leq 2 \\ 0 & \text{if } x \notin [-2, 2] \end{cases}$$

$$\begin{aligned} P(X < 1/2) &= \int_{-\infty}^{1/2} f(x) dx = \int_{-2}^0 f(x) dx + \int_0^{1/2} f(x) dx \\ &= \int_{-2}^0 \frac{1}{2} \left(1 + \frac{x}{2}\right) dx + \int_0^{1/2} \frac{1}{2} \left(1 - \frac{x}{2}\right) dx, \\ &= \frac{1}{2} + \frac{7}{32} = \frac{23}{32} \end{aligned}$$

$$\{X : X^2 > 1\} = \{X : X > 1\} \cup \{X : X < -1\} \Rightarrow$$

$$P(X^2 > 1) = P(X > 1) \cup P(X < -1)$$

$$P(X > 1) = \int_1^{\infty} f(x) dx = \int_1^2 \frac{1}{2} \left(1 - \frac{x}{2}\right) dx = \frac{1}{8}$$

$$P(X < -1) = \int_{-\infty}^{-1} f(x) dx = \int_{-2}^{-1} \frac{1}{2} \left(1 + \frac{x}{2}\right) dx = \frac{1}{8}$$

$$P(X^2 > 1) = \frac{1}{4}$$

□

7 Birleşik Olasılık Dağılımı (Joint Probability Distributions)

Şu ana kadar rasgele değişkeni tek değerliydi. Halbuki rasgele değişkeni çok değerli de olabilir. Örneğin kimyasal bir deneyde, hacim ile çökelti miktarını düşünersek örneklem uzayı ikililerden oluşur.

Tanım 7.1. X kesikli rasgele değişkenine karşılık gelen sayısal değerler x_1, x_2, \dots , Y kesikli rasgele değişkenine karşılık gelen sayısal değerler y_1, y_2, \dots olmak üzere X ve Y kesikli rasgele değişkenlerinin birleşik olasılık yoğunluk fonksiyonu (joint probability mass function - pmf) aşağıdaki koşulları sağlayan $f(x, y) = P(X = x, Y = y)$ fonksiyonudur.

(i) $f(x, y) \geq 0$

(ii) $\sum_x \sum_y f(x, y) = 1$

$A \subset \mathbb{R}^2$ bir bölge olmak üzere $P((X, Y) \in A) = \sum_{(x, y) \in A} f(x, y)$ olarak da verilir.

X/Y	y_1	y_2	\dots	y_n	Toplam
x_1	$f(x_1, y_1)$	$f(x_1, y_2)$	\dots	$f(x_1, y_n)$	$\sum_{j=1}^n f(x_1, y_j)$
x_2	$f(x_2, y_1)$	$f(x_2, y_2)$	\dots	$f(x_2, y_n)$	$\sum_{j=1}^n f(x_2, y_j)$
\vdots	\dots	\dots	\dots	\dots	\dots
x_m	$f(x_m, y_1)$	$f(x_m, y_2)$	\dots	$f(x_m, y_n)$	$\sum_{j=1}^n f(x_m, y_j)$
Toplam	$\sum_{i=1}^m f(x_i, y_1)$	$\sum_{i=1}^m f(x_i, y_2)$	\dots	$\sum_{i=1}^m f(x_i, y_n)$	$1 \rightarrow \text{Genel} - \text{Toplam}$

Örnek 7.2. 2 renkli kalem, 3 mavi, 2 kırmızı ve 3 yeşil kalemin olduğu bir kutudan seçilecektir. Eğer X rasgele değişkeni mavi kalem seçimini, Y ise kırmızı kalem seçimini gösterirse $f(x, y)$ birleşik olasılığı tanımlayınız ve $A = \{(x, y) : x + y < 1\}$ olmak üzere $P((X, Y) \in A)$ olasılığını bulunuz.

Çözüm: olası (x, y) değerleri $(0, 0) \Rightarrow$ ne kırmızı ne de mavi kalem seçimi,

$$f(0, 0) = \frac{C_{3,2}}{C_{8,2}} = \frac{\frac{3!}{1! \cdot 2!}}{28} = \frac{3}{28}$$

$(0, 1) \Rightarrow$ kırmızı kalem seçilirken mavi kalem seçilemez (yani 1 kırmızı 1 adet yeşil seçilir),

$$f(0, 1) = \frac{C_{2,1} * C_{3,1}}{C_{8,2}} = \frac{3}{14}$$

$(1, 0) \Rightarrow$ mavi kalem seçilirken kırmızı kalem seçilemez, (yani 1 mavi 1 adet yeşil seçilir),

$$f(1, 0) = \frac{C_{3,1} * C_{3,1}}{C_{8,2}} = \frac{9}{28}$$

$(1, 1) \Rightarrow$ 1 kırmızı 1 mavi kalem seçilir,

$$f(1, 1) = \frac{C_{2,1} * C_{3,1}}{C_{8,2}} = \frac{3}{14}$$

$(2, 0) \Rightarrow$ 2 mavi kalem seçilir,

$$f(2, 0) = \frac{C_{3,2}}{C_{8,2}} = \frac{3}{28}$$

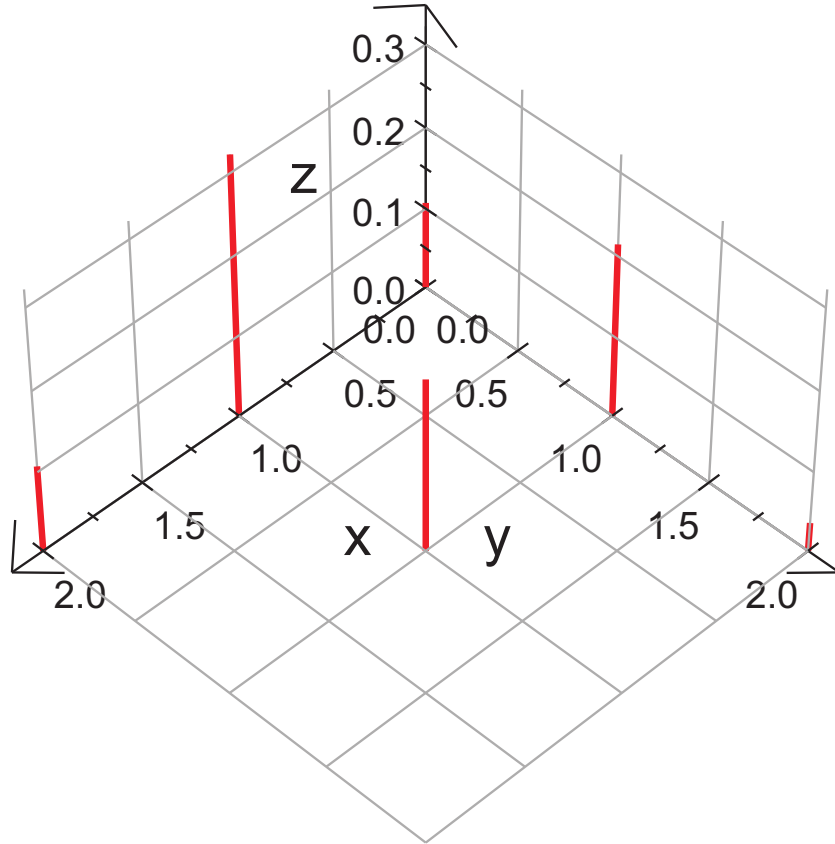
$(0, 2) \Rightarrow$ 2 kırmızı kalem seçilir.

$$f(0, 2) = \frac{C_{2,2}}{C_{8,2}} = \frac{1}{28}$$

$$f(x, y) = \frac{C_{3,x} C_{2,y} C_{3,2-(x+y)}}{C_{8,2}}, x = 0, 1, 2, y = 0, 1, 2, 0 \leq x + y \leq 2$$

olarak belirtebiliriz.

$f(x, y)$	x-mavi			Toplam
y-kırmızı	0	1	2	
0	$\frac{3!}{1! \cdot 2!} = \frac{3}{28}$	$\frac{3!}{2! \cdot 1!} \cdot \frac{3!}{2! \cdot 1!} = \frac{9}{28}$	$\frac{3!}{2! \cdot 1!} = \frac{3}{28}$	$\frac{3}{28} + \frac{9}{28} + \frac{3}{28} = \frac{15}{28}$
1	$\frac{2!}{1! \cdot 1!} \cdot \frac{3!}{2! \cdot 1!} = \frac{3}{14}$	$\frac{2!}{1! \cdot 1!} \cdot \frac{3!}{2! \cdot 1!} = \frac{3}{14}$	0	$\frac{3}{14} + \frac{3}{14} + 0 = \frac{3}{7}$
2	$\frac{2!}{0! \cdot 2!} = \frac{1}{28}$	0	0	$\frac{1}{28} + 0 + 0 = \frac{1}{28}$
Toplam	$\frac{3}{28} + \frac{3}{14} + \frac{1}{28} = \frac{5}{14}$	$\frac{9}{28} + \frac{3}{14} + 0 = \frac{15}{28}$	$\frac{3}{28} + 0 + 0 = \frac{3}{28}$	$\frac{15}{28} + \frac{3}{7} + \frac{1}{28} = 1$ $\frac{5}{14} + \frac{15}{28} + \frac{3}{28} = 1$



(ii) $A = \{(x, y) : x + y < 1\}$ olmak üzere $P((X, Y) \in A)$ olasılığı için

$$\begin{aligned}
 P((X, Y) \in A) &= f(0, 0) + f(0, 1) + f(1, 0) \\
 &= \frac{3}{28} + \frac{9}{28} + \frac{3}{14} = \frac{9}{14} = 0.64286
 \end{aligned}$$

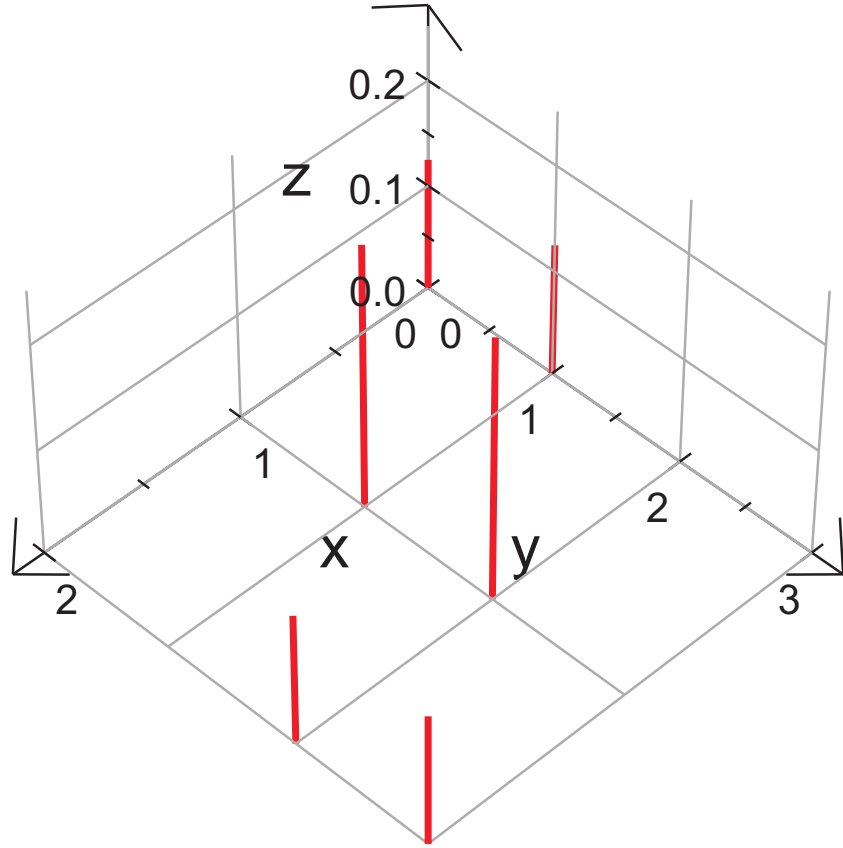
Örnek 7.3. Düzgün bir paranın 3 atılışında gelen turaların sayısı X , ilk iki atıştaki turaların sayısı Y olsun. S örneklem uzayı aşağıdaki gibidir.

Örneklem Uzayı	X	Y
TTT	3	2
TTY	2	2
TYT	2	1
YTT	2	1
TYY	1	1
YYT	1	0
YTY	1	1
YYY	0	0

Buna göre $f(x, y)$ birleşik olasılığı tanımlayınız.

Çözüm:

$X \backslash Y$	0	1	2	Toplam
0	$(YYY) \frac{1}{8}$	0	0	$\frac{1}{8}$
1	$(YYT) \frac{1}{8}$	$(TYY + YTY) \frac{1}{8} + \frac{1}{8} = \frac{1}{4}$	0	$\frac{3}{8}$
2	0	$(TYT + YTT) \frac{1}{8} + \frac{1}{8} = \frac{1}{4}$	$(TTY) \frac{1}{8}$	$\frac{3}{8}$
3	0	0	$(TTT) \frac{1}{8}$	$\frac{1}{8}$
Toplam	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = 1$ $\frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1$



□

Örnek 7.4. X ve Y rasgele değişkenlerinin birleşik olasılık dağılımı

$$f(x, y) = \begin{cases} c(2x + y) & \text{if } 0 \leq x \leq 2, 0 \leq y \leq 3, x, y \in \mathbb{Z} \\ 0 & \text{if } x \notin [0, 2] \cap \mathbb{Z}, y \notin [0, 3] \cap \mathbb{Z} \end{cases}$$

olarak verilmektedir. c değerini bulup $P(X = 2, Y = 1)$, $P(X \geq 1, Y \leq 2)$, olasılıklarını bulunuz.

Çözüm:

X/Y	0	1	2	3	$Toplam$
0	0	c	$2c$	$3c$	$6c$
1	$2c$	$3c$	$4c$	$5c$	$14c$
2	$4c$	$5c$	$6c$	$7c$	$22c$
$Toplam$	$6c$	$9c$	$12c$	$15c$	$42c = 1 \rightarrow Genel - Toplam$

$$42c = 1 \Rightarrow c = \frac{1}{42}$$

$$P(X = 2, Y = 1) = 5c = \frac{5}{42}$$

$$\begin{array}{|c|c|c|} \hline 2c & 3c & 4c \\ \hline 4c & 5c & 6c \\ \hline \end{array} \Rightarrow P(X \geq 1, Y \leq 2) = (f(1, 0) + f(1, 1) + f(1, 2))$$

$$+ (f(2, 0) + f(2, 1) + f(2, 2)) = 24c = \frac{24}{42} = \frac{4}{7} = 0.57143$$

□

Tanım 7.5. X ve Y sürekli rasgele değişkenlerinin olasılık yoğunluk fonksiyonu (probability mass function - pmf) aşağıdaki koşulları sağlayan $f(x, y) = P(X = x, Y = y)$ fonksiyonudur.

$$(i) f(x, y) \geq 0$$

$$(ii) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$$

$A \subset \mathbb{R}^2$ bir bölge olmak üzere $P((X, Y) \in A) = \int \int_A f(x, y)$ olarak da verilir.

Örnek 7.6. Özel bir şirket, müşterilerine yürüyüş gezisi ve araba gezisi imkanı sunmaktadır. Rasgele bir gün seçildiğinde X rasgele değişkeni yürüyüş gezisini, Y rasgele değişkeni de araba gezisinin göstermek üzere birleşik olasılık yoğunluk fonksiyonu

$$f(x, y) = \begin{cases} \frac{2}{5}(2x + 3y) & \text{if } 0 \leq x \leq 1, 0 \leq y \leq 1, x, y \in \mathbb{R} \\ 0 & \text{if } x \notin [0, 1], y \notin [0, 1] \end{cases}$$

ile tanımlanmaktadır. Bu fonksiyonun olasılık yoğunluk fonksiyonu olduğunu kanıtlayınız ve $P(0 < X \leq \frac{1}{2}, \frac{1}{4} < Y < \frac{1}{2})$ olasılığını hesaplayınız.

Çözüm: $f(x, y) \geq 0$ olduğu açıktır.

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy &= \int_0^1 \int_0^1 \frac{2}{5}(2x + 3y) dx dy \\ &= \int_0^1 \frac{2}{5} x^2 + 3xy \Big|_{x=0}^{x=1} dy \\ &= \int_0^1 \frac{2}{5} (1 + 3y) dy \\ &= \frac{2}{5} y + \frac{3}{2} y^2 \Big|_{y=0}^{y=1} \\ &= \frac{2}{5} \cdot 1 + \frac{3}{2} \cdot 1 = 1 \end{aligned}$$

O halde $f(x, y) \geq 0$ olasılık yoğunluk fonksiyonudur.

$$\begin{aligned}
 P\left(0 < X \leq \frac{1}{2}, \frac{1}{4} < Y < \frac{1}{2}\right) &= \int_{\frac{1}{4}}^{\frac{1}{2}} \int_0^{\frac{1}{2}} f(x, y) dx dy \\
 &= \int_{\frac{1}{4}}^{\frac{1}{2}} \int_0^{\frac{1}{2}} \frac{2}{5} (2x + 3y) dx dy \\
 &= \int_{\frac{1}{4}}^{\frac{1}{2}} \frac{2}{5} x^2 + 3xy \Big|_{x=0}^{x=1/2} dy \\
 &= \int_{\frac{1}{4}}^{\frac{1}{2}} \frac{2}{5} \left(\frac{1}{4} + \frac{3y}{2}\right) dy \\
 &= \frac{2}{5} \frac{y}{4} + \frac{3}{4} y^2 \Big|_{y=1/4}^{y=1/2} \\
 &= \frac{1}{10} * \left(\frac{1}{2} + 3 * \frac{1}{2^2} - \frac{1}{4} - 3 * \frac{1}{4^2}\right) \\
 &= \frac{13}{160} = 0.08125
 \end{aligned}$$

□

Örnek 7.7. X ve Y sürekli değişkenlerinin birleşik olasılık yoğunluğu

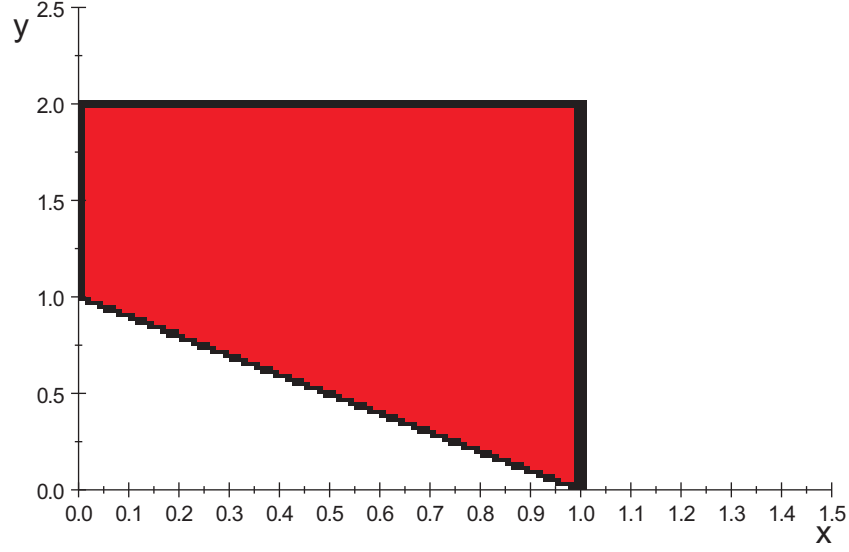
$$f(x, y) = \begin{cases} x^2 + \frac{xy}{3} & \text{if } 0 \leq x \leq 1, 0 \leq y \leq 2, x, y \in \mathbb{R} \\ 0 & \text{if } x \notin [0, 1], y \notin [0, 2] \end{cases}$$

olmak üzere bu fonksiyonun olasılık yoğunluk fonksiyonu olduğunu kanıtlayınız ve $P(X + Y \geq 1)$ olasılığını bulunuz.

Çözüm: $f(x, y) \geq 0$ olduğu açıktır.

$$\begin{aligned}
 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy &= \int_0^2 \int_0^1 \left(x^2 + \frac{xy}{3}\right) dx dy \\
 &= \int_0^2 \frac{x^3}{3} + \frac{x^2 y}{6} \Big|_{x=0}^{x=1} dy \\
 &= \int_0^2 \left(\frac{1}{3} + \frac{y}{6}\right) dy \\
 &= \frac{y}{3} + \frac{y^2}{12} \Big|_{y=0}^{y=2} \\
 &= \frac{2}{3} + \frac{4}{12} = 1
 \end{aligned}$$

O halde $f(x, y) \geq 0$ olasılık yoğunluk fonksiyonudur. Öncelikle $A = \{X + Y \geq 1 : 0 \leq X \leq 1, 0 \leq Y \leq 2\}$ bölgesini çizelim:



I.Yol:

$$\begin{aligned}
 P(X+Y \geq 1) &= \int \int_A f(x, y) dx dy \\
 &= \int_0^1 \int_{1-x}^2 \left(x^2 + \frac{xy}{3}\right) dy dx \\
 &= \int_0^1 x^2 y + \frac{xy^2}{6} \Big|_{y=1-x}^{y=2} dx \\
 &= \int_0^1 \left(2x^2 + \frac{2x}{3} - x^2(1-x) - \frac{x(1-x)^2}{6}\right) dx \\
 &= \frac{1}{6} \int_0^1 (5x^3 + 8x^2 + 3x) dx = \frac{65}{72}
 \end{aligned}$$

2.Yol:

$$\begin{aligned}
 P(X+Y \geq 1) &= 1 - P(X+Y < 1) \\
 &= 1 - \int_0^1 \int_0^{1-x} \left(x^2 + \frac{xy}{3}\right) dy dx \\
 &= 1 - \int_0^1 x^2 y + \frac{xy^2}{6} \Big|_{y=0}^{y=1-x} dx \\
 &= 1 - \int_0^1 \left(x^2(1-x) + \frac{x(1-x)^2}{6}\right) dx \\
 &= 1 - \frac{1}{6} \int_0^1 (-5x^3 + 4x^2 + x) dx \\
 &= 1 - \frac{1}{6} * \frac{7}{12} = \frac{65}{72}
 \end{aligned}$$

Örnek 7.8. X ve Y sürekli değişkenlerinin birleşik olasılık yoğunluğu

$$f(x, y) = \begin{cases} 2e^{-x-2y} & \text{if } x, y \in \mathbb{R}^+ \\ 0 & \text{if } x, y \notin \mathbb{R}^+ \end{cases}$$

olarak verilmektedir. Buna göre $P(X > 1, Y < 1)$, $P(X < Y)$ olasılıklarını bulunuz.

Çözüm:

$$\begin{aligned} P(X > 1, Y < 1) &= \int_0^1 \int_1^\infty 2e^{-x-2y} dx dy \\ &= \int_0^1 -2 e^{-x-2y} \Big|_{x=1}^{x=\infty} dy \\ &= \int_0^1 2e^{-2y-1} dy \\ &= -e^{-2y-1} \Big|_{y=0}^{y=1} \\ &= e^{-1} - e^{-3} \\ P(X < Y) &= \int_0^\infty \int_0^y 2e^{-x-2y} dx dy \\ &= \int_0^\infty -2 e^{-x-2y} \Big|_{x=0}^{x=y} dy \\ &= \int_0^\infty -2 (e^{-3y} - e^{-2y}) dy \\ &= \frac{2}{3} e^{-3y} \Big|_{y=0}^{y=\infty} - e^{-2y} \Big|_{y=0}^{y=\infty} \\ &= -\frac{2}{3} + 1 = \frac{1}{3} = 0.333 \text{ 33} \end{aligned}$$

□

Tanım 7.9. Rasgele değişkenlerinden biri sabit tutulup sadece diğeri üzerinden hesaplanan olasılığa marjinal olasılık denir. X ve Y kesikli rasgele değişkenlerinin marjinal olasılık fonksiyonu

$$g(x) = \sum_y f(x, y) \text{ ve } h(y) = \sum_x f(x, y)$$

X ve Y sürekli rasgele değişkenlerinin marjinal olasılık fonksiyonu

$$g(x) = \int_{-\infty}^{\infty} f(x, y) dy \text{ ve } h(y) = \int_{-\infty}^{\infty} f(x, y) dx$$

olarak tanımlanır.

X/Y	y_1	y_2	\dots	y_n	$Toplam$
x_1	$f(x_1, y_1)$	$f(x_1, y_2)$	\dots	$f(x_1, y_n)$	$\sum_{j=1}^n f(x_1, y_j) = g(x_1)$
x_2	$f(x_2, y_1)$	$f(x_2, y_2)$	\dots	$f(x_2, y_n)$	$\sum_{j=1}^n f(x_2, y_j) = g(x_2)$
\vdots	\dots	\dots	\dots	\dots	\dots
x_m	$f(x_m, y_1)$	$f(x_m, y_2)$	\dots	$f(x_m, y_n)$	$\sum_{j=1}^n f(x_m, y_j) = g(x_m)$
$Toplam$	$\underbrace{\sum_{i=1}^m f(x_i, y_1)}_{h(y_1)}$	$\underbrace{\sum_{i=1}^m f(x_i, y_2)}_{h(y_2)}$	\dots	$\underbrace{\sum_{i=1}^m f(x_i, y_n)}_{h(y_n)}$	$1 \rightarrow Genel - Toplam$

Örnek 7.10. 2 renkli kalem, 3 mavi, 2 kırmızı ve 3 yeşil kalemin olduğu bir kutudan seçilecektir. Eğer X rasgele değişkeni mavi kalem seçimini, Y ise kırmızı kalem seçimini gösterirse marjinal birleşik olasılığı tanımlayınız.

Çözüm:

$f(x, y)$	$x - \text{mavi}$			$h(y)$
$y - \text{kırmızı}$	0	1	2	
0	$\frac{\frac{3!}{1! \cdot 2!}}{28} = \frac{3}{28}$	$\frac{\frac{3!}{2! \cdot 1!} \cdot \frac{3!}{2! \cdot 1!}}{28} = \frac{9}{28}$	$\frac{\frac{3!}{2! \cdot 1!}}{28} = \frac{3}{28}$	$\frac{3}{28} + \frac{9}{28} + \frac{3}{28} = \frac{15}{28} = h(0)$
1	$\frac{\frac{2!}{1! \cdot 1!} \cdot \frac{3!}{1! \cdot 2!}}{28} = \frac{3}{14}$	$\frac{\frac{2!}{1! \cdot 1!} \cdot \frac{3!}{1! \cdot 2!}}{28} = \frac{3}{14}$	0	$\frac{3}{14} + \frac{3}{14} + 0 = \frac{3}{7} = h(1)$
2	$\frac{\frac{2!}{0! \cdot 2!}}{28} = \frac{1}{28}$	0	0	$\frac{1}{28} + 0 + 0 = \frac{1}{28} = h(2)$
$g(x)$	$\frac{3}{28} + \frac{3}{14} + \frac{1}{28} = \frac{5}{14}$	$\frac{9}{28} + \frac{3}{14} + 0 = \frac{15}{28}$	$\frac{3}{28} + 0 + 0 = \frac{3}{28}$	$\frac{15}{28} + \frac{3}{7} + \frac{1}{28} = 1$ $\frac{5}{14} + \frac{15}{28} + \frac{3}{28} = 1$
	$g(0)$	$g(1)$	$g(2)$	

$$g(0) = f(0, 0) + f(0, 1) + f(0, 2)$$

$$g(1) = f(1, 0) + f(1, 1) + f(1, 2)$$

$$g(2) = f(2, 0) + f(2, 1) + f(2, 2)$$

$$h(0) = f(0, 0) + f(1, 0) + f(2, 0)$$

$$h(1) = f(0, 1) + f(1, 1) + f(2, 1)$$

$$h(2) = f(0, 2) + f(1, 2) + f(2, 2)$$

□

Örnek 7.11. Özel bir şirket, müşterilerine yürüyüş gezisi ve araba gezisi imkanı sunmaktadır. Rasgele bir gün seçildiğinde X rasgele değişkeni yürüyüş gezisini, Y rasgele değişkeni de araba gezisinin göstermek üzere birleşik olasılık yoğunluk fonksiyonu

$$f(x, y) = \begin{cases} \frac{2}{5}(2x + 3y) & \text{if } 0 \leq x \leq 1, 0 \leq y \leq 1, x, y \in \mathbb{R} \\ 0 & \text{if } x \notin [0, 1], y \notin [0, 1] \end{cases}$$

ile tanımlanmaktadır. X ve Y sürekli rasgele değişkenlerinin marjinal olasılık fonksiyonunu bulunuz.

Çözüm:

$$\begin{aligned}
 g(x) &= \int_{-\infty}^{\infty} f(x, y) dy = \int_0^1 \frac{2}{5} (2x + 3y) dy \\
 &= \frac{2}{5} 2xy + \frac{3}{2} y^2 \Big|_{y=0}^{y=1} \\
 &= \frac{2}{5} \left(2x + \frac{3}{2} \right), 0 \leq x \leq 1 \Rightarrow \\
 g(x) &= \begin{cases} \frac{2}{5} \left(2x + \frac{3}{2} \right) & \text{if } 0 \leq x \leq 1, x \in \mathbb{R} \\ 0 & \text{if } x \notin [0, 1], \end{cases} \\
 h(y) &= \int_{-\infty}^{\infty} f(x, y) dx = \int_0^1 \frac{2}{5} (2x + 3y) dx \\
 &= \frac{2}{5} x^2 + 3xy \Big|_{x=0}^{x=1} = \frac{2}{5} (1 + 3y), 0 \leq y \leq 1 \\
 h(y) &= \begin{cases} \frac{2}{5} (1 + 3y) & \text{if } 0 \leq y \leq 1, y \in \mathbb{R} \\ 0 & \text{if } y \notin [0, 1], \end{cases}
 \end{aligned}$$

□

Örnek 7.12. Bir mühendis şehir merkezindeki bir dört yol ağzında sabahın erken saatlerinden başlayarak trafik modeli üzerine çalışmaktadır. Mühendis gözleme sabah 5:30 da başlamaktadır. X , kuzey-güney yönünden ilk aracın varış zamanını gösterebilir. Y , doğu-batı yönünden ilk aracın varış zamanını gösterebilir. Zaman saat 5:30 dan sonra saatin kesirleriyle ölçülmektedir. (X, Y) sürekli rasgele değişkeni için yoğunluk fonksiyonu

$$f(x, y) = \frac{1}{x}, 0 < y < x < 1$$

ile verilsin. X ve Y sürekli rasgele değişkenlerinin marjinal olasılık fonksiyonunu bulunuz.

Çözüm:

$$\begin{aligned}
 g(x) &= \int_{-\infty}^{\infty} f(x, y) dy = \int_0^x \frac{1}{x} dy = \frac{1}{x} y \Big|_{y=0}^{y=x} = 1 \\
 h(y) &= \int_{-\infty}^{\infty} f(x, y) dx = \int_y^1 \frac{1}{x} dx = \ln x \Big|_{x=y}^{x=1} = -\ln y
 \end{aligned}$$

□

Örnek 7.13. X ve Y sürekli rasgele değişkenlerinin birleşik olasılık yoğunluk fonksiyonu

$$f(x, y) = \begin{cases} \frac{1}{8} (6 - x - y) & \text{if } 0 \leq x \leq 2, 2 < y \leq 4, x, y \in \mathbb{R} \\ 0 & \text{if } x \notin [0, 2], y \notin (2, 4] \end{cases}$$

olmak üzere

(i) $P(X \leq 1, Y \leq 3)$ olasılığını hesaplayınız

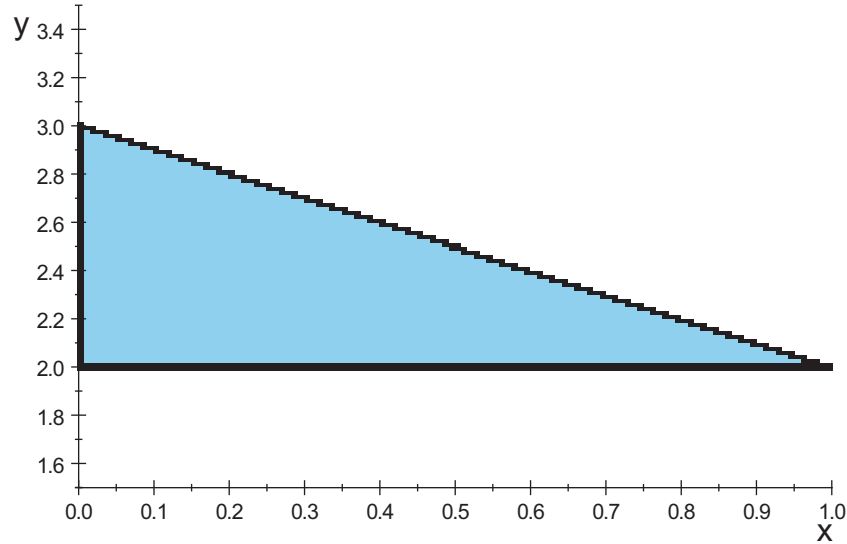
(ii) $P(X + Y \leq 3)$ olasılığını hesaplayınız

(iii) X ve Y sürekli rasgele değişkenlerinin marjinal olasılık fonksiyonunu bulunuz.

Çözüm: (i)

$$\begin{aligned}
 P(X \leq 1, Y \leq 3) &= \int_{-\infty}^1 \int_{-\infty}^3 f(x, y) dy dx = \int_0^1 \int_2^3 \frac{1}{8} (6 - x - y) dy dx \\
 &= \frac{1}{8} \int_0^1 6y - xy - \frac{y^2}{2} \Big|_{y=2}^{y=3} dx = \int_0^1 \left(\frac{7}{16} - \frac{1}{8}x \right) dx \\
 &= \frac{7}{16}x - \frac{1}{16}x^2 \Big|_{x=0}^{x=1} = \frac{3}{8}
 \end{aligned}$$

(ii) Öncelikle $A = \{X + Y \leq 3 : 0 \leq X \leq 2, 2 < Y \leq 4\}$ bölgesini belirleyelim:



$$\begin{aligned}
 P(X + Y \leq 3) &= \int \int_A f(x, y) dx dy = \int_0^1 \int_2^{3-x} \frac{1}{8} (6 - x - y) dy dx \\
 &= \frac{1}{8} \int_0^1 6y - xy - \frac{y^2}{2} \Big|_{y=2}^{y=3-x} dx \\
 &= \frac{1}{8} \int_0^1 \left(6(3-x) - x(3-x) - \frac{(3-x)^2}{2} - 12 + 2x + 2 \right) dx \\
 &= \frac{1}{8} \int_0^1 \left(\frac{1}{2}x^2 - 4x + \frac{7}{2} \right) dx = \frac{1}{8} \left(\frac{1}{6}x^3 - 2x^2 + \frac{7}{2}x \right) \Big|_{x=0}^{x=1} = \frac{5}{24}
 \end{aligned}$$

(iii)

$$\begin{aligned}
 g(x) &= \int_{-\infty}^{\infty} f(x, y) dy = \int_2^4 \frac{1}{8} (6 - x - y) dy \\
 &= \frac{1}{8} 6y - xy - \frac{y^2}{2} \Big|_{y=2}^{y=4} = \frac{3}{4} - \frac{1}{4}x \\
 h(y) &= \int_{-\infty}^{\infty} f(x, y) dx = \int_0^2 \frac{1}{8} (6 - x - y) dx \\
 &= \frac{1}{8} 6x - \frac{x^2}{2} - xy \Big|_{x=0}^{x=2} = \frac{5}{4} - \frac{1}{4}y
 \end{aligned}$$

□

Tanım 7.14. X ve Y kesikli veya sürekli rasgele değişkenleri olmak üzere $X = x$ verildiğine Y rasgele değişkeninin koşullu dağılım fonksiyonu

$$f(y|x) = \frac{f(x, y)}{g(x)}, g(x) > 0$$

ve $Y = y$ verildiğinde X rasgele değişkeninin koşullu dağılım fonksiyonu

$$f(x|y) = \frac{f(x, y)}{h(y)}, h(y) > 0$$

ile tanımlanır. Buna göre $X = x$ verildiğinde Y kesikli rasgele değişkeninin $c \leq Y \leq d$ koşulunu sağlaması olasılığı

$$P(c \leq Y \leq d | X = x) = \sum_{c \leq y \leq d} f(y|x)$$

$X = x$ verildiğinde Y sürekli rasgele değişkeninin $c \leq Y \leq d$ koşulunu sağlaması olasılığı

$$P(c \leq Y \leq d | X = x) = \int_c^d f(y|x) dy$$

şeklinde tanımlanır. Benzer şekilde $Y = y$ verildiğinde X kesikli rasgele değişkeninin $a \leq X \leq b$ koşulunu sağlaması olasılığı

$$P(a \leq X \leq b | Y = y) = \sum_{a \leq x \leq b} f(x|y)$$

$Y = y$ verildiğinde X sürekli rasgele değişkeninin $a \leq X \leq b$ koşulunu sağlaması olasılığı

$$P(a \leq X \leq b | Y = y) = \int_a^b f(x|y) dx$$

olarak tanımlanır.

Örnek 7.15. 2 renkli kalem, 3 mavi, 2 kırmızı ve 3 yeşil kalemin olduğu bir kutudan seçilecektir. Eğer X rasgele değişkeni mavi kalem seçimini, Y ise kırmızı kalem seçimini gösterirse 1 kırmızı kalem seçildiği bilindiğine göre hiç mavi kalem seçilmeme olasılığı ve 1 mavi kalem seçilme olasılığı nedir?

Çözüm:

$f(x, y)$	x – mavi			$h(y)$
y – kırmızı	0	1	2	
0	$\frac{3!}{1! \cdot 2!} = \frac{3}{28}$	$\frac{3!}{2! \cdot 1!} \cdot \frac{3!}{2! \cdot 1!} = \frac{9}{28}$	$\frac{3!}{2! \cdot 1!} = \frac{3}{28}$	$\frac{3}{28} + \frac{9}{28} + \frac{3}{28} = \frac{15}{28} = h(0)$
1	$\frac{2!}{1! \cdot 1!} \cdot \frac{3!}{2! \cdot 1!} = \frac{3}{14}$	$\frac{2!}{1! \cdot 1!} \cdot \frac{3!}{2! \cdot 1!} = \frac{3}{14}$	0	$\frac{3}{14} + \frac{3}{14} + 0 = \frac{3}{7} = h(1)$
2	$\frac{2!}{0! \cdot 2!} = \frac{1}{28}$	0	0	$\frac{1}{28} + 0 + 0 = \frac{1}{28} = h(2)$
$g(x)$	$\frac{3}{28} + \frac{3}{14} + \frac{1}{28} = \frac{5}{14}$	$\frac{9}{28} + \frac{3}{14} + 0 = \frac{15}{28}$	$\frac{3}{28} + 0 + 0 = \frac{3}{28}$	$\frac{15}{28} + \frac{3}{7} + \frac{1}{28} = 1$ $\frac{5}{14} + \frac{15}{28} + \frac{3}{28} = 1$
	$g(0)$	$g(1)$	$g(2)$	

$$P(X = 0|Y = 1) = ?$$

$$P(X = 0|Y = 1) = \frac{f(0,1)}{h(1)}$$

$$P(X = 0|Y = 1) = \frac{3/14}{3/7} = \frac{1}{2} = 0.50$$

$$P(X = 1|Y = 1) = ?$$

$$P(X = 1|Y = 1) = \frac{f(1,1)}{h(1)}$$

$$P(X = 1|Y = 1) = \frac{3/14}{3/7} = \frac{1}{2} = 0.50$$

x	0	1	2
$f(x 1)$	$\frac{1}{2}$	$\frac{1}{2}$	0

□

Örnek 7.16. Bir atom çekirdeğinin parçalanması deneyinde X rasgele değişkeni değişen sıcaklığı, Y rasgele değişkeni ise atom parçacığındaki değişen spektrum oranını göstermek üzere, (X, Y) rasgele değişkenleri için birleşik olasılık dağılımı

$$f(x, y) = \begin{cases} 10xy^2 & \text{if } 0 < x < y < 1, x, y \in \mathbb{R} \\ 0 & \text{otherwise} \end{cases}$$

olarak veriliyor.

(i) $g(x)$, $h(y)$ marjinal olasılık fonksiyonunu bulup $f(y|x)$ koşullu dağılım fonksiyonunu hesaplayınız.

(ii) Sıcaklığın 0.25 birim olduğu bilindiğinde atom parçacığının değişen spektrum oranı yarıdan fazla olma olasılığı nedir?

Çözüm: (i)

$$\begin{aligned}
 g(x) &= \int_{-\infty}^{\infty} f(x, y) dy \\
 &= \int_x^1 10xy^2 dy \\
 &= \frac{10}{3} x y^3 \Big|_{y=x}^{y=1} \\
 &= \frac{10}{3} x (1 - x^3), 0 < x < 1 \\
 g(x) &= \begin{cases} \frac{10}{3} x (1 - x^3) & \text{if } 0 < x < 1, x \in \mathbb{R} \\ 0 & \text{if } x \notin (0, 1), \end{cases} \\
 h(y) &= \int_{-\infty}^{\infty} f(x, y) dx = \int_0^y 10xy^2 dx \\
 &= 5y^2 x^2 \Big|_{x=0}^{x=y} = 5y^2 (y^2), 0 < y < 1 \\
 h(y) &= \begin{cases} 5y^4 & \text{if } 0 < y < 1, y \in \mathbb{R} \\ 0 & \text{if } y \notin (0, 1), \end{cases} \\
 f(y|x) &= \frac{f(x, y)}{g(x)} = \frac{10xy^2}{\frac{10}{3}x(1-x^3)} \\
 &= \frac{3y^2}{(1-x^3)}, 0 < x < y < 1, x, y \in \mathbb{R}
 \end{aligned}$$

(ii)

$$\begin{aligned}
 P(Y > 0.5 \mid X = 0.25) &= \int_{0.5}^1 f(y|x = 0.25) dy \\
 &= \int_{0.5}^1 \frac{3y^2}{(1 - 0.25^3)} dy \\
 &= 0.98438 * y^3 \Big|_{y=0.5}^{y=1} \\
 &= 0.98438 * (1 - 0.5^3) \\
 &= 0.86133
 \end{aligned}$$

□

Örnek 7.17. Birleşik olasılık dağılım fonksiyonu

$$f(x, y) = \begin{cases} \frac{x(1+3y^2)}{4} & \text{if } 0 < x < 2, 0 < y < 1, x, y \in \mathbb{R} \\ 0 & \text{otherwise} \end{cases}$$

olarak verildiğine göre

(i) $g(x), h(y)$ marjinal olasılık fonksiyonunu bulup $f(x|y)$ koşullu dağılım fonksiyonunu hesaplayınız.

(ii) $P\left(\frac{1}{4} < X < \frac{1}{2} \mid Y = \frac{1}{3}\right)$ olasılığı nedir?

Çözüm: (i)

$$\begin{aligned}
 g(x) &= \int_{-\infty}^{\infty} f(x, y) dy \\
 &= \int_0^1 \frac{x(1+3y^2)}{4} dy \\
 &= \frac{1}{4} x y + y^3 \Big|_{y=0}^{y=1} \\
 &= \frac{1}{4} x (2), 0 < x < 2 \\
 g(x) &= \begin{cases} \frac{1}{2}x & \text{if } 0 < x < 2, x \in \mathbb{R} \\ 0 & \text{if } x \notin (0, 2), \end{cases} \\
 h(y) &= \int_{-\infty}^{\infty} f(x, y) dx = \int_0^2 \frac{x(1+3y^2)}{4} dx \\
 &= \frac{(1+3y^2)}{8} x^2 \Big|_{x=0}^{x=2} = \frac{(1+3y^2)}{8} * 2^2, 0 < y < 1 \\
 h(y) &= \begin{cases} \frac{(1+3y^2)}{2} & \text{if } 0 < y < 1, y \in \mathbb{R} \\ 0 & \text{if } y \notin (0, 1), \end{cases} \\
 f(x|y) &= \frac{f(x, y)}{h(y)} = \frac{\frac{x(1+3y^2)}{4}}{\frac{(1+3y^2)}{2}} \\
 &= \frac{x}{2}, 0 < x < 2, 0 < y < 1, x, y \in \mathbb{R}
 \end{aligned}$$

(ii)

$$\begin{aligned}
 P\left(\frac{1}{4} < X < \frac{1}{2} \mid Y = \frac{1}{3}\right) &= \int_{1/4}^{1/2} f\left(x \mid y = \frac{1}{3}\right) dx \\
 &= \int_{1/4}^{1/2} \frac{x}{2} dx \\
 &= \frac{1}{4} * x^2 \Big|_{x=1/4}^{x=1/2} \\
 &= \frac{1}{4} * \left(\frac{1}{2^2} - \frac{1}{4^2}\right) \\
 &= \frac{3}{64} = 4.6875 \times 10^{-2}
 \end{aligned}$$

□

Örnek 7.18. X ve Y kesikli rasgele değişkenlerinin birleşik olasılık yoğunluk fonksiyonu

X/Y	1	3	5
1	0.05	0.05	0
2	0.05	0.1	0.2
3	0.1	0.35	0.1

olarak verilsin.

(i) $g(x), h(y)$ marjinal olasılık fonksiyonunu bulunuz

(ii) $P(Y = 3|X = 2)$ olasılığı nedir?

Çözüm: (i)

$$g(1) = f(1,1) + f(1,3) + f(1,5) = 0.1$$

$$g(2) = f(2,1) + f(2,3) + f(2,5) = 0.35$$

$$g(3) = f(3,1) + f(3,3) + f(3,5) = 0.55$$

$$h(1) = f(1,1) + f(2,1) + f(3,1) = 0.2$$

$$h(3) = f(1,3) + f(2,3) + f(3,3) = 0.5$$

$$h(5) = f(1,5) + f(2,5) + f(3,5) = 0.3$$

(ii)

$$P(Y = 3|X = 2) = \frac{f(3,2)}{g(2)} = \frac{0.1}{0.35} = 0.28571$$

□

Tanım 7.19. X ve Y kesikli veya sürekli rasgele değişkenler, $f(x, y)$ birleşik olasılık yoğunluk fonksiyonu, $g(x), h(y)$ marjinal olasılık fonksiyonları olmak üzere

$$f(x, y) = g(x)h(y)$$

koşulu sağlanıyorsa, X ve Y rasgele değişkenlerine istatistiksel bağımsızdır denir.

Örnek 7.20. 2 renkli kalem, 3 mavi, 2 kırmızı ve 3 yeşil kalemin olduğu bir kutudan seçilecektir. Eğer X rasgele değişkeni mavi kalem seçimini, Y ise kırmızı kalem seçimini gösterirse, X ve Y rasgele değişkenlerinin istatistiksel bağımsız olmadığını gösteriniz.

Çözüm:

$$f(0,1) = \frac{3}{14}$$

$$g(0) = f(0,0) + f(0,1) + f(0,2) = \frac{5}{14}$$

$$h(1) = f(0,1) + f(1,1) + f(2,1) = \frac{3}{7}$$

$$g(0) * h(1) = \frac{5}{14} * \frac{3}{7} = \frac{15}{98} \neq \frac{3}{14} = f(0,1)$$

□

Örnek 7.21. X ve Y kesikli rasgele değişkenleri bağımsızdır ve marjinal olasılık fonksiyonları aşağıdaki gibi verilmiştir:

x	1	2
$g(x)$	0.6	0.4

y	5	10	15
$h(y)$	0.2	0.5	0.3

Buna göre X ve Y kesikli rasgele değişkenlerinin $f(x, y)$ birleşik olasılık yoğunluk fonksiyonunu bulunuz.

Çözüm:

$f(x, y)$	\mathbf{x}		$\mathbf{h(y)}$
\mathbf{y}	1	2	
5	$h(5) * g(1) =$ $0.2 * 0.6 = 0.12$	$h(5) * g(2) =$ $0.2 * 0.4 = 0.08$	$h(5) = 0.2$
10	$h(10) * g(1) =$ $0.5 * 0.6 = 0.3$	$h(10) * g(2) =$ $0.5 * 0.4 = 0.2$	$h(10) = 0.5$
15	$h(15) * g(1) =$ $0.3 * 0.6 = 0.18$	$h(15) * g(2) =$ $0.3 * 0.4 = 0.12$	$h(15) = 0.3$
$g(x)$	$g(1) = 0.6$	$g(2) = 0.4$	

□

Tanım 7.22. X_1, X_2, \dots, X_n kesikli rasgele değişkenlerinin olasılık yoğunluk fonksiyonu (probability mass function - pmf) aşağıdaki koşulları sağlayan $f(x_1, x_2, \dots, x_n) = P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$ fonksiyonudur.

(i) $f(x_1, x_2, \dots, x_n) \geq 0$

(ii) $\sum_{x_1} \sum_{x_2} \dots \sum_{x_n} f(x_1, x_2, \dots, x_n) = 1$

$A \subset \mathbb{R}^n$ bir bölge olmak üzere $P((X_1, X_2, \dots, X_n) \in A) = \sum_{(x,y) \in A} f(x_1, x_2, \dots, x_n)$ olarak da verilir.

Tanım 7.23. X_1, X_2, \dots, X_n sürekli rasgele değişkenlerinin olasılık yoğunluk fonksiyonu (probability mass function - pmf) aşağıdaki koşulları sağlayan $f(x_1, x_2, \dots, x_n) = P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$ fonksiyonudur.

(i) $f(x_1, x_2, \dots, x_n) \geq 0$

(ii) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n = 1$

$A \subset \mathbb{R}^n$ bir bölge olmak üzere $P((X_1, X_2, \dots, X_n) \in A) = \int \int \dots \int_A f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n$ olarak da verilir.

Tanım 7.24. *rasgele değişkenlerinden biri sabit tutulup sadece diğeri üzerinden hesaplanan olasılığa marjinal olasılık denir. X_1, X_2, \dots, X_n kesikli rasgele değişkenlerinin marjinal olasılık fonksiyonları*

$$\begin{aligned} g(x_1) &= \sum_{x_2} \dots \sum_{x_n} f(x_1, x_2, \dots, x_n) \\ g(x_1, x_2) &= \sum_{x_3} \dots \sum_{x_n} f(x_1, x_2, \dots, x_n) \\ g(x_1, x_2, x_3) &= \sum_{x_4} \dots \sum_{x_n} f(x_1, x_2, \dots, x_n) \\ &\dots \end{aligned}$$

X_1, X_2, \dots, X_n sürekli rasgele değişkenlerinin marjinal olasılık fonksiyonları

$$\begin{aligned} g(x_1) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_n) dx_2 \dots dx_n \\ g(x_1, x_2) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_n) dx_3 \dots dx_n \\ g(x_1, x_2, x_3) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_n) dx_4 \dots dx_n \end{aligned}$$

olarak tanımlanır.

Tanım 7.25. X_1, X_2, \dots, X_n kesikli veya sürekli rasgele değişkenleri olmak üzere $X_4 = x_4, X_5 = x_5, \dots, X_n = x_n$ verildiğine X_1, X_2, X_3 rasgele değişkeninin koşullu dağılım fonksiyonu

$$f(x_1, x_2, x_3 | x_4, x_5, \dots, x_n) = \frac{f(x_1, x_2, \dots, x_n)}{g(x_4, x_5, \dots, x_n)}, g(x_4, x_5, \dots, x_n) > 0$$

olarak tanımlanır.

Tanım 7.26. X_1, X_2, \dots, X_n kesikli veya sürekli rasgele değişkenler, $f(x_1, x_2, \dots, x_n)$ birleşik olasılık yoğunluk fonksiyonu, $g_1(x_1), g_2(x_2), \dots, g_n(x_n)$ marjinal olasılık fonksiyonları olmak üzere

$$f(x_1, x_2, \dots, x_n) = g_1(x_1) g_2(x_2) \dots g_n(x_n)$$

koşulu sağlanıyorsa, X_1, X_2, \dots, X_n rasgele değişkenlerine istatistiksel bağımsızdır denir.

Örnek 7.27. Bir ürünün raf ömrü aşağıdaki yoğunluk fonksiyonu ile verilir:

$$f(x) = \begin{cases} e^{-x} & \text{if } x > 0, x \in \mathbb{R} \\ 0 & \text{otherwise} \end{cases}$$

Buna göre marketten seçilen 3 üründen 1.sinin 2 aydan az, 2.sinin 1 ile 3 ay, 3. sününde 2 aydan fazla olma olasılığı nedir? $P(X_1 < 2, 1 < X_2 < 3, X_3 > 2) = ?$

Çözüm:

$$\begin{aligned}
P(X_1 < 2, 1 < X_2 < 3, X_3 > 2) &= \int_0^2 \int_1^3 \int_2^\infty e^{-x_1} e^{-x_2} e^{-x_3} dx_1 dx_2 dx_3 \\
&= -e^{-x_1} \Big|_{x_1=0}^{x_1=2} e^{-x_2} \Big|_{x_2=1}^{x_2=3} e^{-x_3} \Big|_{x_3=2}^{x_3=\infty} \\
&= -(1 - e^{-2}) (e^{-1} - e^{-3}) (-e^{-2}) \\
&= (1 - e^{-2}) * (e^{-1} - e^{-3}) * e^{-2} \\
&= 0.0372
\end{aligned}$$

□

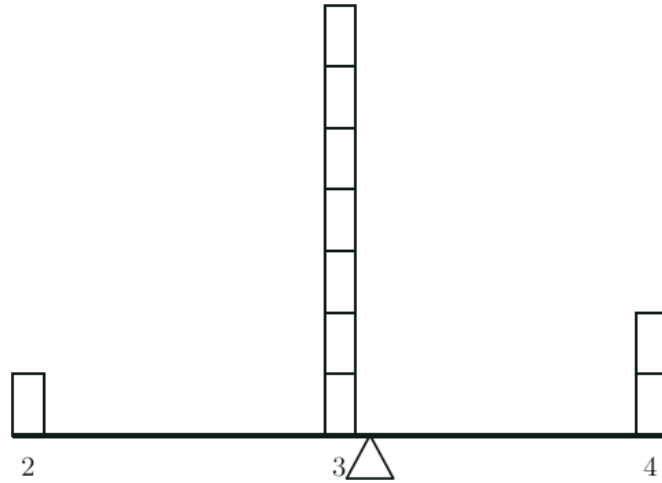
8 Rasgele deęişkenlerinin beklenen deęeri (Mathematical Expextation of random variable)

Benzin řirketi, benzinin varlıęı için belli bölgelerde sondaj çalışmaları yapmaktadır. Çalışmada kullanılan makineler 2,3 ve 4 er saate sondaj çalışmalarını bitirirken bunların işlemin sonucunda petrol bulma olasılıkları 0.1, 0.7 ve 0.2 olarak belirleniyor. Şirket tek bir makineyle de çalışmadıęına göre 10 makine ile çalışılırsa ortalama olarak kaç saatte petrol bulma işlemi gerçekleşir sorusunun cevabını nasıl veririz?

Bunun için 1 makinenin bu üç makineden biri olması durumunda ortalama kaç saatte bitireceğimiz hesaplarız:

$$2 * 0.1 + 3 * 0.7 + 4 * 0.2 = 3.1$$

Bu hesaba göre 1 makine için (bu üç makineden biri olması durumunda) ortalama sondaj çalışmasını bitirme süresi 3.1 saattir.



10 makine için ise $10 * 3.1 = 31$ saatte işmeleri bitimiş oluruz. Bu deęere beklenen deęer (ortalama deęer -mean value)

Zar atıyorsunuz. İlk atışta 6 gelirse kazanıyorsunuz. İlk atışta 6 gelmezse bir daha atıp, gelen zarı daha önce gelen zarla topluyorsunuz. Eęer iki zarın toplamı 6 ya da daha fazlaysa, oyunu kazanıyorsunuz. Yoksa, bir defa daha zar atıyorsunuz. Attığınız zarların toplamı 6 ya da daha fazla olduęunda oyunu kazanıyorsunuz. Eęer zarların toplamı 6'dan küçükse bir defa daha atıyorsunuz. Zarların toplamı 6 ya da daha fazla olana dek bunu böylece sürdürüyorsunuz. Zarların toplamı 6 ya da daha fazla olduęunda - sizi candan kutlarım! - oyunu kazandınız...

Bu oyunu ortalama kaç hamlede kazanırsınız?

Oyun en az 1 hamle sürüyor (ilk zar 6 gelirse); en çok da 6 hamle (6 defa üstüste 1 gelirse). Yani en fazla 6 hamlede oyunu kazanacaęınızdan eminsiniz. Ama ben ortalama hamle sayısını soruyorum.

Diyelim bu oyunu 1 milyar kez oynadınız ve her seferinde kaçınıcı hamlede kazandınızı bir kâğıda yazdınız. Sonra skorlarınızı toplayıp, toplam oyun sayısı olan 1 milyara böldünüz. Kaç civarında bir sayı bulmayı ümit edersiniz?

Önceden bir tahminde bulunmak gerekirse 2-3 arasında bir tahmin oldukça makuldür.

Bulmamız gereken bu sayıya beklenti denir; “kazanan hamle sayısının beklentisi”.

1/6 olasılıkla ilk zar 6 gelir ve oyunu 1 hamlede kazanırız. Bu olgu, beklentiye

$$1/6 \times 1$$

kadar katkı sağlar.

1/6 olasılıkla ilk zar 5 gelir ve ikinci atışta oyunu kesinkes kazanırız; bu durumda oyun 2 hamlede biter. Bu olgu, beklentiye

$$1/6 \times 2$$

kadar katkı sağlar.

Peki ilk zar 4 gelirse ne olur? O zaman ikinci zar 2, 3, 4, 5, 6 gelirse kazanırız. Bu olayın olasılığı da $1/6 \times 5/6$ 'dır. Bu durumda oyun gene 2 hamlede kazanılır. Bu olgu, beklentiye

$$1/6 \times 5/6 \times 2$$

kadar katkı sağlar.

Buna benzer hesapları her türlü kombinasyonla yapmak gerekir. Örneğin 3-2-4 zarlarının olasılığı 1/63'dir ve bu durumda oyun 3 hamlede biter. Bu olayın beklentiye katkısı,

$$1/6^3 \times 3$$

tür. Bütün bu sayıları toplarsak oyunun ortalama kaç hamlede biteceğini buluruz.

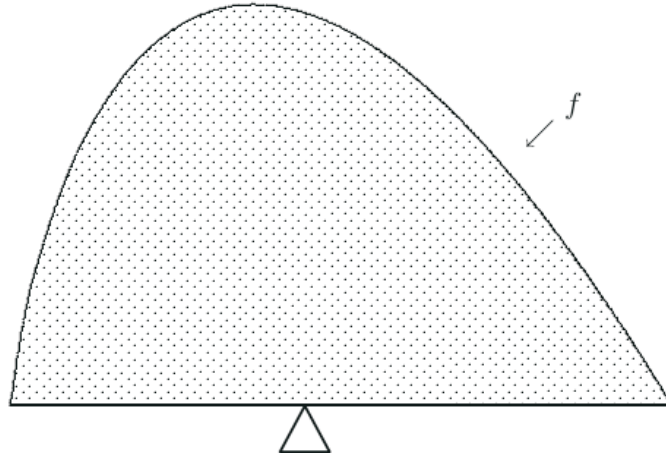
Tanım 8.1. *X kesikli rasgele değişkeninin olasılık dağılım fonksiyonu $f(x)$ olmak üzere X kesikli rasgele değişkeninin beklenti değeri (ortalama değeri- mean of a random variable)*

$$\mu_X = E[X] = \sum_x x f(x)$$

olarak tanımlanırken, X sürekli rasgele değişkeninin olasılık dağılım fonksiyonu $f(x)$ olmak üzere X sürekli rasgele değişkeninin beklenti değeri

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

olarak tanımlanır.



Örnek 8.2. Hilesiz bir zar atışında X rastgele gelen değeri göstermek üzere $E[X]$ beklenen değeri nedir?

Çözüm:

$$P(X = 1) = f(1) = \frac{1}{6},$$

$$P(X = 2) = f(2) = \frac{1}{6},$$

$$P(X = 3) = f(3) = \frac{1}{6},$$

$$P(X = 4) = f(4) = \frac{1}{6},$$

$$P(X = 5) = f(5) = \frac{1}{6},$$

$$P(X = 6) = f(6) = \frac{1}{6}$$

$$\begin{aligned} E[X] &= \sum_{x=1}^6 x f(x) = 1 * \frac{1}{6} + 2 * \frac{1}{6} + 3 * \frac{1}{6} + 4 * \frac{1}{6} + 5 * \frac{1}{6} + 6 * \frac{1}{6} \\ &= \frac{7}{2} \end{aligned}$$

Buradan anıcağımız, bir X rastgele değişkeninin gelmesi olasılığı ortalama $7/2$ atıştan sonra gelmektedir. \square

Örnek 8.3. Bir para 3 kez atılsın. Bulunan turaların sayısı için beklenen değer nedir?

Çözüm:

S	YYY	YYT	YTY	TTY	YTT	TYT	TTY	TTT
X	0	1	1	1	2	2	2	3
$P(X = x)$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

$$f(0) = P(X = 0) = \frac{1}{8}, f(1) = P(X = 1) = \frac{3}{8}$$

$$f(2) = P(X = 2) = \frac{3}{8}, f(3) = P(X = 3) = \frac{1}{8}$$

$$E[X] = \sum_x x f(x) = 0 * \frac{1}{8} + 1 * \frac{3}{8} + 2 * \frac{3}{8} + 3 * \frac{1}{8} = \frac{3}{2}$$

\square

Örnek 8.4. Bir oyuncu bir zar atar. Eğer asal sayı gelirse gelen sayı kadar dolar kazanır, eğer asal sayı gelmezse gelen sayı kadar dolar kaybeder. Oyunun beklenen değeri nedir?

Çözüm:

2	3	5	-1	-4	-6
$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

$$E[X] = -1 * \frac{1}{6} + 2 * \frac{1}{6} + 3 * \frac{1}{6} - 4 * \frac{1}{6} + 5 * \frac{1}{6} - 6 * \frac{1}{6} = -\frac{1}{6}$$

Bu beklenen değer ile bu oyunun oyuncu için uygun olmadığı sonucunu veya her durumda oyuncunun kesin para kaybedeceği sonucunu çıkarabiliriz. \square

Örnek 8.5. İçinde üç beyaz ve iki siyah top bulunan bir kavanozdan yerine koyarak rasgele iki top çekilmiştir. Çekilen her beyaz top için 100 lira kazanılacak ve çekilen her siyah top için 50 lira kaybedilecektir. Oyunda beklenen kar nedir?

Çözüm: B ile çekilen beyaz topu, S ile çekilen siyah topu gösterelim:

$$\begin{aligned}
 P(BB) &= \frac{3}{5} * \frac{3}{5} = \frac{9}{25} (100 + 100 = +200) \\
 P(BS) &= \frac{3}{5} * \frac{2}{5} = \frac{6}{25} (100 - 50 = +50) \\
 P(SB) &= \frac{2}{5} * \frac{3}{5} = \frac{6}{25} (-50 + 100 = +50) \\
 P(SS) &= \frac{2}{5} * \frac{2}{5} = \frac{4}{25} (-50 - 50 = -100) \\
 E[X] &= 200 * \frac{9}{25} + 50 * \frac{6}{25} + 50 * \frac{6}{25} - 100 * \frac{4}{25} \\
 &= 80
 \end{aligned}$$

□

Örnek 8.6. X rasgele değişkeni ilk (6,6) gelinceye kadar bir çift zarın atışlarının sayısı olsun. Buna göre beklenen değeri bulunuz.

Çözüm:

X. atışta gelmesi	1 (ilk atışta (6,6) gelmesi)	2	3	4	...	n
P(X)	$\frac{1}{36}$	$\frac{35}{36} * \frac{1}{6}$	$\frac{35}{36} * \frac{35}{36} * \frac{1}{6}$	$\left(\frac{35}{36}\right)^3 * \frac{1}{36}$...	$\left(\frac{35}{36}\right)^{n-1} * \frac{1}{36}$

$$\begin{aligned}
 E[X] &= \sum_x x f(x) \\
 &= 1 * \frac{1}{36} + 2 * \frac{35}{36} * \frac{1}{6} + 3 * \left(\frac{35}{36}\right)^2 * \frac{1}{36} + 4 * \left(\frac{35}{36}\right)^3 * \frac{1}{36} + \dots + n * \left(\frac{35}{36}\right)^{n-1} * \frac{1}{36} + \dots \\
 &= \sum_{n=1}^{\infty} n * \left(\frac{35}{36}\right)^{n-1} * \frac{1}{36} = \frac{1}{36} \sum_{n=1}^{\infty} n * \left(\frac{35}{36}\right)^{n-1}
 \end{aligned}$$

$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$, $x = \frac{35}{36}$ serisini ele alalım. Seri $x < 1$ olduğundan düzgün yakınsaktır. Buna göre

$$\frac{d}{dx} \left(\sum_{n=0}^{\infty} x^n \right) = \sum_{n=1}^{\infty} n x^{n-1} \Rightarrow \frac{d}{dx} \left(\frac{1}{1-x} \right) = \sum_{n=1}^{\infty} n x^{n-1} = \frac{1}{(x-1)^2}$$

o halde

$$E[X] = \frac{1}{36} * \frac{1}{\left(\frac{35}{36} - 1\right)^2} = 36$$

sonucunu elde ederiz.

□

Örnek 8.7. 3 ü arızalı 7 ürün içerisinde 3 ürün seçilecektir. Arızalı olmayan bir ürünün beklenti değeri nedir?

Çözüm: X kesikli rasgele değişkeni arızasız ürünün seçilmesini göstermek üzere X kesikli rasgele değişkeninin olasılık dağılım fonksiyonu

$$\begin{aligned}
 f(x) &= \frac{C_{4,x} * C_{3,3-x}}{C_{7,3}} \Rightarrow \\
 f(0) &= \frac{C_{4,0} * C_{3,3}}{C_{7,3}} = \frac{\frac{4!}{4! * 0!} * \frac{3!}{0! * 3!}}{\frac{7!}{4! * 3!}} = \frac{1}{35} \\
 f(1) &= \frac{C_{4,1} * C_{3,2}}{C_{7,3}} = \frac{\frac{4!}{3! * 1!} * \frac{3!}{1! * 2!}}{\frac{7!}{4! * 3!}} = \frac{12}{35} \\
 f(2) &= \frac{C_{4,2} * C_{3,1}}{C_{7,3}} = \frac{\frac{4!}{2! * 2!} * \frac{3!}{2! * 1!}}{\frac{7!}{4! * 3!}} = \frac{18}{35} \\
 f(3) &= \frac{C_{4,3} * C_{3,0}}{C_{7,3}} = \frac{\frac{4!}{1! * 3!} * \frac{3!}{3! * 0!}}{\frac{7!}{4! * 3!}} = \frac{4}{35} \\
 \mu_X &= E[X] = 0 * f(0) + 1 * f(1) + 2 * f(2) + 3 * f(3) \\
 &= 0 * \left(\frac{1}{35}\right) + 1 * \left(\frac{12}{35}\right) + 2 * \left(\frac{18}{35}\right) + 3 * \left(\frac{4}{35}\right) \\
 &= \frac{12}{7} = 1.7143
 \end{aligned}$$

yani 3 ü arızalı 7 ürün içerisinde ortalama 1.7 civarında seçimden sonra arızasız ürün seçilir. \square

Örnek 8.8. Bir satış temsilcisi aynı günde 2 randevusu vardır. İlk randevusunda %70 oranında anlaşma yapacağını ve anlaşması durumunda 1000TL komisyon alacaktır. Diğer yandan 2. şirketle %40 oranında anlaşma yapma olasılığını olduğu ve anlaşma yapması durumunda ise 1500TL komisyon alacaktır. Buna göre beklenen komisyon değeri nedir? Her bir randevunun birbirinden bağımsız olduğunu varsayınız.

Çözüm: 4 türlü olasılık vardır. 0TL kazanması, 1000TL, 1500TL ve 2500TL kazanması durumlarına göre olasılık dağılım fonksiyonları:

$$\begin{aligned}
 f(0TL) &= (1 - 0.70) * (1 - 0.40) = 0.18 \\
 f(1000TL) &= 0.70 * (1 - 0.40) = 0.42 \\
 f(1500TL) &= (1 - 0.70) * 0.40 = 0.12 \\
 f(2500TL) &= 0.70 * 0.40 = 0.28
 \end{aligned}$$

Buna göre beklene komisyon değeri

$$\begin{aligned}
 \mu_X &= E[X] = 0 * f(0TL) + 1000 * f(1000TL) \\
 &\quad + 1500 * f(1500TL) + 2500 * f(2500TL) \\
 &= 0 * (0.18) + 1000 * 0.42 \\
 &\quad + 1500 * 0.12 + 2500 * 0.28 \\
 &= 1300TL
 \end{aligned}$$

\square

Örnek 8.9. *Olasılık dağılım fonksiyonu*

$$f(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ 1 & \text{if } 0 < x < 1 \\ 0 & \text{if } x \geq 1 \end{cases}$$

olarak verilen X sürekli rasgele değişkeninin beklenen değeri nedir?

Çözüm:

$$\begin{aligned} \mu_X &= E[X] = \int_{-\infty}^{\infty} x f(x) dx \\ &= \int_0^1 x \cdot 1 dx \\ &= \left. \frac{x^2}{2} \right|_{x=0}^{x=1} = \frac{1}{2} \end{aligned}$$

□

Örnek 8.10. *Olasılık dağılım fonksiyonu*

$$f(x) = \begin{cases} \frac{20000}{x^3} & \text{if } x > 100, x \in \mathbb{R} \\ 0 & \text{otherwise} \end{cases}$$

olarak verilen X sürekli rasgele değişkeninin beklenen değeri nedir?

Çözüm:

$$\begin{aligned} \mu_X &= E[X] = \int_{-\infty}^{\infty} x f(x) dx \\ &= \int_{100}^{\infty} x \frac{20000}{x^3} dx \\ &= \int_{100}^{\infty} \frac{20000}{x^2} dx \\ &= -20000 \left. \frac{1}{x} \right|_{x=100}^{x=\infty} \\ &= 200 \end{aligned}$$

□

Tanım 8.11. X kesikli rasgele değişkeninin olasılık dağılım fonksiyonu $f(x)$ olmak üzere $g(X)$ rasgele değişkeninin beklenti değeri

$$\mu_{g(X)} = E[g(X)] = \sum_x g(x) f(x)$$

olarak tanımlanırken, X sürekli rasgele değişkeninin olasılık dağılım fonksiyonu $f(x)$ olmak üzere $g(X)$ rasgele değişkeninin beklenti değeri

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

olarak tanımlanır.

Örnek 8.12. X rasgele değişkenin olasılık dağılımı aşağıdaki gibi veriliyor.

x	1	2	3	4	5	6
$f(x) = P(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Buna göre $E[X^2]$, $E[2X]$, $E[2X - 1]$ beklenti değerlerini hesaplayınız.

Çözüm:

x	1	2	3	4	5	6
$f(x) = P(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
$g(X) = X^2$	1	$2^2 = 4$	9	16	25	36
$h(X) = 2X$	2	4	6	8	10	12
$q(X) = 2X - 1$	1	3	5	7	9	11

$$E[X^2] = \sum_x g(x) f(x) = 1 * \frac{1}{6} + 4 * \frac{1}{6} + 9 * \frac{1}{6} + 16 * \frac{1}{6} + 25 * \frac{1}{6} + 36 * \frac{1}{6} = \frac{91}{6}$$

$$E[2X] = \sum_x h(x) f(x) = 2 * \frac{1}{6} + 4 * \frac{1}{6} + 6 * \frac{1}{6} + 8 * \frac{1}{6} + 10 * \frac{1}{6} + 12 * \frac{1}{6} = 7$$

$$E[2X - 1] = \sum_x q(x) f(x) = 1 * \frac{1}{6} + 3 * \frac{1}{6} + 5 * \frac{1}{6} + 7 * \frac{1}{6} + 9 * \frac{1}{6} + 11 * \frac{1}{6} = 6$$

□

Örnek 8.13. Güneşli bir cuma günü, saat 16:00-17:00 arası araba yıkamaya giden arabaların sayısı X olmak üzere olasılık dağılımı aşağıdaki gibi veriliyor:

x	4	5	6	7	8	9
$f(x) = P(X = x)$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$

$g(X) = 5(2X - 1)$ ile arabaların odemiş oldukları ücret verilmektedir. Verilen bu zaman diliminde kazanılması beklenen para miktarı nedir?

Çözüm:

$$\begin{aligned}
 \mu_{g(X)} &= E[g(X)] = \sum_{x=4}^9 g(x) f(x) \\
 &= 5 * (2 * 4 - 1) * \frac{1}{12} + 5 * (2 * 5 - 1) * \frac{1}{12} \\
 &\quad + 5 * (2 * 6 - 1) * \frac{1}{4} + 5 * (2 * 7 - 1) * \frac{1}{4} \\
 &\quad + 5 * (2 * 8 - 1) * \frac{1}{6} + 5 * (2 * 9 - 1) * \frac{1}{6} \\
 &= \frac{190}{3} = 63.333
 \end{aligned}$$

□

Örnek 8.14. X sürekli rasgele değişkeninin olasılık dağılımı aşağıdaki gibi veriliyor:

$$f(x) = \begin{cases} \frac{x^2}{3} & \text{if } -1 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

$g(X) = 4X + 3$ rasgele değişkeninin beklentisi nedir?

Çözüm:

$$\begin{aligned} \mu_{g(X)} &= E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx \\ &= \int_{-1}^2 (4x + 3) \frac{x^2}{3} dx \\ &= \frac{1}{3} x^3 (x + 1) \Big|_{x=-1}^{x=2} \\ &= \frac{1}{3} * (2^3 * 3 - (-1)^3 (-1 + 1)) \\ &= 8 \end{aligned}$$

□

Tanım 8.15. X ve Y kesikli rasgele değişkenler, $f(x, y)$ birleşik olasılık yoğunluk fonksiyonu olmak üzere $g(X, Y)$ rasgele değişkeninin beklenti değeri

$$\mu_{g(X, Y)} = E[g(X, Y)] = \sum_x \sum_y g(x, y) f(x, y)$$

olarak tanımlanırken, X ve Y sürekli rasgele değişkenler, $f(x, y)$ birleşik olasılık yoğunluk fonksiyonu olmak üzere $g(X, Y)$ rasgele değişkeninin beklenti değeri

$$\mu_{g(X, Y)} = E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) dx dy$$

olarak tanımlanır.

Örnek 8.16. X ve Y kesikli rasgele değişkenlerin $f(x, y)$ birleşik olasılık yoğunluk fonksiyonu aşağıdaki gibi verilsin:

$f(x, y)$	x		
y	0	1	2
0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$
1	$\frac{3}{14}$	$\frac{3}{14}$	0
2	$\frac{1}{28}$	0	0

Buna göre $g(X, Y) = XY$ rasgele değişkeninin beklenti değerini hesaplayınız.

Çözüm:

$$\begin{aligned}
 \mu_{g(X,Y)} &= E[g(X,Y)] = \sum_{x=0}^2 \sum_{y=0}^2 g(x,y) f(x,y) \\
 &= (0)(0)f(0,0) + (0)(1)f(0,1) + (0)(2)f(0,2) \\
 &\quad + (1)(0)f(1,0) + (1)(1)f(1,1) + (1)(2)f(1,2) \\
 &\quad + (2)(0)f(2,0) + (2)(1)f(2,1) + (2)(2)f(2,2) \\
 &= (0)(0)\frac{3}{28} + (0)(1)\frac{3}{14} + (0)(2)\frac{1}{28} \\
 &\quad + (1)(0)\frac{9}{28} + (1)(1)\frac{3}{14} + (1)(2)0 \\
 &\quad + (2)(0)\frac{3}{28} + (2)(1)0 + (2)(2)0 \\
 &= \frac{3}{14} = 0.21429
 \end{aligned}$$

□

Örnek 8.17. X ve Y sürekli rasgele değişkeninin olasılık dağılımı aşağıdaki gibi veriliyor:

$$f(x,y) = \begin{cases} \frac{x(1+3y^2)}{4} & \text{if } 0 < x < 2, 0 < y < 1, x, y \in \mathbb{R} \\ 0 & \text{otherwise} \end{cases}$$

$E\left[\frac{Y}{x}\right]$ beklentisi nedir?

Çözüm:

$$\begin{aligned}
 g(X,Y) &= \frac{Y}{x} \\
 \mu_{g(X,Y)} &= E[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f(x,y) dx dy \\
 &= \int_0^2 \int_0^1 \frac{y}{x} \frac{x(1+3y^2)}{4} dy dx \\
 &= \frac{1}{4} \int_0^2 \frac{y^2}{2} + \frac{3y^4}{4} \Big|_{y=0}^{y=1} dx \\
 &= \frac{1}{4} \int_0^2 \left(\frac{1}{2} + \frac{3}{4} \right) dx \\
 &= \frac{1}{4} * 2 * \frac{5}{4} = \frac{5}{8} = 0.625
 \end{aligned}$$

□

9 Rasgele değişkenlerinin varyansı ve kovaryansı (Variance and Covariance of Random Variables)

X rasgele değişkeninin beklenen değeri (veya ortalama değeri) istatistikte çok önemlidir çünkü olasılık dağılımının merkezini verir fakat dağılımın şekli hakkında bize tam net bilgi vermemektedir. Örneğin 1, 2, 3

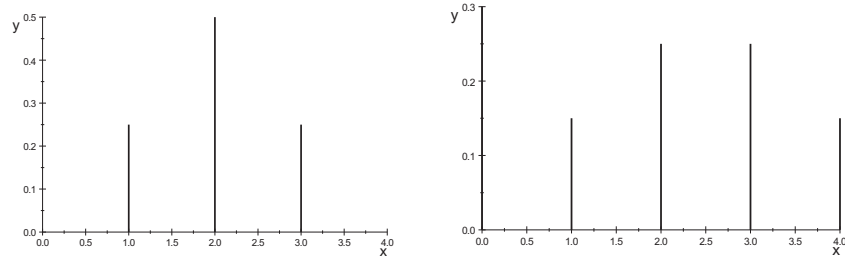
numaralı makineler 1, 2, 3 saatte işi bitirme kapasitesine sahipken seçilme olasılığı sırasıyla 0.25, 0.5, 0.25 iken işin bitme süresi ortalama olarak

$$E[X] = 1 * 0.25 + 2 * 0.5 + 3 * 0.25 = 2$$

hesaplanırken başka bir deneyde 0, 1, 2, 3, 4 nolu makineler 0, 1, 2, 3, 4 saatte işi bitirme kapasitesine sahipken seçilme olasılığı sırasıyla 3/10, 3/20, 1/4, 1/4, 3/20 olarak verilirken işin bitme süresi ortalama olarak

$$E[X] = 0 * \frac{3}{10} + 1 * \frac{3}{20} + 2 * \frac{1}{4} + 3 * \frac{1}{4} + 4 * \frac{3}{20} = 2$$

olarak hesaplanmaktadır. Sonuçta her iki deneyde de ortalama iş bitirilme süresi 2 saattir. Ancak aynı ortalama değeri bulmamıza rağmen bu sonuç iki deneydeki değerlerin değişme niteliği veya dağılımı hakkında kesin bilgi verememektedir.



Bir başka örnek olarak ortalama 500\$ lık gelir elde edeceğiniz bir işe yatırım yapacaksınız. Buna göre 450\$ yatırım yaptığınızda %50 oranında , 550\$ yatırım yaparkende %50 oranında bu geliri elde etme olasılığınız var ki ortalama olarak $450 * \frac{50}{100} + 550 * \frac{50}{100} = 500$ \$ geliri elde etmiş oluyorsunuz. Bu durumda 50\$ kazancı %50 olasılıkla ve 50\$ zarar içinde %50 olasılığınız vardır. O halde işe yatırım yapmada herhangi bir tereddüt yaşamazsınız. Fakat şöyle bir durum olsaydı : 0\$ yatırım yaptığınızda %50 oranında , 1000\$ yatırım yaparkende %50 oranında bu geliri elde etme olasılığınız var ki ortalama olarak $0 * \frac{50}{100} + 1000 * \frac{50}{100} = 500$ \$ geliri elde etmiş oluyorsunuz. Bu durumda 500\$ kazancı %50 olasılıkla ve 500\$ zarar içinde %50 olasılığınız vardır. O halde bu işe yatırım yapmada pek çok kişi tereddüt yaşayabilir. Bunun için X rasgele değişkeninin dağılım şekline varyans (değişiklik) denilcek ve beklenen değer ile standart sapmasının karesinin beklenen değeri olarak tanımlanacak.

Tanım 9.1. X kesikli veya sürekli rasgele değişkenler, $f(x)$ olasılık yoğunluk fonksiyonu için, X in varyansı (The variance of X)

$$\sigma_X^2 = \text{Var}(X) = E[(X - \mu_x)^2] = \sum_x (x - \mu_x)^2 f(x), \quad X \text{ kesikli ise,}$$

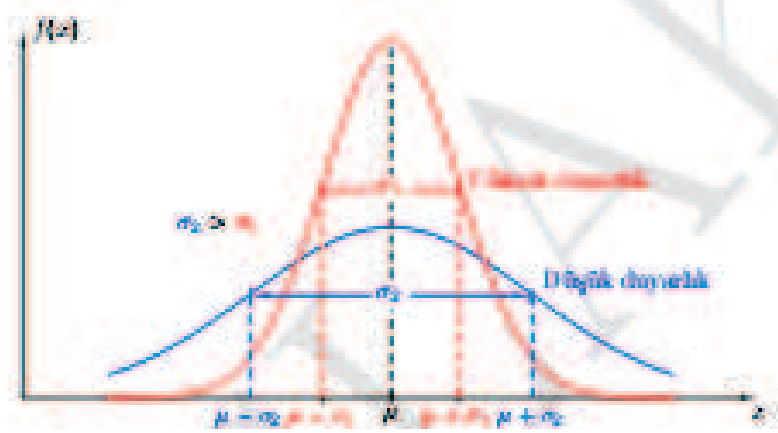
$$\sigma_X^2 = \text{Var}(X) = E[(X - \mu_x)^2] = \int_{-\infty}^{\infty} (x - \mu_x)^2 f(x) dx, \quad X \text{ sürekli ise,}$$

$$\mu_x = E[X],$$

olarak tanımlanır. $\text{Var}(X)$ in pozitif kareköküne ise X in standart sapması (standard deviation of X) denir.

$$\sigma_X = \sqrt{\text{Var}(X)} \rightarrow \text{standart sapma}$$

Büyük standart sapma değerleri X rastgele değişkeninin μ ye göre daha geniş alana yayıldığını, küçük olanlar ise saçılmanın daraldığını gösterir. Buradan da, bir ölçme uygulamasında standart sapmanın, sonuçların yüksek yada düşük duyarlılık (precision) olarak yorumlanmasına sebep olmaktadır..



Örnek 9.2. Bir para üç kez atıldığında gelebilecek turların sayısını X rasgele değişkeni ile gösterelim. Buna göre Varyansı ve standart sapmayı hesaplayınız.

Çözüm:

S	YYY	YYT	YTY	TTY	YTT	TYT	TTY	TTT
X	0	1	1	1	2	2	2	3
$P(X = x)$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
$f(x)$	$\frac{1}{8}$	$\frac{3}{8}$			$\frac{3}{8}$		$\frac{1}{8}$	

$$\Rightarrow \mu_x = E[X] = 0 * \frac{1}{8} + 1 * \frac{3}{8} + 2 * \frac{3}{8} + 3 * \frac{1}{8} = \frac{3}{2}$$

X	0	1	2	3
$f(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$
$(x - \mu_x)^2$	$(-\frac{3}{2})^2 = \frac{9}{4}$	$(1 - \frac{3}{2})^2 = \frac{1}{4}$	$(2 - \frac{3}{2})^2 = \frac{1}{4}$	$(3 - \frac{3}{2})^2 = \frac{9}{4}$
$(x - \mu_x)^2 f(x)$	$\frac{9}{4} * \frac{1}{8} = \frac{9}{32}$	$\frac{1}{4} * \frac{3}{8} = \frac{3}{32}$	$\frac{1}{4} * \frac{3}{8} = \frac{3}{32}$	$\frac{9}{4} * \frac{1}{8} = \frac{9}{32}$

$$\sigma_X^2 = Var(X) = \sum_x (x - \mu_x)^2 f(x) = \frac{9}{32} + \frac{3}{32} + \frac{3}{32} + \frac{9}{32} = \frac{3}{4}$$

$$\sigma_X = \sqrt{\frac{3}{4}} = \frac{\sqrt{3}}{2}$$

□

Örnek 9.3. X rasgele değişkeni, bir firmada iş-günlerinde ofis işleri için kullanılan arabaların sayısını göstermek üzere A firması için olasılık dağılım fonksiyonu

x	1	2	3
$f(x)$	0.3	0.4	0.3

B firması için olasılık dağılım fonksiyonu

x	0	1	2	3	4
$f(x)$	0.2	0.1	0.3	0.3	0.1

olarak verildiğine göre B şirketi için olasılık dağılımın varyansının A şirketinde büyük olduğunu gösteriniz.

Çözüm: Aslında dağılım fonksiyonundan da görüleceği üzere B şirketinde kullanılan otomobillerin çeşitliliği (varyansı) A şirketinden daha fazladır ancak şimdi bunu formülüne göre gösterelim.

A şirketi için öncelikle $\mu = E[X]$ beklenen değerini bulalım:

$$\mu = E[X] = 1 * 0.3 + 2 * 0.4 + 3 * 0.3 = 2$$

Buna göre A şirketi için varyans

$$\begin{aligned} Var(X) &= E[(X - \mu)^2] = \sum_{x=1}^3 (x - \mu)^2 f(x) \\ &= (1 - 2)^2 f(1) + (2 - 2)^2 f(2) + (3 - 2)^2 f(3) \\ &= (1 - 2)^2 * 0.3 + (2 - 2)^2 * 0.4 + (3 - 2)^2 * 0.3 \\ &= 0.6 \end{aligned}$$

B şirketi için öncelikle $\mu = E[X]$ beklenen değerini bulalım:

$$\mu = E[X] = 0 * 0.2 + 1 * 0.1 + 2 * 0.3 + 3 * 0.3 + 4 * 0.1 = 2$$

Buna göre B şirketi için varyans

$$\begin{aligned} Var(X) &= E[(X - \mu)^2] = \sum_{x=0}^4 (x - \mu)^2 f(x) \\ &= (0 - 2)^2 f(0) + (1 - 2)^2 f(1) + (2 - 2)^2 f(2) + (3 - 2)^2 f(3) + (4 - 2)^2 f(4) \\ &= (0 - 2)^2 * 0.2 + (1 - 2)^2 * 0.1 + (2 - 2)^2 * 0.3 + (3 - 2)^2 * 0.3 + (4 - 2)^2 * 0.1 \\ &= 1.6 \end{aligned}$$

□

Teorem 9.4. X kesikli veya sürekli rasgele değişkenler, $f(x)$ olasılık yoğunluk fonksiyonu için

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

eşitliği sağlanır.

İspat: X kesikli ise

$$\begin{aligned} Var(X) &= E[(X - \mu)^2] = \sum_x (x - \mu)^2 f(x), \\ &= \sum_x (x^2 - 2\mu x + \mu^2) f(x), \mu - sabit \\ &= \sum_x x^2 f(x) - 2\mu \underbrace{\sum_x x f(x)}_{=\mu=E[X]} + \mu^2 \underbrace{\sum_x f(x)}_{=1} \quad (f(x) - \text{olasılık} - \text{yoğunluk} - \text{fonksiyonu}) \\ &= E[X^2] - 2\mu^2 + \mu^2 = E[X^2] - \mu^2 \end{aligned}$$

X sürekli ise

$$\begin{aligned}
 Var(X) &= E[(X - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx \\
 &= \int_{-\infty}^{\infty} (x^2 - 2\mu x + \mu^2) f(x) dx, \mu - sabit \\
 &= \int_{-\infty}^{\infty} x^2 f(x) dx - 2\mu \underbrace{\int_{-\infty}^{\infty} x f(x) dx}_{=\mu=E[X]} + \mu^2 \underbrace{\int_{-\infty}^{\infty} f(x) dx}_{=1} (f(x) - pdf) \\
 &= E[X^2] - 2\mu^2 + \mu^2 = E[X^2] - \mu^2
 \end{aligned}$$

Örnek 9.5. X bir makinenin bozuk olan parçalarını göstermek üzere, aşağıdaki olasılık yoğunluk fonksiyonuna sahiptir. Buna göre $Var(X)$ değerini hesaplayınız.

x	0	1	2	3
$f(x)$	0.51	0.38	0.1	0.01

Çözüm: I.Yol:

x	0	1	2	3
$f(x)$	0.51	0.38	0.1	0.01
$x * f(x)$	$0 * 0.51 = 0$	$1 * 0.38 = 0.38$	$2 * 0.1 = 0.2$	$3 * 0.01 = 0.03$
	$\mu_x = E[X] = 0 + 0.38 + 0.2 + 0.03 = 0.61$			
$(x - \mu_x)^2$	$(0 - 0.61)^2 = 0.3721$	$(1 - 0.61)^2 = 0.1521$	$(2 - 0.61)^2 = 1.9321$	$(3 - 0.61)^2 = 5.7121$
$(x - \mu_x)^2 f(x)$	$0.51 * 0.3721 = 0.18977$	$0.38 * 0.1521 = 5.7798 \times 10^{-2}$	$0.1 * 1.9321 = 0.19321$	$0.01 * 5.7121 = 5.7121 \times 10^{-2}$
	$Var(X) = E[(X - \mu_X)^2] = \sum_{x=0}^4 (x - \mu_x)^2 f(x) = 0.18977 + 5.7798 \times 10^{-2} + 0.19321 + 5.7121 \times 10^{-2}$			
	$Var(X) = 0.4979$			

II.Yol-Teorem 9.4 ı kullanarak :

x	0	1	2	3
$f(x)$	0.51	0.38	0.1	0.01
$x * f(x)$	$0 * 0.51 = 0$	$1 * 0.38 = 0.38$	$2 * 0.1 = 0.2$	$3 * 0.01 = 0.03$
	$\mu_x = E[X] = 0 + 0.38 + 0.2 + 0.03 = 0.61$			
x^2	0	1	4	9
$x^2 * f(x)$	$0.51 * 0 = 0$	$0.38 * 1 = 0.38$	$0.1 * 4 = 0.4$	$0.01 * 9 = 0.09$
	$E[X^2] = 0 + 0.38 + 0.4 + 0.09 = 0.87$			
	$Var(X) = E[X^2] - \mu_X^2 = 0.87 - 0.61^2 = 0.4979$			

□

Örnek 9.6. Düzgün 6 yüzlü bir zar atıldığında üst yüzde görünecek sayı X olmak üzere $Var(X)$ değerini hesaplayınız.

Çözüm:

x	$f(x) = P(X = x)$	$x * f(x)$	$x^2 * f(x)$
1	$\frac{1}{6}$	$\frac{1}{6}$	$1^2 * \frac{1}{6} = \frac{1}{6}$
2	$\frac{1}{6}$	$\frac{2}{6}$	$2^2 * \frac{1}{6} = \frac{4}{6}$
3	$\frac{1}{6}$	$\frac{3}{6}$	$3^2 * \frac{1}{6} = \frac{9}{6}$
4	$\frac{1}{6}$	$\frac{4}{6}$	$4^2 * \frac{1}{6} = \frac{16}{6}$
5	$\frac{1}{6}$	$\frac{5}{6}$	$5^2 * \frac{1}{6} = \frac{25}{6}$
6	$\frac{1}{6}$	$\frac{6}{6}$	$6^2 * \frac{1}{6} = \frac{36}{6}$
		$Toplam : \mu_x = E[X] = \frac{21}{6}$	$Toplam : E[X^2] = \frac{91}{6}$

$$\sigma_X^2 = Var(X) = E[X^2] - \mu_X^2 = \frac{91}{6} - \left(\frac{21}{6}\right)^2 = \frac{35}{3}$$

□

Örnek 9.7. İçinde 1 den N ye kadar numaranmış N top bulunan bir kavanozdan rasgele bir top seçiliyor. X rasgele değişkeni top üzerindeki sayı olsun. Buna göre $E[X]$, $E[X^2]$ ve $Var(X)$ değerlerini hesaplayınız.

Çözüm:

x	$f(x) = P(X = x)$	$x * f(x)$	$x^2 * f(x)$
1	$\frac{1}{N}$	$\frac{1}{N}$	$1^2 * \frac{1}{N}$
2	$\frac{1}{N}$	$\frac{2}{N}$	$2^2 * \frac{1}{N}$
\vdots	\vdots	\vdots	\vdots
N	$\frac{1}{N}$	$\frac{N}{N}$	$N^2 * \frac{1}{N}$
		$Toplam : \mu_x = E[X] = \frac{1+2+\dots+N}{N}$ $= \frac{N(N+1)}{2N} = \frac{N+1}{2}$	$Toplam : E[X^2]$ $= \frac{1^2+2^2+\dots+N^2}{N}$ $= \frac{N(N+1)(2N+1)/6}{N}$ $= \frac{(N+1)(2N+1)}{6}$

$$\sigma_X^2 = Var(X) = E[X^2] - \mu_X^2 = \frac{(N+1)(2N+1)}{6} - \left(\frac{N+1}{2}\right)^2 = \frac{1}{12}N^2 - \frac{1}{12}$$

□

Örnek 9.8. İçme suyu üreten bir şirketin İstanbul şubesindeki haftalık depo ihtiyaçları X sürekli rasgele değişkeni olarak gösterirsek, olasılık dağılım fonksiyonu

$$f(x) = \begin{cases} 2(x-1) & \text{if } 1 < x < 2, x \in \mathbb{R} \\ 0 & \text{otherwise} \end{cases}$$

olarak verilir. Buna göre X rasgele değişkeninin beklenen değeri (ortalama olarak ihtiyacı) ve varyansını bulunuz.

Çözüm: 1.yol:

$$\begin{aligned}
 \mu &= E[X] = \int_{-\infty}^{\infty} x f(x) dx = \int_1^2 x * 2(x-1) dx \\
 &= \left. \frac{1}{3} x^2 (2x-3) \right|_{x=1}^{x=2} = \frac{4}{3} - \left(-\frac{1}{3} \right) = \frac{5}{3} \\
 Var(X) &= E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x) dx \\
 &= \int_1^2 \left(x - \frac{5}{3} \right)^2 * 2(x-1) dx \\
 &= \left. \frac{1}{2} x^4 - \frac{26}{9} x^3 + \frac{55}{9} x^2 - \frac{50}{9} x \right|_{x=1}^{x=2} \\
 &= -\frac{16}{9} - \left(-\frac{11}{6} \right) = \frac{1}{18}
 \end{aligned}$$

2.Yol

$$\begin{aligned}
 \mu &= E[X] = \int_{-\infty}^{\infty} x f(x) dx = \int_1^2 x * 2(x-1) dx \\
 &= \left. \frac{1}{3} x^2 (2x-3) \right|_{x=1}^{x=2} = \frac{4}{3} - \left(-\frac{1}{3} \right) = \frac{5}{3} \\
 E[X^2] &= \int_{-\infty}^{\infty} x^2 f(x) dx = \int_1^2 x^2 * 2(x-1) dx \\
 &= \left. \frac{1}{6} x^3 (3x-4) \right|_{x=1}^{x=2} = \frac{8}{3} - \left(-\frac{1}{6} \right) = \frac{17}{6} \\
 Var(X) &= E[X^2] - \mu^2 = \frac{17}{6} - \left(\frac{5}{3} \right)^2 = \frac{1}{18}
 \end{aligned}$$

□

Tanım 9.9. X kesikli rasgele değişkeninin olasılık dağılım fonksiyonu $f(x)$ olmak üzere $g(X)$ rasgele değişkeninin varyansı

$$Var(g(X)) = E[(g(X) - \mu_{g(X)})^2] = \sum_x (g(X) - \mu_{g(X)})^2 f(x)$$

olarak tanımlanırken, X sürekli rasgele değişkeninin olasılık dağılım fonksiyonu $f(x)$ olmak üzere $g(X)$ rasgele değişkeninin varyansı

$$Var(g(X)) = E[(g(X) - \mu_{g(X)})^2] = \int_{-\infty}^{\infty} (g(X) - \mu_{g(X)})^2 f(x) dx$$

olarak tanımlanır. Burada $\mu_{g(X)} = E[g(X)]$ dır.

Örnek 9.10. X rasgele değişkeni olmak üzere $g(X) = 2X + 3$ rastlantı değişkenin varyansını aşağıdaki olasılık dağılım fonksiyonu ile birlikte hesaplayınız.

x	0	1	2	3
$f(x)$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{8}$

Çözüm:

x	0	1	2	3
$f(x)$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{8}$
$g(X) = 2X + 3$	3	5	7	9
$g(x)f(x)$	$3 * \frac{1}{4}$	$5 * \frac{1}{8}$	$7 * \frac{1}{2}$	$9 * \frac{1}{8}$
	$\mu_{g(X)} = E[g(X)] = \sum_x g(x)f(x)$			
	$\mu_{g(X)} = 3 * \frac{1}{4} + 5 * \frac{1}{8} + 7 * \frac{1}{2} + 9 * \frac{1}{8} = 6$			
$(g(X) - \mu_{g(X)})^2$	$(3 - 6)^2 = 9$	$(5 - 6)^2 = 1$	$(7 - 6)^2 = 1$	$(9 - 6)^2 = 9$
$(g(x) - \mu_{g(x)})^2 f(x)$	$9 * \frac{1}{4}$	$1 * \frac{1}{8}$	$1 * \frac{1}{2}$	$9 * \frac{1}{8}$
	$Var(g(X)) = E[(g(X) - \mu_{g(X)})^2] = \sum_x (g(X) - \mu_{g(X)})^2 f(x)$			
	$Var(g(X)) = 9 * \frac{1}{4} + 1 * \frac{1}{8} + 1 * \frac{1}{2} + 9 * \frac{1}{8} = 4$			

□

Örnek 9.11. X sürekli rasgele değişkeninin olasılık dağılımı aşağıdaki gibi veriliyor:

$$f(x) = \begin{cases} \frac{x^2}{3} & \text{if } -1 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

 $g(X) = 4X + 3$ rasgele değişkeninin varyansı nedir?**Çözüm:**

$$\begin{aligned}
\mu_{g(X)} &= E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx \\
&= \int_{-1}^2 (4x+3) \frac{x^2}{3} dx \\
&= \frac{1}{3} x^3 (x+1) \Big|_{x=-1}^{x=2} \\
&= \frac{1}{3} * (2^3 * 3 - (-1)^3 (-1+1)) \\
&= 8 \\
Var(g(X)) &= E[(g(X) - \mu_{g(X)})^2] = \int_{-\infty}^{\infty} (g(X) - \mu_{g(X)})^2 f(x) dx \\
&= \int_{-1}^2 (4x+3-8)^2 \frac{x^2}{3} dx = \frac{16}{15} x^5 - \frac{10}{3} x^4 + \frac{25}{9} x^3 \Big|_{x=-1}^{x=2} \\
&= \frac{136}{45} - \left(-\frac{323}{45}\right) = \frac{51}{5} \\
f(x) &= \frac{16}{15} x^5 - \frac{10}{3} x^4 + \frac{25}{9} x^3
\end{aligned}$$

□

Tanım 9.12. X ve Y kesikli rasgele değişkenlerinin birleşik olasılık dağılım fonksiyonu $f(x, y)$ olmak üzere X ve Y nin kovaryansı (covariance of X and Y)

$$Cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] = \sum_x \sum_y (x - \mu_X)(y - \mu_Y) f(x, y)$$

olarak tanımlanırken, X ve Y sürekli rasgele değişkenlerinin birleşik olasılık dağılım fonksiyonu $f(x, y)$ olmak üzere X ve Y nin kovaryansı

$$\text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y) f(x, y) dx dy$$

olarak tanımlanır. Burada $\mu_X = E[X]$, $\mu_Y = E[Y]$ dır.

İki rasgele değişkeninin kovaryansı, bu ikisi arasındaki doğal bir durumdur. Eğer X in büyük değerleri, Y nin büyük değerleri ile sonuçlanıyorsa $(x - \mu_X)$ farkının pozitifliği $(y - \mu_Y)$ nin de pozitif olması sonucunu vericektir ve $(x - \mu_X)(y - \mu_Y)$ çarpımı da pozitif olacaktır veya tersine X in küçük değerleri, Y nin küçük değerleri ile sonuçlanıyorsa $(x - \mu_X)$ farkının negatifliği $(y - \mu_Y)$ nin de negatif olması sonucunu vericektir ve $(x - \mu_X)(y - \mu_Y)$ çarpımı yine de pozitif olacaktır. Diğer yandan X in büyük değerleri, Y nin küçük değerleri ile sonuçlanıyorsa $(x - \mu_X)$ farkının pozitifliği $(y - \mu_Y)$ nin de negatif olması sonucunu ile $(x - \mu_X)(y - \mu_Y)$ çarpımı da negatif olacaktır. Böylece kovaryansın işareti, iki bağımsız rasgele değişkeninin arasındaki bağıntının pozitif veya negatifliğini göstermektedir. Değişkenler istatistiksel bağımsız olduğunda ise kovaryans sıfır olacaktır. Bunun tersi her zaman doğru olmayabilir. Yani kovaryansı sıfır olan iki değişken istatistiksel olarak lineer bağımsız olmak zorunda değildir.

Teorem 9.13. X ve Y kesikli (sürekli) rasgele değişkenlerinin beklenti değerleri sırasıyla $\mu_X = E[X]$, $\mu_Y = E[Y]$ olmak üzere

$$\text{Cov}(X, Y) = E[XY] - \mu_X \mu_Y$$

bağıntısı doğrudur.

İspat: X ve Y kesikli rasgele değişkenlerinin birleşik olasılık dağılım fonksiyonu $f(x, y)$ olmak üzere X ve Y nin kovaryansı (covariance of X and Y)

$$\begin{aligned} \text{Cov}(X, Y) &= E[(X - \mu_X)(Y - \mu_Y)] = \sum_x \sum_y (x - \mu_X)(y - \mu_Y) f(x, y) \\ &= \sum_x \sum_y xyf(x, y) - \sum_x \sum_y \mu_X yf(x, y) - \sum_x \sum_y \mu_Y xf(x, y) + \sum_x \sum_y \mu_X \mu_Y f(x, y), \\ \mu_X, \mu_Y - \text{sabit}, \mu_Y &= E[Y] = \sum_x \sum_y yf(x, y), \mu_X = E[X] = \sum_x \sum_y xf(x, y), \\ &= \sum_x \sum_y xyf(x, y) - \underbrace{\mu_X \sum_x \sum_y yf(x, y)}_{=\mu_Y} - \underbrace{\mu_Y \sum_x \sum_y xf(x, y)}_{=\mu_X} + \underbrace{\mu_X \mu_Y \sum_x \sum_y f(x, y)}_{=1} \\ &= \sum_x \sum_y xyf(x, y) - 2\mu_X \mu_Y + \mu_X \mu_Y = \sum_x \sum_y xyf(x, y) - \mu_X \mu_Y = E[XY] - \mu_X \mu_Y \end{aligned}$$

olarak tanımlanırken, X ve Y sürekli rasgele değişkenlerinin birleşik olasılık dağılım fonksiyonu $f(x, y)$ olmak üzere X ve Y nin kovaryansı

$$\begin{aligned}
Cov(X, Y) &= E[(X - \mu_X)(Y - \mu_Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y) f(x, y) dx dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f(x, y) dx dy - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mu_X y f(x, y) dx dy - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mu_Y x f(x, y) dx dy + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mu_X \mu_Y f(x, y) dx dy \\
&\quad \mu_X, \mu_Y - \text{sabit}, \mu_Y = E[Y] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x, y) dy dx, \mu_X = E[X] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x, y) dx dy, \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f(x, y) dx dy - \underbrace{\mu_X \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x, y) dx dy}_{=\mu_Y} - \underbrace{\mu_Y \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x, y) dx dy}_{=\mu_X} + \underbrace{\mu_X \mu_Y \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy}_{=1} \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f(x, y) dx dy - 2\mu_X \mu_Y + \mu_X \mu_Y = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f(x, y) dx dy - \mu_X \mu_Y = E[XY] - \mu_X \mu_Y
\end{aligned}$$

Örnek 9.14. 2 renkli kalem, 3 mavi, 2 kırmızı ve 3 yeşil kalemin olduğu bir kutudan seçilecektir. Eğer X rasgele değişkeni mavi kalem seçimini, Y ise kırmızı kalem seçimini gösterirse X ve Y nin kovaryansını hesaplayınız.

Çözüm:

$f(x, y)$	x-mavi			h(y)
y-kırmızı	0	1	2	
0	$\frac{3!}{1! \cdot 2!} = \frac{3}{28}$	$\frac{3!}{2! \cdot 1!} \cdot \frac{3!}{2! \cdot 1!} = \frac{9}{28}$	$\frac{3!}{2! \cdot 1!} = \frac{3}{28}$	$\frac{3}{28} + \frac{9}{28} + \frac{3}{28} = \frac{15}{28} = h(0)$
1	$\frac{2!}{1! \cdot 1!} \cdot \frac{3!}{2! \cdot 1!} = \frac{3}{14}$	$\frac{2!}{1! \cdot 1!} \cdot \frac{3!}{1! \cdot 1!} = \frac{3}{14}$	0	$\frac{3}{14} + \frac{3}{14} + 0 = \frac{3}{7} = h(1)$
2	$\frac{2!}{0! \cdot 2!} = \frac{1}{28}$	0	0	$\frac{1}{28} + 0 + 0 = \frac{1}{28} = h(2)$
$g(x)$	$\frac{3}{28} + \frac{3}{14} + \frac{1}{28} = \frac{5}{14}$	$\frac{9}{28} + \frac{3}{14} + 0 = \frac{15}{28}$	$\frac{3}{28} + 0 + 0 = \frac{3}{28}$	$\frac{15}{28} + \frac{3}{7} + \frac{1}{28} = 1$ $\frac{5}{14} + \frac{15}{28} + \frac{3}{28} = 1$
	$g(0)$	$g(1)$	$g(2)$	

$$\begin{aligned}
\mu_X &= \sum_{x=0}^2 \sum_{y=0}^2 x f(x, y) = \sum_{x=0}^2 x \underbrace{\sum_{y=0}^2 f(x, y)}_{=g(x)} = 0 * g(0) + 1 * g(1) + 2 * g(2) \\
&= 0 * \frac{5}{14} + 1 * \frac{15}{28} + 2 * \frac{3}{28} = \frac{3}{4}
\end{aligned}$$

$$\begin{aligned}
\mu_Y &= \sum_{x=0}^2 \sum_{y=0}^2 y f(x, y) = \sum_{y=0}^2 y \underbrace{\sum_{x=0}^2 f(x, y)}_{=h(y)} = 0 * h(0) + 1 * h(1) + 2 * h(2) \\
&= 0 * \frac{15}{28} + 1 * \frac{3}{7} + 2 * \frac{1}{28} = \frac{1}{2}
\end{aligned}$$

$$\begin{aligned}
E[XY] &= \sum_{x=0}^2 \sum_{y=0}^2 xyf(x, y) = 0 * 0 * f(0, 0) + 0 * 1 * f(0, 1) + 0 * 2 * f(0, 2) + 1 * 0 * f(1, 0) + 1 * 1 * f(1, 1) \\
&\quad + 1 * 2 * f(1, 2) + 2 * 0 * f(2, 0) + 2 * 1 * f(2, 1) + 2 * 2 * f(2, 2) \\
&= 0 * 0 * \frac{3}{28} + 0 * 1 * \frac{3}{14} + 0 * 2 * \frac{1}{28} + 1 * 0 * \frac{9}{28} + 1 * 1 * \frac{3}{14} \\
&\quad + 1 * 2 * 0 + 2 * 0 * \frac{3}{28} + 2 * 1 * 0 + 2 * 2 * 0 \\
&= \frac{3}{14} \\
Cov(X, Y) &= E[XY] - \mu_X \mu_Y = \frac{3}{14} - \frac{3}{4} * \frac{1}{2} = -\frac{9}{56}
\end{aligned}$$

□

Örnek 9.15. Bir koşu maratonunda X ile yarışı tamamlayan bayan koşucuların oranını, Y ile de yarışı tamamlayan erkek koşucuların oranını göstermek üzere birleşik olasılık dağılım fonksiyonu $f(x, y)$ aşağıdaki gibi veriliyor:

$$f(x, y) = \begin{cases} 8xy & \text{if } 0 \leq y \leq x \leq 1, x, y \in \mathbb{R} \\ 0 & \text{otherwise} \end{cases}$$

Buna göre X ve Y nin kovaryansını hesaplayınız.

Çözüm:

$$\begin{aligned}
\mu_X &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xf(x, y) dx dy = \int_{-\infty}^{\infty} x \underbrace{\int_{-\infty}^{\infty} f(x, y) dy}_{=g(x)} dx \\
g(x) &= \int_{-\infty}^{\infty} f(x, y) dy = \int_0^x 8xy dy = 4x^3, 0 \leq x \leq 1 \Rightarrow \\
g(x) &= \begin{cases} 4x^3 & \text{if } 0 \leq x \leq 1, x \in \mathbb{R} \\ 0 & \text{otherwise} \end{cases} \Rightarrow \\
\mu_X &= \int_{-\infty}^{\infty} xg(x) dx = \int_0^1 x4x^3 dx = \frac{4}{5} \\
\mu_Y &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} yf(x, y) dx dy = \int_{-\infty}^{\infty} y \underbrace{\int_{-\infty}^{\infty} f(x, y) dx}_{=h(y)} dy \\
h(y) &= \int_{-\infty}^{\infty} f(x, y) dx = \int_y^1 8xy dx = 4y x^2 \Big|_{x=y}^{x=1} = 4y(1 - y^2) \\
h(y) &= \begin{cases} 4y(1 - y^2) & \text{if } 0 \leq y \leq 1, y \in \mathbb{R} \\ 0 & \text{otherwise} \end{cases} \\
\mu_Y &= \int_{-\infty}^{\infty} yh(y) dy = \int_0^1 y4y(1 - y^2) dy = \frac{4}{3}y^3 - \frac{4}{5}y^5 \Big|_{y=0}^{y=1} \\
&= \frac{8}{15} - 0 = \frac{8}{15}
\end{aligned}$$

$$\begin{aligned}
E[XY] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f(x, y) dy dx = \int_0^1 \int_0^x xy 8xy dy dx \\
&= \int_0^1 8x^2 \int_0^x y^2 dy dx = \int_0^1 8x^2 \frac{y^3}{3} \Big|_{y=0}^{y=x} dx \\
&= \frac{8}{3} \int_0^1 x^5 dx = \frac{8}{3} * \frac{1}{6} = \frac{4}{9} \\
Cov(X, Y) &= E[XY] - \mu_X \mu_Y \\
&= \frac{4}{9} - \frac{4}{5} * \frac{8}{15} = \frac{4}{225}
\end{aligned}$$

□

Kovaryans iki rastgele değişken hakkında doğal bir bağıntı versede $Cov(X, Y)$ sayısı, bu bağıntının geçerliliği hakkında net bir bilgi vermez çünkü, $Cov(X, Y)$ serbest bir ölçek değildir. $Cov(X, Y)$ değeri hem X in hemde Y nin ölçümlerine bağlıdır. Kovaryansın serbest ölçek ile verilen sabitine korelasyon katsayısı denir.

Tanım 9.16. X ve Y kesikli (sürekli) rasgele değişkenlerinin varyansı $Var(X)$, $Var(Y)$ ve bu iki değişkenin kovaryansı $Cov(X, Y)$ olmak üzere

$$Corr(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)} * \sqrt{Var(Y)}}$$

sayısına X ve Y kesikli (sürekli) rasgele değişkenlerinin korelasyon katsayısı denir. Ve korelasyon katsayısı X ve Y nin ölçümünün serbesttir.

Örnek 9.17. 2 renkli kalem, 3 mavi, 2 kırmızı ve 3 yeşil kalemin olduğu bir kutudan seçilecektir. Eğer X rasgele değişkeni mavi kalem seçimini, Y ise kırmızı kalem seçimini gösterirse X ve Y nin korelasyon katsayısını hesaplayınız.

Çözüm: Bakınız Örnek 9.14

olası (x, y) değerleri $(0, 0) \Rightarrow$ ne kırmızı ne de mavi kalem seçimi, $(0, 1) \Rightarrow$ kırmızı kalem seçilirken mavi kalem seçilemez (yani 1 kırmızı 1 adet yeşil seçilir), $(1, 0) \Rightarrow$ mavi kalem seçilirken kırmızı kalem seçilemez, $(1, 1) \Rightarrow$ 1 kırmızı 1 mavi kalem seçilir, $(2, 0) \Rightarrow$ 2 mavi kalem seçilir, $(0, 2) \Rightarrow$ 2 kırmızı kalem seçilir.

$f(x, y)$	x-mavi			$h(y)$
y-kırmızı	0	1	2	
0	$\frac{\frac{3!}{1! * 2!}}{28} = \frac{3}{28}$	$\frac{\frac{3!}{2! * 1!} * \frac{3!}{2! * 1!}}{28} = \frac{9}{28}$	$\frac{\frac{3!}{2! * 1!}}{28} = \frac{3}{28}$	$\frac{3}{28} + \frac{9}{28} + \frac{3}{28} = \frac{15}{28} = h(0)$
1	$\frac{\frac{2!}{1! * 1!} * \frac{3!}{2! * 1!}}{28} = \frac{3}{14}$	$\frac{\frac{2!}{1! * 1!} * \frac{3!}{2! * 1!}}{28} = \frac{3}{14}$	0	$\frac{3}{14} + \frac{3}{14} + 0 = \frac{3}{7} = h(1)$
2	$\frac{\frac{2!}{0! * 2!}}{28} = \frac{1}{28}$	0	0	$\frac{1}{28} + 0 + 0 = \frac{1}{28} = h(2)$
$g(x)$	$\frac{3}{28} + \frac{3}{14} + \frac{1}{28} = \frac{5}{14}$	$\frac{9}{28} + \frac{3}{14} + 0 = \frac{15}{28}$	$\frac{3}{28} + 0 + 0 = \frac{3}{28}$	$\frac{15}{28} + \frac{3}{7} + \frac{1}{28} = 1$ $\frac{5}{14} + \frac{15}{28} + \frac{3}{28} = 1$
	$g(0)$	$g(1)$	$g(2)$	

$$\begin{aligned}\mu_X &= \sum_{x=0}^2 \sum_{y=0}^2 x f(x, y) = \sum_{x=0}^2 x \underbrace{\sum_{y=0}^2 f(x, y)}_{=g(x)} = 0 * g(0) + 1 * g(1) + 2 * g(2) \\ &= 0 * \frac{5}{14} + 1 * \frac{15}{28} + 2 * \frac{3}{28} = \frac{3}{4}\end{aligned}$$

$$\begin{aligned}\mu_Y &= \sum_{x=0}^2 \sum_{y=0}^2 y f(x, y) = \sum_{y=0}^2 y \underbrace{\sum_{x=0}^2 f(x, y)}_{=h(y)} = 0 * h(0) + 1 * h(1) + 2 * h(2) \\ &= 0 * \frac{15}{28} + 1 * \frac{3}{7} + 2 * \frac{1}{28} = \frac{1}{2}\end{aligned}$$

$$\begin{aligned}E[XY] &= \sum_{x=0}^2 \sum_{y=0}^2 xy f(x, y) = 0 * 0 * f(0, 0) + 0 * 1 * f(0, 1) + 0 * 2 * f(0, 2) + 1 * 0 * f(1, 0) + 1 * 1 * f(1, 1) \\ &\quad + 1 * 2 * f(1, 2) + 2 * 0 * f(2, 0) + 2 * 1 * f(2, 1) + 2 * 2 * f(2, 2) \\ &= 0 * 0 * \frac{3}{28} + 0 * 1 * \frac{3}{14} + 0 * 2 * \frac{1}{28} + 1 * 0 * \frac{9}{28} + 1 * 1 * \frac{3}{14} \\ &\quad + 1 * 2 * 0 + 2 * 0 * \frac{3}{28} + 2 * 1 * 0 + 2 * 2 * 0 \\ &= \frac{3}{14}\end{aligned}$$

$$\mathbf{Cov}(X, Y) = \mathbf{E}[XY] - \mu_X \mu_Y = \frac{3}{14} - \frac{3}{4} * \frac{1}{2} = -\frac{9}{56}$$

$$\mathbf{Var}(X) = E[(X - \mu_X)^2] = E[X^2] - (\mu_X)^2,$$

$$\begin{aligned}E[X^2] &= \sum_{x=0}^2 \sum_{y=0}^2 x^2 f(x, y) = \sum_{x=0}^2 x^2 \underbrace{\sum_{y=0}^2 f(x, y)}_{=g(x)} = 0^2 * g(0) + 1^2 * g(1) + 2^2 * g(2) \\ &= 0^2 * \frac{5}{14} + 1^2 * \frac{15}{28} + 2^2 * \frac{3}{28} = \frac{27}{28} \Rightarrow \mathbf{Var}(X) = \mathbf{E}[X^2] - (\mu_X)^2 = \frac{27}{28} - \left(\frac{3}{4}\right)^2 = \frac{45}{112}\end{aligned}$$

$$\mathbf{Var}(Y) = E[(Y - \mu_Y)^2] = E[Y^2] - (\mu_Y)^2,$$

$$\begin{aligned}E[Y^2] &= \sum_{x=0}^2 \sum_{y=0}^2 y^2 f(x, y) = \sum_{y=0}^2 y^2 \underbrace{\sum_{x=0}^2 f(x, y)}_{=h(y)} = 0^2 * h(0) + 1^2 * h(1) + 2^2 * h(2) \\ &= 0^2 * \frac{15}{28} + 1^2 * \frac{3}{7} + 2^2 * \frac{1}{28} = \frac{4}{7} \Rightarrow \mathbf{Var}(Y) = \mathbf{E}[Y^2] - (\mu_Y)^2 = \frac{4}{7} - \left(\frac{1}{2}\right)^2 = \frac{9}{28}\end{aligned}$$

$$\mathbf{Corr}(X, Y) = \frac{\mathbf{Cov}(X, Y)}{\sqrt{\mathbf{Var}(X)} * \sqrt{\mathbf{Var}(Y)}} = \frac{-\frac{9}{56}}{\sqrt{\frac{45}{112}} * \sqrt{\frac{9}{28}}} = -\frac{1}{5}\sqrt{5} = -0.44721$$

□

Örnek 9.18. Bir koşu maratonunda X ile yarışı tamamlayan bayan koşucuların oranını, Y ile de yarışı tamamlayan erkek koşucuların oranını göstermek üzere birleşik olasılık dağılım fonksiyonu $f(x, y)$ aşağıdaki gibi veriliyor:

$$f(x, y) = \begin{cases} 8xy & \text{if } 0 \leq y \leq x \leq 1, x, y \in \mathbb{R} \\ 0 & \text{otherwise} \end{cases}$$

Buna göre X ve Y nin korelasyon katsayısını hesaplayınız.

Çözüm: Bakınız Örnek 9.15

$$\mu_X = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x, y) dx dy = \int_{-\infty}^{\infty} x \underbrace{\int_{-\infty}^{\infty} f(x, y) dy}_{=g(x)} dx$$

$$g(x) = \int_{-\infty}^{\infty} f(x, y) dy = \int_0^x 8xy dy = 4x^3, 0 \leq x \leq 1 \Rightarrow$$

$$g(x) = \begin{cases} 4x^3 & \text{if } 0 \leq x \leq 1, x \in \mathbb{R} \\ 0 & \text{otherwise} \end{cases} \Rightarrow$$

$$\mu_X = \int_{-\infty}^{\infty} x g(x) dx = \int_0^1 x 4x^3 dx = \frac{4}{5}$$

$$\mu_Y = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x, y) dx dy = \int_{-\infty}^{\infty} y \underbrace{\int_{-\infty}^{\infty} f(x, y) dx}_{=h(y)} dy$$

$$h(y) = \int_{-\infty}^{\infty} f(x, y) dx = \int_y^1 8xy dx = 4y x^2 \Big|_{x=y}^{x=1} = 4y(1 - y^2)$$

$$h(y) = \begin{cases} 4y(1 - y^2) & \text{if } 0 \leq y \leq 1, y \in \mathbb{R} \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned} \mu_Y &= \int_{-\infty}^{\infty} y h(y) dy = \int_0^1 y 4y(1 - y^2) dy = \frac{4}{3} y^3 - \frac{4}{5} y^5 \Big|_{y=0}^{y=1} \\ &= \frac{8}{15} - 0 = \frac{8}{15} \end{aligned}$$

$$\begin{aligned} E[XY] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f(x, y) dy dx = \int_0^1 \int_0^x xy 8xy dy dx \\ &= \int_0^1 8x^2 \int_0^x y^2 dy dx = \int_0^1 8x^2 \frac{y^3}{3} \Big|_{y=0}^{y=x} dx \\ &= \frac{8}{3} \int_0^1 x^5 dx = \frac{8}{3} * \frac{1}{6} = \frac{4}{9} \end{aligned}$$

$$Cov(X, Y) = E[XY] - \mu_X \mu_Y$$

$$\Rightarrow \mathbf{Cov}(X, Y) = \frac{4}{9} - \frac{4}{5} * \frac{8}{15} = \frac{4}{225}$$

$$\begin{aligned}
Var(X) &= E[(X - \mu_X)^2] = E[X^2] - (\mu_X)^2 \\
E[X^2] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^2 f(x, y) dx dy = \int_{-\infty}^{\infty} x^2 \underbrace{\int_{-\infty}^{\infty} f(x, y) dy}_{=g(x)} dx \\
&= \int_0^1 x^2 4x^3 dx = \left. \frac{2}{3} x^6 \right|_{x=0}^1 = \frac{2}{3} \Rightarrow \\
\mathbf{Var}(X) &= \frac{2}{3} - \left(\frac{4}{5}\right)^2 = \frac{2}{75}
\end{aligned}$$

$$\begin{aligned}
Var(Y) &= E[(Y - \mu_Y)^2] = E[Y^2] - (\mu_Y)^2, \\
E[Y^2] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y^2 f(x, y) dx dy = \int_{-\infty}^{\infty} y^2 \underbrace{\int_{-\infty}^{\infty} f(x, y) dx}_{=h(y)} dy \\
&= \int_0^1 y^2 4y(1 - y^2) dy = \left. y^4 - \frac{2}{3} y^6 \right|_{y=0}^{y=1} = \frac{1}{3} \Rightarrow \\
\mathbf{Var}(Y) &= \frac{1}{3} - \left(\frac{8}{15}\right)^2 = \frac{11}{225} \Rightarrow \\
Corr(X, Y) &= \frac{Cov(X, Y)}{\sqrt{Var(X)} * \sqrt{Var(Y)}} = \frac{\frac{4}{225}}{\sqrt{\frac{2}{75}} * \sqrt{\frac{11}{225}}} = \frac{2}{33} \sqrt{66} = 0.49237
\end{aligned}$$

□

10 Rasgele değişkenlerinin lineer kombinasyonlarının varyansı ve ortalaması (Means and Variances of Linear Combinations of Random Variables)

Teorem 10.1. *a ve b sabitler, X kesikli (sürekli) rasgele değişken olmak üzere*

$$E[aX + b] = aE[X] + b$$

İspat: X kesikli rasgele değişkeninin olasılık dağılım fonksiyonu $f(x)$ olmak üzere $aX + b$ kesikli rasgele değişkeninin beklenti değeri

$$\begin{aligned}
E[aX + b] &= \sum_x (ax + b) f(x) \\
&= a \sum_x x f(x) + b \underbrace{\sum_x f(x)}_{=1} \\
&= aE[X] + b
\end{aligned}$$

X sürekli rasgele değişkeninin olasılık dağılım fonksiyonu $f(x)$ olmak üzere X sürekli rasgele değişkeninin beklenti değeri

$$\begin{aligned} E[aX + b] &= \int_{-\infty}^{\infty} (ax + b) f(x) dx \\ &= a \int_{-\infty}^{\infty} x f(x) dx + b \underbrace{\int_{-\infty}^{\infty} f(x) dx}_{=1} \\ &= aE[X] + b \end{aligned}$$

Sonuç 10.2. b sabit olmak üzere $E[b] = b$.

İspat: Teorem 10.1'de $a = 0$ yazarak istenen sonucu elde ederiz.

Sonuç 10.3. a sabit olmak üzere $E[aX] = aE[X]$

İspat: Teorem 10.1'de $b = 0$ yazarak istenen sonucu elde ederiz.

Örnek 10.4. Güneşli bir cuma günü, saat 16:00-17:00 arası araba yıkamaya giden arabaların sayısı X olmak üzere olasılık dağılımı aşağıdaki gibi veriliyor:

x	4	5	6	7	8	9
$f(x)$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$

$g(X) = 5(2X - 1)$ ile arabaların odemiş oldukları ücret verilmektedir. Verilen bu zaman diliminde kazanılması beklenen para miktarını Teorem 10.1'i kullanarak bulunuz.

Çözüm: Bu örneği Örnek 8.13'te hesaplamıştık! ve $E[5(2X - 1)] = \frac{190}{3}$ sonucunu elde etmiştik.

$$\begin{aligned} E[5(2X - 1)] &= 10E[X] - 5 \\ E[X] &= 4 * \frac{1}{12} + 5 * \frac{1}{12} + 6 * \frac{1}{4} + 7 * \frac{1}{4} + 8 * \frac{1}{6} + 9 * \frac{1}{6} \\ &= \frac{41}{6} \\ E[5(2X - 1)] &= 10 * \frac{41}{6} - 5 = \frac{190}{3} \end{aligned}$$

□

Örnek 10.5. X sürekli rasgele değişkeninin olasılık dağılımı aşağıdaki gibi veriliyor:

$$f(x) = \begin{cases} \frac{x^2}{3} & \text{if } -1 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

$g(X) = 4X + 3$ rasgele değişkeninin beklentisini Teorem 10.1'i kullanarak bulunuz.

Çözüm: Bu örneği Örnek 8.14'te hesaplamıştık! ve $E[g(X)] = 8$ olarak elde edilmiştir.

$$\begin{aligned}
 E[g(X)] &= 4E[X] + 3 \\
 E[X] &= \int_{-1}^2 x \frac{x^2}{3} dx \\
 &= \frac{1}{12} x^4 \Big|_{x=-1}^{x=2} \\
 &= \frac{1}{12} * (2^4 - (-1)^4) \\
 &= \frac{5}{4} \\
 E[g(X)] &= 4E[X] + 3 \\
 &= 5 + 3 = 8
 \end{aligned}$$

□

Teorem 10.6. X kesikli (sürekli) rasgele değişken olmak üzere

$$E[g(X) \pm h(X)] = E[g(X)] \pm E[h(X)]$$

İspat: X kesikli rasgele değişkeninin olasılık dağılım fonksiyonu $f(x)$ olmak üzere

$$\begin{aligned}
 E[g(X) \pm h(X)] &= \sum_x (g(x) \pm h(x)) f(x) \\
 &= \sum_x g(x) f(x) \pm \sum_x h(x) f(x) \\
 &= E[g(X)] \pm E[h(X)]
 \end{aligned}$$

X sürekli rasgele değişkeninin olasılık dağılım fonksiyonu $f(x)$ olmak üzere

$$\begin{aligned}
 E[g(X) \pm h(X)] &= \int_{-\infty}^{\infty} (g(x) \pm h(x)) f(x) dx \\
 &= \int_{-\infty}^{\infty} g(x) f(x) dx \pm \int_{-\infty}^{\infty} h(x) f(x) dx \\
 &= E[g(X)] \pm E[h(X)]
 \end{aligned}$$

Örnek 10.7. X kesikli rasgele değişkeninin olasılık dağılım fonksiyonu

x	0	1	2	3
$f(x)$	$\frac{1}{3}$	$\frac{1}{2}$	0	$\frac{1}{6}$

olarak veriliyor. Buna göre $E[(X-1)^2]$ beklentisini hesaplayınız.

Çözüm:

$$\begin{aligned}
E[(X-1)^2] &= E[X^2 - 2X + 1] = E[X^2] - E[2X] + E[1] \\
&= E[X^2] - 2E[X] + 1 \\
E[X^2] &= \sum_{x=0}^3 x^2 f(x) = 0^2 * \frac{1}{3} + 1^2 * \frac{1}{2} + 2^2 * 0 + 3^2 * \frac{1}{6} = 2 \\
E[X] &= \sum_{x=0}^3 x f(x) = 0 * \frac{1}{3} + 1 * \frac{1}{2} + 2 * 0 + 3 * \frac{1}{6} = 1 \\
E[(X-1)^2] &= E[X^2] - 2E[X] + 1 = 2 - 2 + 1 = 1
\end{aligned}$$

□

Örnek 10.8. X sürekli rasgele değişkeninin olasılık dağılım fonksiyonu

$$f(x) = \begin{cases} 2(x-1) & \text{if } 1 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

olarak veriliyor. Buna göre $E[X^2 + X - 2]$ beklentisini hesaplayınız.**Çözüm:**

$$\begin{aligned}
E[X^2 + X - 2] &= E[X^2] + E[X] - 2 \\
E[X^2] &= \int_1^2 x^2 2(x-1) dx \\
&= \frac{1}{6} x^3 (3x-4) \Big|_{x=1}^{x=2} \\
&= \frac{8}{3} - \left(-\frac{1}{6}\right) = \frac{17}{6} \\
E[X] &= \int_1^2 x 2(x-1) dx \\
&= \frac{1}{3} x^2 (2x-3) \Big|_{x=1}^{x=2} \\
&= \frac{4}{3} - \left(-\frac{1}{3}\right) = \frac{5}{3} \\
E[X^2 + X - 2] &= \frac{17}{6} + \frac{5}{3} - 2 = \frac{5}{2}
\end{aligned}$$

□

Teorem 10.9. X, Y kesikli (sürekli) rasgele değişken olmak üzere

$$E[g(X, Y) \pm h(X, Y)] = E[g(X, Y)] \pm E[h(X, Y)].$$

İspat: X, Y kesikli rasgele değişkeninin olasılık dağılım fonksiyonu $f(x, y)$ olmak üzere

$$\begin{aligned}
 E[g(X, Y) \pm h(X, Y)] &= \sum_x \sum_y (g(x, y) \pm h(x, y)) f(x, y) \\
 &= \sum_x \sum_y g(x, y) f(x, y) \pm \sum_x \sum_y h(x, y) f(x, y) \\
 &= E[g(X, Y)] \pm E[h(X, Y)]
 \end{aligned}$$

X, Y sürekli rasgele değişkeninin olasılık dağılım fonksiyonu $f(x, y)$ olmak üzere

$$\begin{aligned}
 E[g(X, Y) \pm h(X, Y)] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (g(x, y) \pm h(x, y)) f(x, y) dx dy \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) dx dy \pm \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) f(x, y) dx dy \\
 &= E[g(X, Y)] \pm E[h(X, Y)]
 \end{aligned}$$

Sonuç 10.10. X, Y kesikli (sürekli) rasgele değişken olmak üzere

$$E[g(X) \pm h(Y)] = E[g(X)] \pm E[h(Y)].$$

İspat: Teorem 10.9'de $g(X, Y) = g(X), h(X, Y) = h(Y)$ olarak seçersek istenilen sonucu elde ederiz.

Sonuç 10.11. X, Y kesikli (sürekli) rasgele değişken olmak üzere

$$E[X \pm Y] = E[X] \pm E[Y].$$

İspat: Yukarıdaki sonuçtan $g(X) = X, h(Y) = Y$ yazarak istenilen sonucu elde ederiz.

Örnek 10.12. Hilesiz bir çift zar atışında X rastgele gelen değerlerin toplamını göstermek üzere $E(X)$ beklenen değeri nedir?

Çözüm: 1.yol: Bir zar atışında Y rastgele gelen değerini gösterirsek yukarıda $E[Y] = \frac{7}{2}$ olarak bulmuştuk. Z 2.zarın rastgele gelen değerini gösterirsek $E[Z] = \frac{7}{2}$ olur. $X = Y + Z \Rightarrow E[X] = \frac{7}{2} + \frac{7}{2} = 7$ olarak elde ederiz.

2.yol:

$$\begin{aligned}
f(2) &= P((1,1)) = \frac{1}{6} * \frac{1}{6} = \frac{1}{36} \\
f(3) &= P(\{(1,2), (2,1)\}) = \frac{1}{36} + \frac{1}{36} = \frac{2}{36} \\
f(4) &= P(\{(1,3), (2,2), (3,1)\}) = \frac{3}{36} \\
f(5) &= P(\{(1,4), (2,3), (3,2), (4,1)\}) = \frac{4}{36} \\
f(6) &= P(\{(1,5), (2,4), (3,3), (4,2), (5,1)\}) = \frac{5}{36} \\
f(7) &= P(\{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}) = \frac{6}{36} \\
f(8) &= P(\{(2,6), (3,5), (4,4), (5,3), (6,2)\}) = \frac{5}{36} \\
f(9) &= P(\{(3,6), (4,5), (5,4), (6,3)\}) = \frac{4}{36} \\
f(10) &= P(\{(4,6), (5,5), (6,4)\}) = \frac{3}{36} \\
f(11) &= P(\{(5,6), (6,5)\}) = \frac{2}{36} \\
f(12) &= P((6,6)) = \frac{1}{36} \Rightarrow \\
E[X] &= 2 * \frac{1}{36} + 3 * \frac{2}{36} + 4 * \frac{3}{36} + 5 * \frac{4}{36} + 6 * \frac{5}{36} + \\
&\quad 7 * \frac{6}{36} + 8 * \frac{5}{36} + 9 * \frac{4}{36} + 10 * \frac{3}{36} + 11 * \frac{2}{36} + 12 * \frac{1}{36} \\
&= 7
\end{aligned}$$

□

Teorem 10.13. *X, Y kesikli (sürekli) rasgele değişkenleri bağımsız ise*

$$E[XY] = E[X] E[Y]$$

İspat: X, Y kesikli (sürekli) rasgele değişkenleri bağımsız ise, X, Y kesikli rasgele değişkeninin olasılık dağılım fonksiyonu $f(x, y)$ olmak üzere

$$f(x, y) = g(x) h(y)$$

olarak yazılır. Burada $g(x), h(y)$ marjinal olasılık dağılım fonksiyonlarıdır.

$$\begin{aligned}
 E[XY] &= \sum_x \sum_y xyf(x, y) \\
 &= \sum_x \sum_y xyg(x)h(y), g(x) = \sum_y f(x, y), h(y) = \sum_x f(x, y) \\
 &= \left(\sum_x xg(x) \right) \left(\sum_y yh(y) \right) \\
 &= \left(\sum_x x \sum_y f(x, y) \right) \left(\sum_y y \sum_x f(x, y) \right) \\
 &= \sum_x \sum_y xf(x, y) + \sum_x \sum_y yf(x, y) \\
 &= E[X]E[Y]
 \end{aligned}$$

X, Y sürekli rasgele değişkeninin olasılık dağılım fonksiyonu $f(x, y)$ olmak üzere

$$\begin{aligned}
 E[XY] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x, y) dx dy \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyg(x)h(y) dx dy \\
 &= \left(\int_{-\infty}^{\infty} xg(x) dx \right) \left(\int_{-\infty}^{\infty} yh(y) dy \right) \\
 g(x) &= \int_{-\infty}^{\infty} f(x, y) dy, h(y) = \int_{-\infty}^{\infty} f(x, y) dx \Rightarrow \\
 E[XY] &= \left(\int_{-\infty}^{\infty} xg(x) dx \right) \left(\int_{-\infty}^{\infty} yh(y) dy \right) \\
 &= \left(\int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f(x, y) dy dx \right) \left(\int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} f(x, y) dx dy \right) \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xf(x, y) dx dy + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} yf(x, y) dx dy \\
 &= E[X]E[Y]
 \end{aligned}$$

Sonuç 10.14. X, Y kesikli (sürekli) rasgele değişkenleri bağımsız ise $Cov(X, Y) = 0$

İspat: Teorem 9.13'den

$$Cov(X, Y) = E[XY] - \mu_X \mu_Y$$

Ayrıca Teorem 10.13'den X, Y kesikli (sürekli) rasgele değişkenleri bağımsız ise

$$E[XY] = E[X]E[Y] = \mu_X \mu_Y$$

ifadesini yukarıdaki kovaryans formülünde yazarsak istenilen sonucu elde ederiz.

Teorem 10.15. X ve Y kesikli (sürekli) rasgele değişkenleri, a, b, c sabitler olmak üzere

$$Var(aX + bY + c) = a^2 Var(X) + b^2 Var(Y) + 2abCov(X, Y)$$

İspat:

$$\begin{aligned}
Var(aX + bY + c) &= E \left[((aX + bY + c) - \mu_{aX+bY+c})^2 \right] \\
\mu_{aX+bY+c} &= E[aX + bY + c] \\
&= aE[X] + bE[Y] + c \\
&= a\mu_X + b\mu_Y + c \Rightarrow \\
(aX + bY + c) - \mu_{aX+bY+c} &= (aX + bY + c) - (a\mu_X + b\mu_Y + c) \\
&= a(X - \mu_X) + b(Y - \mu_Y) \Rightarrow \\
Var(aX + bY + c) &= E \left[((aX + bY + c) - \mu_{aX+bY+c})^2 \right] \\
&= E \left[(a(X - \mu_X) + b(Y - \mu_Y))^2 \right] \\
&= E \left[a^2(X - \mu_X)^2 + b^2(Y - \mu_Y)^2 + 2ab(X - \mu_X)(Y - \mu_Y) \right] \\
&= E \left[a^2(X - \mu_X)^2 \right] + E \left[b^2(Y - \mu_Y)^2 \right] + E \left[2ab(X - \mu_X)(Y - \mu_Y) \right] \\
&= a^2 E \left[(X - \mu_X)^2 \right] + b^2 E \left[(Y - \mu_Y)^2 \right] + 2ab E \left[(X - \mu_X)(Y - \mu_Y) \right] \\
&= a^2 Var(X) + b^2 Var(Y) + 2ab Cov(X, Y)
\end{aligned}$$

Sonuç 10.16. *X ve Y kesikli (süreklı) rasgele değişkenleri olmak üzere*

$$Var(aX + c) = a^2 Var(X)$$

İspat: Teorem 10.15'den $b = 0$ seçersek istenen sonucu elde ederiz.

Sonuç 10.17. *X ve Y kesikli (süreklı) rasgele değişkenleri olmak üzere*

$$Var(X + c) = Var(X)$$

İspat: Teorem 10.15'den $a = 1, b = 0$ seçersek istenen sonucu elde ederiz.

Sonuç 10.18. *X ve Y kesikli (süreklı) rasgele değişkenleri olmak üzere*

$$Var(aX) = a^2 Var(X)$$

İspat: Teorem 10.15'den $c = 0, b = 0$ seçersek istenen sonucu elde ederiz.

Örnek 10.19. *X rasgele değişkeninin varyansı 0.5 olsun. Buna göre $2X, X/2+3$ değişkenlerinin varyansını bulunuz.*

Çözüm:

$$\begin{aligned}
Var(2X) &= 4Var(X) = 4 * 0.5 = 2 \\
Var\left(\frac{X}{2} + 3\right) &= \frac{1}{4}Var(X) = \frac{1}{4} * 0.5 = 0.125
\end{aligned}$$

□

Sonuç 10.20. X ve Y kesikli (sürekli) rasgele değişkenleri bağımsız olmak üzere

$$\text{Var}(aX + bY) = a^2 \text{Var}(X) + b^2 \text{Var}(Y)$$

İspat: Teorem 10.15'den $c = 0$ seçersek, bağımsız değişkenler için Sonuç 10.14'den $\text{Cov}(X, Y) = 0$ olduğundan, istenen sonucu elde ederiz.

Sonuç 10.21. X ve Y kesikli (sürekli) rasgele değişkenleri bağımsız olmak üzere

$$\text{Var}(aX - bY) = a^2 \text{Var}(X) + b^2 \text{Var}(Y)$$

İspat: Teorem 10.15'den $c = 0$ seçersek, bağımsız değişkenler için Sonuç 10.14'den $\text{Cov}(X, Y) = 0$ olduğundan, istenen sonucu elde ederiz.

11 Moment ve Moment üreten fonksiyonlar (Moments and Moment-Generating Functions)

Moment kabaca konuşacak olursak, bir noktalar kümesinin nasıl bir şekil gösterdiğini anlatan sayısal bir ölçüttür ve X rasgele değişkeninin merkez etrafında (sıfır civarında) n . momenti $g(X) = X^n$ fonksiyonun beklenen değeridir.

Tanım 11.1. X rasgele değişkeninin n . momenti

$$m_n = E[X^n] = \begin{cases} \sum_x x^n f(x) & , \quad X - \text{kesikli rasgele değişkeni} \\ \int_{-\infty}^{\infty} x^n f(x) dx & , \quad X - \text{sürekli rasgele değişkeni} \end{cases}$$

olarak tanımlanır. Buna göre X rasgele değişkeninin 1.momenti beklenen değerine eşittir.

Tanım 11.2. X rasgele değişkeni ve beklenti değeri $\mu = E[X]$ olmak üzere n . merkezi momenti

$$E[(X - \mu)^n] = \begin{cases} \sum_x (x - \mu)^n f(x) & , \quad X - \text{kesikli rasgele değişkeni} \\ \int_{-\infty}^{\infty} (x - \mu)^n f(x) dx & , \quad X - \text{sürekli rasgele değişkeni} \end{cases}$$

olarak tanımlanır.

Örnek 11.3. X sürekli rasgele değişkeninin olasılık dağılım fonksiyonu $f(x) = ce^{-(x-3)^2}$ olarak tanımlandığına göre n -tek sayısı için n . merkezi momentini bulunuz.

Çözüm: Öncelikle c sabitini belirleyelim:

$$\begin{aligned}\int_{-\infty}^{\infty} ce^{-(x-3)^2} dx &= 1 \Rightarrow 1 = c \int_{-\infty}^{\infty} ce^{-u^2} du = c\sqrt{\pi} \Rightarrow c = \frac{1}{\sqrt{\pi}} \\ \mu &= E[X] = \int_{-\infty}^{\infty} xf(x) dx = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} xe^{-(x-3)^2} dx \\ &\stackrel{(x-3)=u}{=} \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} (u+3) e^{-u^2} du = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} ue^{-u^2} du + \frac{3}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-u^2} du \\ &= -\frac{1}{2\sqrt{\pi}} e^{-u^2} \Big|_{u=-\infty}^{u=\infty} + 3 = 3 \\ E[(X-3)^n] &= \int_{-\infty}^{\infty} (x-3)^n f(x) dx = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} (x-3) e^{-(x-3)^2} dx \\ &\stackrel{(x-3)=u}{=} \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} u^n e^{-u^2} du = 0, n - \text{tek}\end{aligned}$$

n -tek sayısı için $u^n e^{-u^2}$ fonksiyonu tek fonksiyondur ve tek fonksiyonunun simetrik aralıkta integrali sıfırdır.

□

Tanım 11.4. X rasgele değişkeninin moment üreten fonksiyonu $E[e^{tX}]$ ile tanımlanıp $M_X(t)$ ile gösterilir:

$$M_X(t) = E[e^{tX}] = \begin{cases} \sum_x e^{tx} f(x) & , \quad X - \text{kesikli rasgele değişkeni} \\ \int_{-\infty}^{\infty} e^{tx} f(x) dx & , \quad X - \text{sürekli rasgele değişkeni} \end{cases}$$

Teorem 11.5.

$$\left. \frac{dM_X(t)}{dt} \right|_{t=0} = m_n$$

İspat:

$$\begin{aligned}\frac{dM_X(t)}{dt} &= \begin{cases} \sum_x x e^{tx} f(x) & , \quad X - \text{kesikli rasgele değişkeni} \\ \int_{-\infty}^{\infty} x e^{tx} f(x) dx & , \quad X - \text{sürekli rasgele değişkeni} \end{cases} \Rightarrow \\ \left. \frac{d^n M_X(t)}{dt^n} \right|_{t=0} &= \begin{cases} \sum_x x^n f(x) & , \quad X - \text{kesikli rasgele değişkeni} \\ \int_{-\infty}^{\infty} x^n f(x) dx & , \quad X - \text{sürekli rasgele değişkeni} \end{cases} = m_n\end{aligned}$$

Örnek 11.6. X sürekli rasgele değişkeninin olasılık dağılım fonksiyonu

$$f(x) = \begin{cases} e^{-x} & , \quad x > 0 \\ 0 & , \quad x \leq 0 \end{cases}$$

olarak tanımlandığına göre X rasgele değişkeninin moment üreten fonksiyonunu bulunuz.

Çözüm:

$$\begin{aligned}M_X(t) &= E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f(x) dx = \int_0^{\infty} e^{tx} e^{-x} dx = \int_0^{\infty} e^{(t-1)x} dx = \frac{1}{t-1} e^{(t-1)x} \Big|_{x=0}^{x=\infty} = \frac{1}{1-t}, t-1 < 0 \Rightarrow \\ M_X(t) &= \frac{1}{1-t}, t < 1\end{aligned}$$

□

Teorem 11.7. (i)

$$\forall a \in \mathbb{R}, M_{X+a}(t) = e^{at} M_X(t)$$

(ii)

$$\forall a \in \mathbb{R}, M_{aX}(t) = M_X(at)$$

İspat: (i)

$$\begin{aligned} M_{X+a}(t) &= E[e^{t(X+a)}] = \begin{cases} \sum_x e^{t(x+a)} f(x) & , \quad X - \text{kesikli rasgele deęişkeni} \\ \int_{-\infty}^{\infty} e^{t(x+a)} f(x) dx & , \quad X - \text{sürekli rasgele deęişkeni} \end{cases} \\ &= e^{at} \begin{cases} \sum_x e^{tx} f(x) & , \quad X - \text{kesikli rasgele deęişkeni} \\ \int_{-\infty}^{\infty} e^{tx} f(x) dx & , \quad X - \text{sürekli rasgele deęişkeni} \end{cases} \\ &= e^{at} M_X(t) \end{aligned}$$

(ii)

$$\begin{aligned} M_{aX}(t) &= E[e^{t(aX)}] = \begin{cases} \sum_x e^{t(ax)} f(x) & , \quad X - \text{kesikli rasgele deęişkeni} \\ \int_{-\infty}^{\infty} e^{t(ax)} f(x) dx & , \quad X - \text{sürekli rasgele deęişkeni} \end{cases} \\ &= \begin{cases} \sum_x e^{(at)x} f(x) & , \quad X - \text{kesikli rasgele deęişkeni} \\ \int_{-\infty}^{\infty} e^{(at)x} f(x) dx & , \quad X - \text{sürekli rasgele deęişkeni} \end{cases} \\ &= M_X(at) \end{aligned}$$

Örnek 11.8. X sürekli rasgele deęişkeninin olasılık dağılım fonksiyonu

$$f(x) = \begin{cases} e^{-x} & , \quad x > 0 \\ 0 & , \quad x \leq 0 \end{cases}$$

olarak tanımlandığına göre $3 - 2X$ rasgele deęişkeninin moment üreten fonksiyonunu bulunuz.

Çözüm:

$$M_{3-2X}(t) = e^{3t} M_{-2X}(t) = e^{3t} M_X(-2t)$$

Örnek 11.6'da X rasgele deęişkeninin moment üreten fonksiyonunu $M_X(t) = \frac{1}{1-t}, t < 1$ olarak elde etmiştik. Buna göre $M_X(-2t) = \frac{1}{1+2t}, -2t < 1 \Rightarrow M_X(-2t) = \frac{1}{1+2t}, t > -\frac{1}{2}$,

$$M_{3-2X}(t) = e^{3t} M_X(-2t) = \frac{e^{3t}}{1+2t}, t > -\frac{1}{2}$$

olarak elde ederiz. □

12 Chebysev Eşitsizliği ve Büyük Sayılar Yasası (CHEBYSHEV'S INEQUALITY AND THE LAW OF LARGE NUMBERS)

Bir A olayı p olasılığı ile gerçekleşiyorsa A 'nın ortalama olarak gerçekleşme sayısı, bağımsız denemeler arttıkça p olasılığına yakınsar. Bu kavram büyük sayılar yasası ile ifade edilirken bu yasanın kanıtında Chebysev eşitsizliği kullanılmaktadır.

Önerme 12.1. (Markov Eşitsizliği) Eğer X rasgele değişkeni için $X \geq 0$ ise $\forall a > 0$ için

$$P(X \geq a) \leq \frac{E[X]}{a} \quad (12.1)$$

İspat:

$$\begin{aligned} E[X] &= \begin{cases} \sum_x x f(x) & , \quad X - \text{kesikli rasgele değişkeni} \\ \int_{-\infty}^{\infty} x f(x) dx & , \quad X - \text{sürekli rasgele değişkeni} \end{cases} \\ &= \begin{cases} \sum_{0 \leq x < a} x f(x) + \sum_{x \geq a} x f(x) & , \quad X - \text{kesikli rasgele değişkeni} \\ \int_0^a x f(x) dx + \int_a^{\infty} x f(x) dx & , \quad X - \text{sürekli rasgele değişkeni} \end{cases} \\ &\geq \begin{cases} \sum_{x \geq a} x f(x) & , \quad X - \text{kesikli rasgele değişkeni} \\ \int_a^{\infty} x f(x) dx & , \quad X - \text{sürekli rasgele değişkeni} \end{cases} , (x f(x) \geq 0 - \text{oldugundan}) \\ &\geq \begin{cases} \sum_{x \geq a} a f(x) & , \quad X - \text{kesikli rasgele değişkeni} \\ \int_a^{\infty} a f(x) dx & , \quad X - \text{sürekli rasgele değişkeni} \end{cases} \\ &= \begin{cases} a \sum_{x \geq a} f(x) & , \quad X - \text{kesikli rasgele değişkeni} \\ a \int_a^{\infty} f(x) dx & , \quad X - \text{sürekli rasgele değişkeni} \end{cases} \\ &= a P(X \geq a) \Rightarrow P(X \geq a) \leq \frac{E[X]}{a} \end{aligned}$$

Teorem 12.2. (Chebysev Eşitsizliği) X rasgele değişkeni için beklenti değeri $\mu = E[X]$ olmak üzere $\forall a > 0$ için

$$P(|X - \mu| \geq a) \leq \frac{Var(X)}{a^2}$$

İspat: $|X - \mu|^2 \geq 0$ olduğundan (12.1) Markov eşitsizliğini kullanırsak

$$P(|X - \mu|^2 \geq a^2) \leq \frac{E[|X - \mu|^2]}{a^2} = \frac{Var(X)}{a^2}$$

Ayrıca

$$|X - \mu|^2 \geq a^2 \Leftrightarrow |X - \mu| \geq a$$

olduğundan

$$\begin{aligned} P(|X - \mu|^2 \geq a^2) &= P(|X - \mu| \geq a) \Rightarrow \\ P(|X - \mu| \geq a) &= P(|X - \mu|^2 \geq a^2) \leq \frac{Var(X)}{a^2} \end{aligned}$$

sonucunu elde ederiz.

Örnek 12.3. $\mu_x = 7$ ve $\sigma_x = 2$ alarak $P(3 < X < 11)$ olasılığının en küçük değerini bulunuz.

Çözüm:

$$\begin{aligned} P(3 < X < 11) &= P(|X - 7| < 4) = 1 - P(|X - 7| \geq 4) \\ P(|X - 7| \geq 4) &\leq \frac{Var(X)}{4^2} = \frac{4}{4^2} = \frac{1}{4} \Rightarrow -P(|X - 7| \geq 4) \geq -\frac{1}{4} \Rightarrow \\ 1 - P(|X - 7| \geq 4) &\geq 1 - \frac{1}{4} = \frac{3}{4} \end{aligned}$$

□

Örnek 12.4. Bir fabrikada bir hafta boyunca üretilen ürünlerin sayısının ortalama değeri (beklenen değeri) 50 olarak veriliyor.

- (i) Buna göre haftalık olarak üretilen ürünlerin sayısının 75'i aşması olasılığı hakkında ne söylenebilir?
(ii) Haftalık üretimlerin varyansı 25 olduğuna göre haftalık üretilen ürünlerin sayısının 40 ile 60 arasında kalması olasılığı hakkında ne söylenebilir?

Çözüm: $X \rightarrow$ haftalık olarak üretilen ürünler

(i) $P(X \geq 75) = ?$

$$P(X \geq 75) \leq \frac{E[X]}{75} = \frac{50}{75} = \frac{2}{3} = \%0.66667$$

(ii) $P(|X - 50| \geq 10) = ?$

$$P(|X - 50| \geq 10) \leq \frac{Var(X)}{10^2} = \frac{25}{100} = \frac{1}{4} = \%25$$

□

Teorem 12.5. (Büyük sayılar Yasası - The Law of Large Numbers) Ortalaması (beklenen değeri) μ ve varyansı σ^2 olan X_1, X_2, \dots bağımsız rasgele değişkenleri olmak üzere $\forall \varepsilon > 0$ için

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{X_1 + X_2 + \dots + X_n}{n} - \mu\right| \geq \varepsilon\right) = 0$$

veya

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{X_1 + X_2 + \dots + X_n}{n} - \mu\right| < \varepsilon\right) = 1$$

İspat: Sonuç 10.14'den

$$\begin{aligned} E\left[\frac{X_1 + X_2 + \dots + X_n}{n}\right] &= \frac{E[X_1] + E[X_2] + \dots + E[X_n]}{n} = \frac{\mu + \mu + \dots + \mu}{n} = \frac{n\mu}{n} = \mu \\ Var\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right) &= \frac{1}{n^2} Var(X_1 + X_2 + \dots + X_n) = \frac{Var(X_1) + Var(X_2) + \dots + Var(X_n)}{n^2} \\ &= \frac{\sigma^2 + \sigma^2 + \dots + \sigma^2}{n^2} = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n} \end{aligned}$$

Chebysev eşitsizliğinden

$$\begin{aligned}
 0 &\leq P\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| \geq \varepsilon\right) \leq \frac{\text{Var}\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right)}{\varepsilon^2} = \frac{\frac{\sigma^2}{n}}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \Rightarrow \\
 0 &\leq \lim_{n \rightarrow \infty} P\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| \geq \varepsilon\right) \leq \lim_{n \rightarrow \infty} \frac{\sigma^2}{n\varepsilon^2} = 0 \Rightarrow \\
 0 &= \lim_{n \rightarrow \infty} P\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| \geq \varepsilon\right) \\
 P\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| < \varepsilon\right) &= 1 - P\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| \geq \varepsilon\right) \Rightarrow \\
 \lim_{n \rightarrow \infty} P\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| < \varepsilon\right) &= 1 - \lim_{n \rightarrow \infty} P\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| \geq \varepsilon\right) \\
 &= 1
 \end{aligned}$$

Büyük sayılar yasasını şu şekilde de yorumlayabiliriz: Örneğin hileli olmayan bir para atıyorsunuz. Varsayalım ki 100 kez attınız kaç defa yazı kaç defa tura gelmesini beklersiniz. Genel beklenti aslında yarı yarıya şeklindedir. Veya 48 tura ike 52 yazı şeklinde de olabilir kısaca yarı yarıya veya bu aralığa yakın bir değer beklenirken akla şu soru da elbette gelebilir. Ya 90 yazı ve 10 tura gelirse? Veya daha kötüsü ya hepsi birde yazı gelirse? Ki böyle bir durumda genelde çok iyi düşüncelere varmayız ve genelde de paranın hileli olduğunu veya zar atanın hile yaptığını öne süreriz. Bu ise büyük yasalar kanunu ile açıklanabilir. Bir olayı ne kadar çok denerseniz teorik olarak beklenen ortalama değere o kadar çok yakın bir sonuç elde ederiz.

Örnek 12.6. X_1, X_2, \dots, X_n olaylarını

$$X_i = \begin{cases} 1 & , \quad \text{sonuç başarılı ise} \\ 0 & , \quad \text{sonuç başarısız ise} \end{cases}$$

olarak verelim (Örnek olarak para atma deneyinde eğer yazı gelirse 1 tura gelirse 0 gibi ve X_i bu durumda i . atış olarak düşünebiliriz.) Ve olayın başarılı olması ihtimali %30, başarısız olması ihtimali ise %70 dir. Buna göre

$$E[X_i] = 1 * 0.3 = 0.3$$

yani başarılı olmasının beklentisi %30 dur.

$$\begin{aligned}
 \text{Var}(X_i) &= E[(X_i - \mu)^2] = (1 - 0.3) * 0.3 = 0.21 \\
 E\left[\frac{X_1 + X_2 + \cdots + X_n}{n}\right] &= \frac{E[X_1] + E[X_2] + \cdots + E[X_n]}{n} = \frac{n \cdot 0.3}{n} = 0.3 \\
 \text{Var}\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right) &= \frac{\text{Var}(X_1) + \text{Var}(X_2) + \cdots + \text{Var}(X_n)}{n^2} \\
 &= \frac{n \cdot 0.21}{n^2} = \frac{0.21}{n}
 \end{aligned}$$

Chebysev eşitsizliğinden $\varepsilon = 0.1$

$$0 \leq P\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - 0.3\right| \geq 0.1\right) \leq \frac{\text{Var}\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right)}{\varepsilon^2} = \frac{\frac{0.21}{n}}{0.01} = \frac{21}{n}$$

Eğer $n = 100$ olarak alırsak (yani 100 kez para attığımızda düşünebiliriz!)

$$\begin{aligned} 0 &\leq P\left(\left|\frac{X_1 + X_2 + \cdots + X_{100}}{100} - 0.3\right| \geq 0.1\right) \leq \frac{21}{100} = 0.21 \Rightarrow \\ P\left(\left|\frac{X_1 + X_2 + \cdots + X_{100}}{100} - 0.3\right| < 0.1\right) &\geq 0.79 \Rightarrow \\ P\left(0.2 < \frac{X_1 + X_2 + \cdots + X_{100}}{100} < 0.4\right) &\geq 0.79 \end{aligned}$$

(Yani 100 denemeden sonra yazı gelme olasılığı %79 dan fazladır). Veya $n = 1000$ olarak alırsak

$$\begin{aligned} 0 &\leq P\left(\left|\frac{X_1 + X_2 + \cdots + X_{1000}}{1000} - 0.3\right| \geq 0.1\right) \leq \frac{21}{1000} = 0.021 \Rightarrow \\ P\left(\left|\frac{X_1 + X_2 + \cdots + X_{1000}}{1000} - 0.3\right| < 0.1\right) &\geq 0.979 \Rightarrow \\ P\left(0.2 < \frac{X_1 + X_2 + \cdots + X_{1000}}{1000} < 0.4\right) &\geq 0.979 \end{aligned}$$

(Yani 1000 denemeden sonra yazı gelme olasılığı %97 dan fazladır.)

13 Önemli Kesikli Dağılım Fonksiyonları

13.1 Binom Dağılımı (Binomial Distributions)

Tanım 13.1. Bir deneyde başarı sağlanması durumunda X raslantı değişkeni için $X = 1$ veya başarısızlık sağlanıyorsa $X = 0$ değerini alıyorsa bu raslantı değişkenine Bernoulli¹ raslantı değişkeni denir. Buna göre olasılık dağılım fonksiyonu

$$\begin{aligned} f(0) &= P(X = 0) = 1 - p, \quad p \in (0, 1) \\ f(1) &= P(X = 1) = p \end{aligned}$$

Ve beklenti ise

$$E[X] = 1 * f(1) + 0 * f(0) = p$$

Örnek 13.2. Örneğin bir para havaya atıldığında yazı, tura gelmesi, herhangi bir seçimi kazanma, kayıp etme, ürünün defolu olması ya da, defolu olmaması, beğenmek ya da beğenmemek. Buna göre bunların hepsi birer Bernoulli raslantı değişkeni olarak adlandırılabilir.

Tanım 13.3. X Bernoulli raslantı değişkeninin başarılı olması olasılığı p , başarısız olması durumu ise $1 - p$ olmak üzere X raslantı değişkeninin, n lineer bağımsız denemesinin binom dağılım fonksiyonu

$$P(X = x) = b(x; n, p) = C_{n,x} p^x (1 - p)^{n-x}, \quad x = 0, 1, 2, \dots, n$$

ile tanımlanır.

Örnek 13.4. Bir fabrikada üretilen ürünlerden arızalı olanların (D) seçilmesi durumunu 1 ile arızalı olmayanlarında (N) seçilmemesini 0 ile gösterirsek, arızalı gelmesi olasılığı %25 olduğuna göre 3 ürün seçilmesi

¹James Bernoulli - İsviçreli Matematikçi - 27 Aralık 1654/6 Ocak 1655 – 16 Ağustos 1705

durumunda binom dağılım fonksiyonunu belirleyiniz.

S	DDD	DDN	DND	NDD	DNN	NDN	NND	NNN
X	3	2	2	2	1	1	1	0

$$f(0) = P(X=0) = P(NNN) = \frac{3}{4} * \frac{3}{4} * \frac{3}{4} = \frac{27}{64}$$

$$f(1) = P(X=1) = P(NND) + P(NDN) + P(DNN) = 3 * \frac{3}{4} * \frac{3}{4} * \frac{1}{4} = \frac{27}{64}$$

$$f(2) = P(X=2) = P(NDD) + P(DND) + P(DDN) = 3 * \frac{3}{4} * \frac{1}{4} * \frac{1}{4} = \frac{9}{64}$$

$$f(3) = P(X=3) = P(DDD) = \frac{1}{4} * \frac{1}{4} * \frac{1}{4} = \frac{1}{64}$$

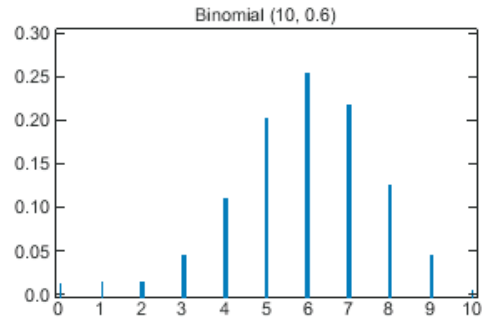
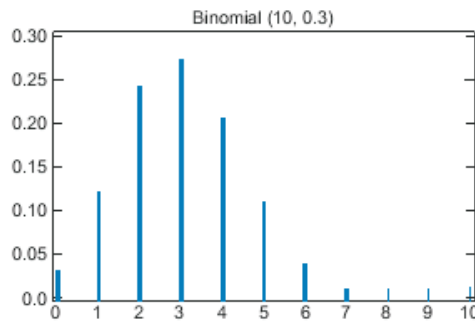
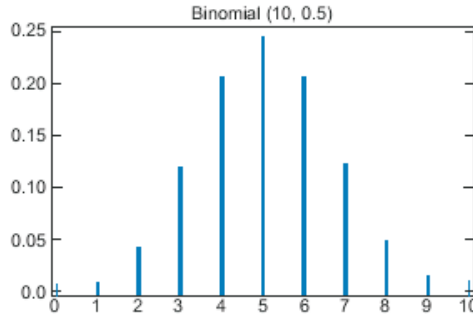
$$n=3, p=\frac{1}{4} \Rightarrow b\left(x; 3, \frac{1}{4}\right) = C_{3,x} \left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right)^{3-x}, x=0, 1, 2, 3 \Rightarrow$$

$$P(X=0) = b\left(0; 3, \frac{1}{4}\right) = C_{3,0} \left(\frac{1}{4}\right)^0 \left(\frac{3}{4}\right)^3 = \frac{27}{64}$$

$$P(X=1) = b\left(1; 3, \frac{1}{4}\right) = C_{3,1} \left(\frac{1}{4}\right)^1 \left(\frac{3}{4}\right)^2 = \frac{27}{64}$$

$$P(X=2) = b\left(2; 3, \frac{1}{4}\right) = C_{3,2} \left(\frac{1}{4}\right)^2 \left(\frac{3}{4}\right)^1 = \frac{9}{64}$$

$$P(X=3) = b\left(3; 3, \frac{1}{4}\right) = C_{3,3} \left(\frac{1}{4}\right)^3 \left(\frac{3}{4}\right)^0 = \frac{1}{64}$$



Önerme 13.5.

$$\sum_{x=0}^n b(x; n, p) = 1$$

İspat:

$$1 = 1^n = (p + (1 - p))^n = \sum_{x=0}^n C_{n,x} p^x (1 - p)^{n-x} = \sum_{x=0}^n b(x; n, p)$$

Tanım 13.6.

$$P(X \leq d) = \sum_{x=0}^d b(x; n, p)$$

toplamına Binom toplamı denir. Buna göre

$$\begin{aligned} P(X > d) &= 1 - P(X \leq d) = 1 - \sum_{x=0}^d b(x; n, p) \\ P(c \leq X \leq d) &= P(X \leq d) - P(X \leq c-1) \\ &= \sum_{x=0}^d b(x; n, p) - \sum_{x=0}^{c-1} b(x; n, p) \\ &= \sum_{x=c}^d b(x; n, p) \end{aligned}$$

Örnek 13.7. Bir A bölgesinde yaşayan insanların ölümcül bir hastalığa yakalanarak ölme riski %60 olup bu bölgede rasgele seçilen 15 kişiden

(i) en az 10 kişinin bu hastalıktan dolayı ölmemesi olasılığı nedir?

(ii) 3 ile kişinin bu hastalıktan ölmemesi olasılığı nedir?

(iii) 5 kişinin bu hastalıktan ölme olasılığı nedir?

Çözüm: X raslantı değişkeni ile hastalıktan dolayı ölmeyen insanları gösterisek:

(i)

$$\begin{aligned} P(X \geq 10) &= \sum_{x=10}^{15} b(x, 15, 0.4) = C_{15,10} 0.4^{10} 0.6^5 + C_{15,11} 0.4^{11} 0.6^4 + C_{15,12} 0.4^{12} 0.6^3 + C_{15,13} 0.4^{13} 0.6^2 \\ &\quad + C_{15,14} 0.4^{14} 0.6^1 + C_{15,15} 0.4^{15} 0.6^0 = 0.0338 \end{aligned}$$

(ii)

$$\begin{aligned} P(3 \leq X \leq 8) &= \sum_{x=3}^8 b(x, 15, 0.4) \\ &= C_{15,3} 0.4^3 0.6^{12} + C_{15,4} 0.4^4 0.6^{11} + C_{15,5} 0.4^5 0.6^{10} + C_{15,6} 0.4^6 0.6^9 + C_{15,7} 0.4^7 0.6^8 + C_{15,8} 0.4^8 0.6^7 \\ &= \frac{15!}{3!12!} 0.4^3 0.6^{12} + \frac{15!}{4!11!} 0.4^4 0.6^{11} + \frac{15!}{5!10!} 0.4^5 0.6^{10} + \frac{15!}{6!9!} 0.4^6 0.6^9 + \frac{15!}{7!8!} 0.4^7 0.6^8 + \frac{15!}{8!7!} 0.4^8 0.6^7 \\ &= 0.87784 \end{aligned}$$

(iii)

$$P(X = 5) = b(5, 15, 0.4) = C_{15,5} 0.4^5 0.6^{10} = \frac{15!}{5!10!} 0.4^5 0.6^{10} = 0.18594$$

□

Örnek 13.8. Bir mağazada elektronik ürünlerin arızalı olması ihtimali %3 olduğu bilindiğine göre kargo ile gönderilmek üzere rasgele seçilen 20 ürünün en az birinin arızalı olması olasılığı nedir? Ayda 10 defa kargo ile gönderilmek üzere 20 ürün seçilmesi durumunda 3 gönderimde en az bir arızalı ürün seçilmesi olasılığı nedir?

Çözüm: X arızalı ürün seçimini göstermek üzere

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) = 1 - b(0, 20, 0.03) \\ &= 1 - (0.97)^{20} = 0.45621 \end{aligned}$$

Her bir kargo ile gönderimdeki en az bir arızalı ürün seçilmesi olasılığı 0.45621 olduğuna göre 3 gönderimde bunun bulunması, Y ile kargo seçimini göstermek üzere

$$\begin{aligned} P(Y = 3) &= b(3, 10, 0.45621) \\ &= C_{10,3} 0.45621^3 (1 - 0.45621)^7 \\ &= \frac{10!}{3!7!} 0.45621^3 (1 - 0.45621)^7 \\ &= 0.16021 \end{aligned}$$

□

Teorem 13.9. X Bernoulli raslantı değişkeninin başarılı olması olasılığı p ve olasılık dağılım fonksiyonu $b(x, n, p)$ Binom dağılımı olsun. Buna göre

$$\begin{aligned} E[X] &= np \\ Var(X) &= np(1 - p) \end{aligned}$$

İspat: i. denemeyi I_i ile gösterelim. I_i de başarılı olan sonuçları 1 veya başarısızlıkları da 0 ile ve bunların olasılıklarını da p ve 1-p göstermek üzere n bağımsız denemenin sonucundaki X raslantı değişkeninin,

$$X = I_1 + I_2 + \dots + I_n$$

olarak yazabiliriz. Ayrıca,

$$\begin{aligned} E[I_i] &= 1 * p + 0 * (1 - p) = p \Rightarrow \\ E[X] &= E[I_1 + I_2 + \dots + I_n] \\ &= E[I_1] + E[I_2] + \dots + E[I_n] \\ &= np \end{aligned}$$

olarak elde ederiz. Teorem 9.4'dan

$$\begin{aligned}
 Var(X) &= Var(I_1 + I_2 + \dots + I_n) \\
 &= Var(I_1) + Var(I_2) + \dots + Var(I_n) \\
 Var(I_i) &= E[I_i^2] - p^2 \\
 &= 0^2 * (1-p) + 1^2 * p - p^2 \\
 &= p - p^2 = p(1-p) \Rightarrow \\
 Var(X) &= p(1-p) + p(1-p) + \dots + p(1-p) \\
 &= np(1-p)
 \end{aligned}$$

Örnek 13.10. Bir içme suyu kuyusunda %30 oranında suda kirlilik tespit edilmiştir. Rasgele seçilen 10 içme suyu kuyusundan 3 ünde kirlilik rastlanması olasılığı nedir?

Çözüm:

$$b(3, 10, 0.3) = C_{10,3} 0.3^3 0.7^7 = \frac{10!}{3!7!} 0.3^3 0.7^7 = 0.26683$$

□

Örnek 13.11. Hilesiz bir zar 7 kez atılıyor. 5 yada 6 nın 3 defa gelmesi olasılığı nedir?

Çözüm:

$$\begin{aligned}
 P(\{5, 6\}) &= P(5) + P(6) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3} \\
 b\left(3, 7, \frac{1}{3}\right) &= C_{7,3} \left(\frac{1}{3}\right)^3 \left(\frac{2}{3}\right)^4 = \frac{7!}{3!4!} \left(\frac{1}{3}\right)^3 \left(\frac{2}{3}\right)^4 = \frac{560}{2187}
 \end{aligned}$$

□

Örnek 13.12. Bir A bölgesinde yaşayan insanların ölümcül bir hastalığa yakalanarak ölme riski %60 olup bu bölgede rasgele seçilen 15 kişiden hastaların bu hastalığa yakalanmayanların 2.206 ile 9.794 arasında olması olasılığı nedir?

Çözüm:

$$E[X] = np = 15 * 0.4 = 6 \Rightarrow$$

$$6 - 2.206 = 3.794$$

$$9.794 - 6 = 3.794 \Rightarrow$$

$$P(|X - 6| < 3.794) = 1 - P(|X - 6| \geq 3.794)$$

$$Var(X) = np(1-p) = 15 * 0.4 * 0.6 = 3.6$$

Chebyshev eşitsizliğinden

$$\begin{aligned}
 P(|X - \mu| \geq a) &\leq \frac{Var(X)}{a^2} \Rightarrow \\
 P(|X - 6| \geq 3.794) &\leq \frac{Var(X)}{(3.794)^2} \Rightarrow \\
 P(|X - 6| \geq 3.794) &\leq \frac{3.6}{(3.794)^2} = 0.25010 \Rightarrow \\
 1 - P(|X - 6| \geq 3.794) &\geq 1 - 0.25010 = 0.7499 \Rightarrow \\
 P(|X - 6| < 3.794) &= P(2.206 < X < 9.794) \geq 0.7499
 \end{aligned}$$

□

13.2 Hipergeometrik Dağılım (Hypergeometric Distribution)

Hipergeometrik olasılık dağılımı binom olasılık dağılımı ile yakından ilişkilidir. Söz konusu iki olasılık dağılımı arasındaki fark hipergeometrik olasılık dağılımında denemelerin bağımsız olmamasıdır. Bu durum başarı sonucu elde etme olasılığının her deneme için farklı olmasına neden olmaktadır. Hipergeometrik dağılımda da iki adet olası sonuç vardır: Başarı ve başarısızlık ve denemeler birbirine bağımlıdır.

Örnek 13.13. *Bir kavanozda 5'i beyaz ve 8'i siyah renkli olmak üzere 13 bilye vardır. İade edilmeksizin 3 bilye çekiliyor. Bu durumda “X hipergeometrik rassal değişkeni” çekilen beyaz (siyah) renkli bilyelerin sayısıdır.*

Tanım 13.14. *X, raslantı değişkeninin hipergeometrik dağılım fonksiyonu, N çeşitten, n adet seçilenlerin içinden k tane başarılı olması ve (n - k) tane başarısız olması durumunda*

$$h(x, N, n, k) = \frac{C_{k,x} C_{N-k, n-x}}{C_{N,n}} = \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}}$$

olarak tanımlanır.

Örnek 13.15. *12 kişilik bir yönetim kurulunun 7 üyesi belirli bir konuda alınacak karar tasarısının lehinde, 5'i ise aleyhinde düşünmektedir. Bu kuruldan rassal olarak seçilen 4 kişilik komisyonda karar tasarısının lehinde düşünen üye sayısının,*

- İki olması*
- En az üç olması*
- En fazla iki olması*

Çözüm: X=tasarısının lehinde düşünen üye sayısı

$$N = 12$$

$$n = 4$$

$$k = 7$$

a)

$$P(X = 2) = h(2, N, n, k) = \frac{C_{7,2}C_{5,2}}{C_{12,4}} = \frac{7!}{2!5!} \frac{5!}{2!3!} \frac{4!8!}{12!} = \frac{14}{33} = 0.424\ 24$$

b)

$$\begin{aligned} P(X \geq 3) &= h(3, N, n, k) + h(4, N, n, k) \\ &= \frac{C_{7,3}C_{5,1}}{C_{12,4}} + \frac{C_{7,4}C_{5,0}}{C_{12,4}} \\ &= \frac{7!}{3!4!} \frac{5!}{1!4!} \frac{4!8!}{12!} + \frac{7!}{4!3!} \frac{5!}{0!5!} \frac{4!8!}{12!} \\ &= \frac{14}{33} = 0.424\ 24 \end{aligned}$$

c)

$$\begin{aligned} P(X \leq 2) &= h(0, N, n, k) + h(1, N, n, k) + h(2, N, n, k) \\ &= \frac{C_{7,0}C_{5,4}}{C_{12,4}} + \frac{C_{7,1}C_{5,3}}{C_{12,4}} + \frac{C_{7,2}C_{5,2}}{C_{12,4}} \\ &= \frac{7!}{0!7!} \frac{5!}{1!4!} \frac{4!8!}{12!} + \frac{7!}{1!6!} \frac{5!}{2!3!} \frac{4!8!}{12!} + \frac{7!}{2!5!} \frac{5!}{2!3!} \frac{4!8!}{12!} \\ &= \frac{19}{33} = 0.575\ 76 \end{aligned}$$

□

Teorem 13.16. *X raslantı değişkeninin dağılım fonksiyonu $h(x, N, n, k)$ kipergeometrik dağılımı olsun. Buna göre*

$$\begin{aligned} E[X] &= \frac{nk}{N} \\ Var(X) &= \frac{N-n}{N-1} n \frac{k}{N} \left(1 - \frac{k}{N}\right) \end{aligned}$$

İspat: i. denemeyi I_i ile gösterelim. I_i de başarılı olan sonuçları 1 veya başarısızlıkları da 0 ile ve n bağımsız denemenin sonucundaki X raslantı değişkeninin,

$$X = I_1 + I_2 + \dots + I_n$$

olarak yazabiliriz. Ayrıca, N çeşitten, n adet seçilenlerin içinden k tane başarılı olması

$$\begin{aligned} P(I_i = 1) &= \frac{k}{N} \\ E[I_i] &= 1 * \frac{k}{N} + 0 * \left(1 - \frac{k}{N}\right) = \frac{k}{N} \Rightarrow \\ E[X] &= E[I_1 + I_2 + \dots + I_n] \\ &= E[I_1] + E[I_2] + \dots + E[I_n] \\ &= n \frac{k}{N} \end{aligned}$$

olarak elde ederiz. Teorem 10.15'dan

$$\begin{aligned}
 Var(aX + bY + c) &= a^2 Var(X) + b^2 Var(Y) + 2ab Cov(X, Y) \\
 Var(X) &= Var(I_1 + I_2 + \dots + I_n) \\
 &= \sum_{i=1}^n Var(I_i) + \sum_{i+j=1}^n Cov(I_i, I_j) \\
 Var(I_i) &= E[I_i^2] - \left(\frac{k}{N}\right)^2 \\
 &= 0^2 * \left(1 - \frac{k}{N}\right) + 1^2 * \frac{k}{N} - \left(\frac{k}{N}\right)^2 \\
 &= \frac{k}{N} - \left(\frac{k}{N}\right)^2 = \frac{k}{N} \left(1 - \frac{k}{N}\right) \\
 Cov(X, Y) &= E[XY] - \mu_X \mu_Y \\
 Cov(I_i, I_j) &= E[I_i I_j] - \mu_{I_i} \mu_{I_j} \\
 E[I_i I_j] &= P(X_i = 1, X_j = 1) = \frac{k}{N} * \frac{k-1}{N-1} \Rightarrow \\
 Cov(I_i, I_j) &= \frac{k}{N} * \frac{k-1}{N-1} - \left(\frac{k}{N}\right)^2 \\
 &= \frac{k}{N} \left(\frac{k-1}{N-1} - \frac{k}{N}\right) = -\frac{1}{N^2} k \frac{N-k}{N-1} \\
 &= -\frac{1}{N(N-1)} k \left(1 - \frac{k}{N}\right) \\
 Var(X) &= \left[n \frac{k}{N} \left(1 - \frac{k}{N}\right) - n(n-1) \left(-\frac{1}{N(N-1)} k \left(1 - \frac{k}{N}\right)\right) \right] \\
 &= n \frac{k}{N} \left(1 - \frac{k}{N}\right) \left[1 + \frac{n-1}{N-1}\right] \\
 &= n \frac{k}{N} \left(1 - \frac{k}{N}\right) \frac{N-n}{N-1}
 \end{aligned}$$

13.3 Poisson Dağılımı (Poisson Distribution)

Poisson olasılık dağılımı kesikli olasılık dağılımlarından bir diğeridir. Poisson olasılık dağılımı, belli bir zaman diliminde ya da belli bir fiziksel mekanda oluşan olay sayısı ile ilgili olasılıkları tanımlamakta kullanılmaktadır.

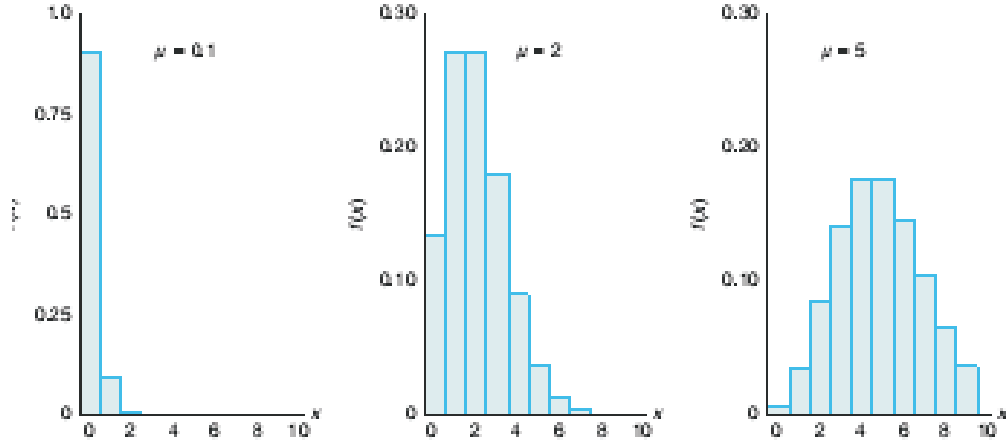
Örnek 13.17. *Ankara'da bir hafta içinde meydana gelen trafik kazası sayısı, Bir süper markette belirli bir kasiyere bir saatte gelen müşteri sayısı, Bir hafta içerisinde herhangi bir fabrika da üretim sürecinde kullanılan bir makinanın bozulma sayısı, Maden işçilerinin bir yıl içerisinde yaptıkları grev sayısı, Boğaz Köprüsü'ndeki herhangi bir gışeden 12 saatte geçen araç sayısı*

Tanım 13.18. *X raslantı değişkeninin Poisson dağılım fonksiyonu*

$$P(X = i) = p(i, \lambda) = e^{-\lambda} \frac{\lambda^i}{i!}, i = 0, 1, 2, 3, \dots$$

Burada λ belirli bir zaman diliminde x 'in gerçekleşme olasılıklarının beklenen değeridir. Ve Poisson dağılımı gerçekten bir olasılık dağılım fonksiyonudur, çünkü:

$$\sum_{i=0}^{\infty} p(i, \lambda) = e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = e^{-\lambda} e^{\lambda} = 1$$



Örnek 13.19. Bir laboratuvar deneyinde radyoaktif elementlerin 1 milisaniyede tepki vermesi ortalama olarak 4 tür (1 milisaniyede ortalama 4 element tepki veriyor.) Buna göre 1 milisaniyede 6 elementin tepki vermesi olasılığı nedir?

Çözüm:

$$\begin{aligned} \lambda &= 4 \\ P(X=6) &= p(6, 4) = e^{-4} \frac{4^6}{6!} = 0.10420 \end{aligned}$$

□

Örnek 13.20. Belli bir bölgedeki otobanda ortalama kaza sayısı haftada 3 olarak belirlenmiştir. Buna göre haftada en az 1 kaza yapma olasılığı nedir?

Çözüm:

$$\begin{aligned} \lambda &= 3 \\ P(X \geq 1) &= 1 - P(X=0) \\ &= 1 - p(0, 3) \\ &= 1 - e^{-3} \frac{3^0}{0!} \\ &= 0.95021 \end{aligned}$$

□

Örnek 13.21. Bir sigorta şirketinin günlük olarak sigortaladığı müşterilerinin sayısı 5 tir. Buna göre göre en çok 3 müşterinin sigortalanması olasılığı nedir? Gelecek 5 günden herhangi 3 ünde 4 müşterinin sigortalanması olasılığı nedir?

Çözüm:

$$\begin{aligned}
 \lambda &= 5 \\
 P(X \leq 3) &= p(0, 5) + p(1, 5) + p(2, 5) \\
 &= e^{-5} \frac{5^0}{0!} + e^{-5} \frac{5^1}{1!} + e^{-5} \frac{5^2}{2!} \\
 &= 0.12465 \\
 &\quad 1 - gunde \\
 P(X = 4) &= p(4, 5) \\
 &= e^{-5} \frac{5^4}{4!} = 0.175475 \\
 &\quad Y - gun \\
 P(Y = 3) &= \binom{5}{3} (0.175475)^3 (1 - 0.175475)^2 \\
 &= \frac{5!}{2!3!} (0.175475)^3 (1 - 0.175475)^2 \\
 &= 0.036733
 \end{aligned}$$

□

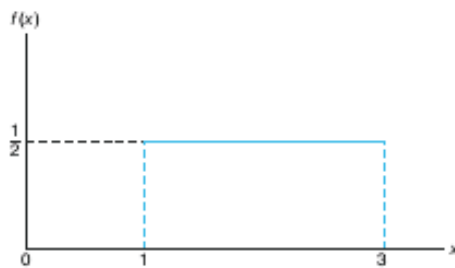
14 Önemli Sürekli Dağılım Fonksiyonları

14.1 Düzgün Sürekli Dağılım (Continuous Uniform Distribution)

Tanım 14.1. $[A, B]$ aralığında X sürekli raslantı değişkeninin düzgün sürekli olasılık dağılım fonksiyonu

$$f(x; A, B) = \begin{cases} \frac{1}{B-A} & , \quad A \leq x \leq B \\ 0 & , \quad \text{Aksi durum} \end{cases}$$

ile tanımlanır.



Örnek 14.2. Bir şirket, konferans salonunu, 4 saatten fazla kiralamamaktadır. X raslantı değişkeni, bir konferansın süresini vermek üzere sürekli olasılık dağılım fonksiyonunu tanımlayıp, herhangi bir konferansın en az 3 saat olması olasılığı nedir?

Çözüm: X raslantı değişkeninin sürekli olasılık dağılım fonksiyonunu

$$f(x; 0, 4) = \begin{cases} \frac{1}{4} & , \quad 0 \leq x \leq 4 \\ 0 & , \quad \text{Aksi durum} \end{cases}$$

ile tanımlayabiliriz. Buna göre

$$P(X \geq 3) = \int_3^{\infty} f(x; 0, 4) dx = \int_3^4 \frac{1}{4} dx = \frac{1}{4}$$

□

Teorem 14.3. $[A, B]$ aralığında X sürekli raslantı değişkeninin düzgün sürekli olasılık dağılım fonksiyonu $f(x; A, B)$ olmak üzere X in beklenen değeri ve varyansını aşağıdaki şekilde veririz:

$$\begin{aligned} \mu_X &= E[X] = \frac{A+B}{2} \\ \text{Var}(X) &= \frac{(B-A)^2}{12} \end{aligned}$$

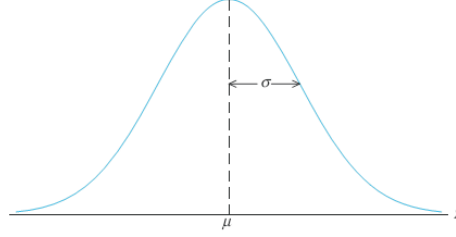
İspat:

$$\begin{aligned} \mu_X = E[X] &= \int_{-\infty}^{\infty} x f(x; A, B) dx = \int_A^B \frac{x}{B-A} dx = \\ &= \frac{1}{B-A} \left. \frac{x^2}{2} \right|_{x=A}^{x=B} = \frac{B^2 - A^2}{2(B-A)} = \frac{A+B}{2} \\ \text{Var}(X) = E[(X - \mu_X)^2] &= \int_{-\infty}^{\infty} (x - \mu_X)^2 f(x; A, B) dx \\ &= \int_A^B \frac{(x - \frac{A+B}{2})^2}{B-A} dx \\ &= \frac{1}{B-A} \int_A^B \left(x - \frac{A+B}{2} \right)^2 dx, x - \frac{A+B}{2} = u \\ &= \frac{1}{B-A} \int_{-(B-A)/2}^{(B-A)/2} u^2 du \\ &= \frac{1}{B-A} \left. \frac{u^3}{3} \right|_{u=-(B-A)/2}^{u=(B-A)/2} \\ &= \frac{1}{B-A} * 2 * \frac{(B-A)^3}{3*8} = \frac{(B-A)^2}{12} \end{aligned}$$

14.2 Normal Dağılım (Normal Distribution)

İstatistikte çok önemli bir yeri olan dağılımdır. Grafiğine normal eğri (normal curve) denir. Örneğin, yağmur yağış miktarının modellenmesi gibi meteorolojik deneyler normal dağılım ile ifade edilir. 1733 yılında Abraham DeMoivre normal eğrilerin denklemlerini geliştirmiştir. Pek çok istatistik kaynaklarında normal dağılım,

Karl Friedrich Gauss (1777–1855) anısına, Gauss dağılımı (the Gaussian distribution) da denir. Gauss, aynı büyüklüklerin tekrarlanması hatasını normal eğriler ile ilişkilendirmiştir.



Tanım 14.4. X sürekli raslantı değişkeninin beklenen değeri $E[X] = \mu_X$, $\sigma^2 = \text{Var}(X)$ olmak üzere normal dağılımı $n(x; \mu, \sigma)$ ile gösterilir ve

$$n(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu_X)^2}{2\sigma^2}}, -\infty < x < \infty$$

ile tanımlanır. Ve aşağıdaki özelliklere sahiptir:

- (1) $n(x; \mu, \sigma)$ normal dağılımı maksimum değerini $x = \mu$ noktasında alır.
- (2) $n(x; \mu, \sigma)$ normal dağılımı $x = \mu$ n doğrusuna göre simetriktir.
- (3) $n(x; \mu, \sigma)$ normal dağılımının yatay asimptotu x -eksenidir.
- (4) $n(x; \mu, \sigma)$ normal dağılımı ve x -ekseni arasında kalan bölgenin alanın 1 dir.

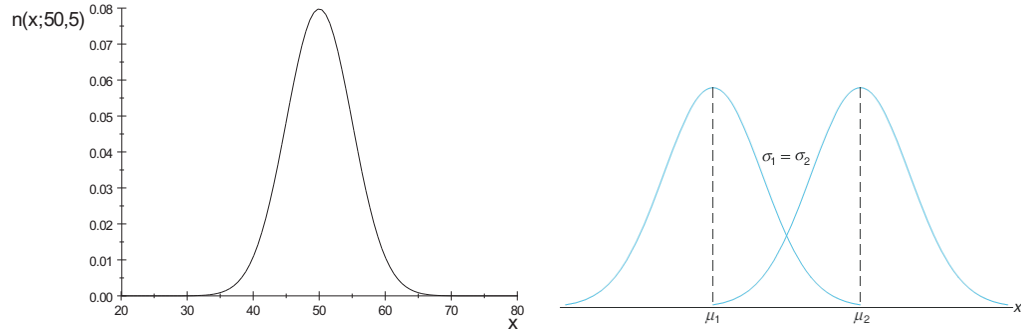


FIGURE 14.1. (a) Normal dağılım eğrisi, $\mu = 50, \sigma = 5$, (b) $\sigma_1 = \sigma_2, \mu_1 < \mu_2$ olduğunda normal dağılım eğrisi

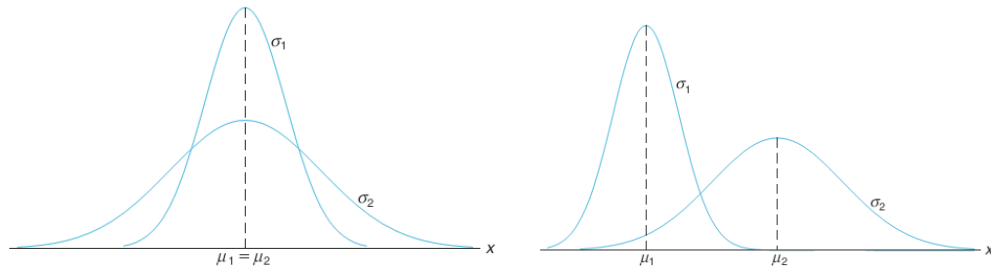


FIGURE 14.2. (a) $\mu_1 = \mu_2, \sigma_1 < \sigma_2$ olduğunda normal dağılım eğrisi, (b) $\mu_1 < \mu_2, \sigma_1 < \sigma_2$ olduğunda normal dağılım eğrisi

Tanım 14.5. X sürekli raslantı değişkeninin normal dağılımı $n(x; \mu, \sigma)$ olmak üzere

$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} n(x; \mu, \sigma) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{(x-\mu_X)^2}{2\sigma^2}} dx$$

ile tanımlanır.

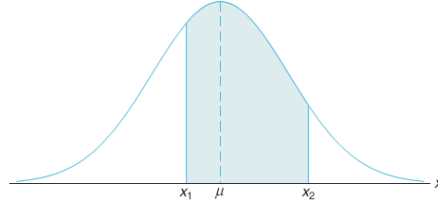


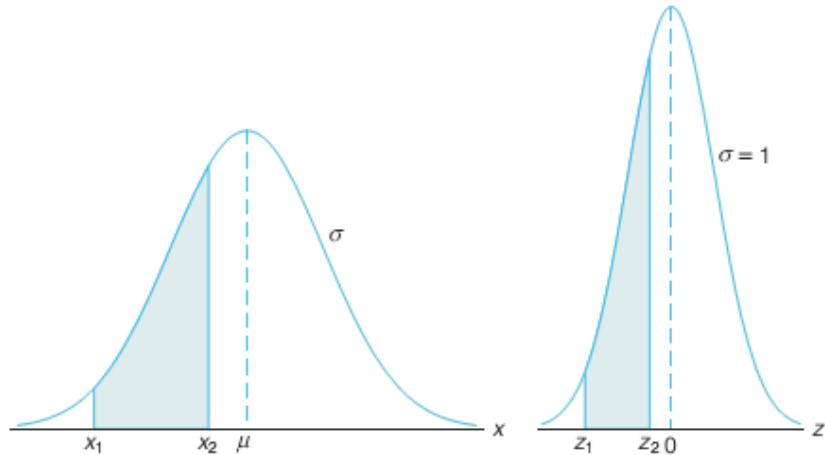
FIGURE 14.3. $P(x_1 < X < x_2)$

Tanım 14.6. X sürekli raslantı değişkeninin beklenen değeri $E[X] = 0, Var(X) = 1$ ise $n(x; 0, 1)$ normal dağılımına standart normal dağılım denir.

$$n(x; 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

Lemma 14.7.

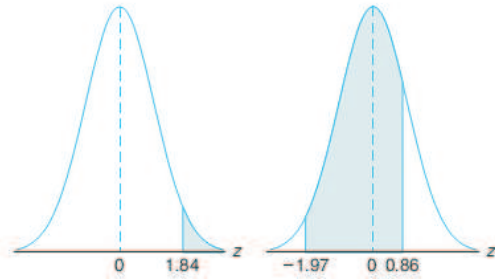
$$\begin{aligned} P(x_1 < X < x_2) &= \int_{x_1}^{x_2} n(x; \mu, \sigma) dx \\ &= \int_{z_1}^{z_2} n(z; 0, 1) dz, \\ z_1 &= \frac{x_1 - \mu_X}{\sigma}, \\ z_2 &= \frac{x_2 - \mu_X}{\sigma} \end{aligned}$$



İspat:

$$\begin{aligned}
 P(x_1 < X < x_2) &= \int_{x_1}^{x_2} n(x; \mu, \sigma) dx \\
 &= \frac{1}{\sigma\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{(x-\mu_X)^2}{2\sigma^2}} dx, \\
 z &= \frac{x - \mu_X}{\sigma} \Rightarrow \sigma dz = dx \\
 P(x_1 < X < x_2) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{z_1}^{z_2} e^{-\frac{z^2}{2}} \sigma dz \\
 &= \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{-\frac{z^2}{2}} dz = \int_{z_1}^{z_2} n(z; 0, 1) dz \\
 &= P(z_1 < Z < z_2)
 \end{aligned}$$

Örnek 14.8. Standart normal dağılım için $P(Z > 1.84)$, $P(-1.97 < Z < 0.86)$ olasılıklarını hesaplayınız.



Çözüm: $P(Z \leq 1.84)$ değeri için $(1.84 = 1.8 + 0.04)$ tabloda dikey sütunda $z = 1.8$ değerini bulduktan sonra yatay sütunda $z = 0.04$ değerinin kesiştiği değeri bulmaktır. Buna göre $P(Z \leq 1.84) = 0.9671$ dir.

$$P(Z > 1.84) = 1 - P(Z \leq 1.84)$$

$$= 1 - 0.9671$$

$$P(-1.97 < Z < 0.86) = P(Z < 0.86) - P(Z \leq -1.97)$$

$P(Z < 0.86)$ değeri için $(0.86 = 0.8 + 0.06)$ tabloda dikey sütunda $z = 0.8$ değerini bulduktan sonra yatay sütunda $z = 0.06$ değerinin kesiştiği değeri bulmaktır. Buna göre $P(Z \leq 0.86) = 0.8051$ dir. $P(Z \leq -1.97)$ değeri için $(-1.97 = -1.9 - .07)$ tabloda dikey sütunda $z = -1.9$ değerini bulduktan sonra yatay sütunda $z = 0.07$ değerinin kesiştiği değeri bulmaktır. Buna göre $P(Z \leq -1.97) = 0.0244$ dir. Buna göre

$$P(-1.97 < Z < 0.86) = P(Z < 0.86) - P(Z \leq -1.97)$$

$$= 0.8051 - 0.0244 = 0.7807$$

□

Örnek 14.9. Standart normal dağılım için

(i) $P(Z > k) = 0.3015$

(ii) $P(k < Z < -0.18) = 0.4197$ değerlerini veren k nedir?

Çözüm: (i)

$$P(Z > k) = 0.3015 = 1 - P(Z \leq k) \Rightarrow$$

$$P(Z \leq k) = 1 - 0.3015 = 0.6985$$

0.69 değerinde olan $z=0.5$ ve satırlarını takip edersek $z=0.52$ sonucunu elde ederiz. Böylece $k = 0.52$ değerini buluruz.

(ii)

$$\begin{aligned} 0.4197 &= P(k < Z < -0.18) = P(Z < -0.18) - P(Z \leq k) \\ &= 0.4286 - P(Z \leq k) \Rightarrow P(Z \leq k) = 0.4286 - 0.4197 \\ &= 0.0089 \end{aligned}$$

değerini alan tabloda $z = -2.37$ değeridir. O halde $k = -2.37$. □

Örnek 14.10. X sürekli raslantı değişkeninin beklenen değeri $\mu_X = 50, \sigma = 10$ olmak üzere $P(45 < X < 62)$ olasılığını bulunuz.

Çözüm: $x_1 = 45, x_2 = 62,$

$$\begin{aligned} P(45 < X < 62) &= \int_{x_1}^{x_2} n(x; 50, 10) dx \\ &= \int_{z_1}^{z_2} n(z; 0, 1) dz \\ &= P(z_1 < Z < z_2) \end{aligned}$$

$$z_1 = \frac{x_1 - \mu_X}{\sigma} = \frac{45 - 50}{10} = -0.5,$$

$$z_2 = \frac{x_2 - \mu_X}{\sigma} = \frac{62 - 50}{10} = 1.2$$

$$P(45 < X < 62) = P(-0.5 < Z < 1.2)$$

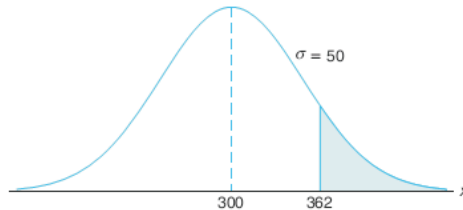
$$= P(Z < 1.2) - P(Z \leq -0.5)$$

$$= 0.8849 - 0.3085$$

$$= 0.5764$$

□

Örnek 14.11. X sürekli raslantı değişkeninin beklenen değeri $\mu_X = 300, \sigma = 50$ olmak üzere X raslantı değişkeninin 362 den büyük olması olasılığı nedir?



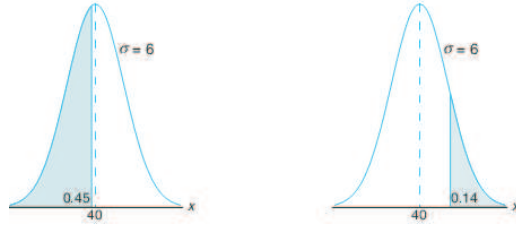
Çözüm:

$$z_2 = \frac{362 - 300}{50} = 1.24$$

$$\begin{aligned} P(X > 362) &= P(Z > 1.24) \\ &= 1 - P(Z \leq 1.24) \\ &= 1 - 0.8925 \\ &= 0.1075 \end{aligned}$$

□

Örnek 14.12. X sürekli raslantı değişkeninin beklenen değeri $\mu_X = 40, \sigma = 6$ olmak üzere şekildeki alanların %45 olması ve %14 olması için x değeri ne olmalıdır?



Çözüm:

$$\begin{aligned} P(X < x) &= 0.45 \Rightarrow x = ? \\ P(X < x) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-\mu_X)^2}{2\sigma^2}} dt \\ &= \frac{1}{6\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-40)^2}{2 \cdot 6^2}} dt \\ z &= \frac{t-40}{6} \Rightarrow 6dz = dt \\ &= \frac{1}{6\sqrt{2\pi}} \int_{-\infty}^{\frac{x-40}{6}} e^{-\frac{z^2}{2}} 6dz \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-40}{6}} e^{-\frac{z^2}{2}} dz \\ &= P\left(Z < \frac{x-40}{6}\right) = 0.45 \end{aligned}$$

Bu ise tabloda baktığımızda $P(Z < -0.13) = 0.4483$ değeriyle birlikte

$$\frac{x-40}{6} \simeq -0.13 \Rightarrow x \simeq 6 \cdot (-0.13) + 40 = 39.22$$

olmalıdır. Benzer şekilde

$$\begin{aligned}
 P(X > x) &= 0.14 \Rightarrow x = ? \\
 P(X > x) &= P\left(Z > \frac{x-40}{6}\right) = 0.14 \Rightarrow \\
 0.14 &= 1 - P\left(Z \leq \frac{x-40}{6}\right) \Rightarrow \\
 P\left(Z \leq \frac{x-40}{6}\right) &= 1 - 0.14 = 0.86 \Rightarrow
 \end{aligned}$$

tabloda $P(Z|1.08)=0.8599$ olması ile birlikte

$$\frac{x-40}{6} \simeq 1.08 \Rightarrow x \simeq 6 * 1.08 + 40 = 46.48$$

□

Örnek 14.13. 3 yıllık ortalama pil ömrüne sahip bir bataryanın standart sapması 0.5 olarak veriliyor. Bataryanın ömrünün normal dağılıma sahip olduğu bilindiğine göre bir bataryanın ömrünün 2.3 yıldan az olması olasılığı nedir?

Çözüm: $\mu_X = 3, \sigma = 0.5 \Rightarrow P(X < 2.3)$

$$\begin{aligned}
 P(X < 2.3) &= \frac{1}{0.5\sqrt{2\pi}} \int_{-\infty}^{2.3} e^{-\frac{(t-3)^2}{2*0.5^2}} dt \\
 z &= \frac{t - \mu_X}{\sigma} = \frac{t - 3}{0.5} \\
 &= \frac{1}{0.5\sqrt{2\pi}} \int_{-\infty}^{\frac{2.3-3}{0.5}} e^{-\frac{z^2}{2}} 0.5 dz \\
 &= P\left(Z < \frac{2.3-3}{0.5} = -1.4\right) \\
 &= 0.0808
 \end{aligned}$$

□

Örnek 14.14. Bir elektrik fabrikası, bir lambanın ışık vermesini normal dağılımlı olarak verip, ortalama ömrünün 800 saat ve standart sapmasının 40 olduğunu belirtmiştir. Buna göre bir lambanın 778 ile 834 saatleri arasında ışık vermesi olasılığı nedir?

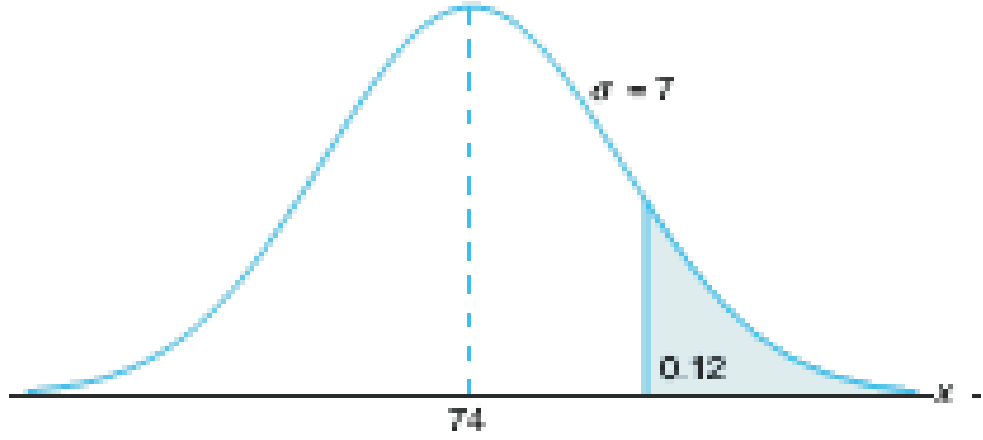
Çözüm:

$$\begin{aligned}
 \mu_x &= 800, \sigma = 40 \\
 P(778 < X < 834) &= P\left(\frac{778-800}{40} < Z < \frac{834-800}{40}\right) \\
 &= P(-0.55 < Z < 0.85) \\
 &= P(Z < 0.85) - P(Z < -0.55) \\
 &= 0.8023 - 0.2912 = 0.5111
 \end{aligned}$$

□

Örnek 14.15. Bir sınıfta not ortalaması 74 ve standart sapması 7 olarak veriliyor. Sınıftaki A alanların yüzdesi %12, ve sınıftaki not dağılımı normal dağılım ile verildiğine göre sınıftaki en düşük A notu ve en yüksek B notu nedir?

Çözüm:



$$\begin{aligned}
 \mu_X &= 74, \sigma = 7 \Rightarrow P(X > k) = 0.12 \Rightarrow \\
 P\left(Z > \frac{k - 74}{7}\right) &= 0.12 \Rightarrow 1 - P\left(Z \leq \frac{k - 74}{7}\right) = 0.12 \Rightarrow \\
 P\left(Z \leq \frac{k - 74}{7}\right) &= 1 - 0.12 = 0.88 \Rightarrow P(Z \leq 1.17) = 0.8790 \Rightarrow \\
 \frac{k - 74}{7} &\simeq 1.17 \Rightarrow k = 7 * 1.17 + 74 = 82.19
 \end{aligned}$$

O halde en düşük A notu 83 olup en yüksek B notu da 82 dir.

□

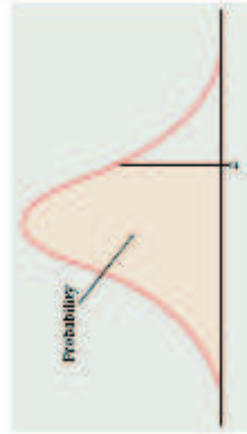


Table entry for z is the area under the standard normal curve to the left of z .

TABLE A. Standard normal probabilities (continued)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7020	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9383	.9395	.9408	.9419	.9429	.9441
1.6	.9452	.9463	.9473	.9483	.9493	.9503	.9513	.9522	.9531	.9540
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9978	.9979	.9980	.9981	.9982	.9983
2.9	.9984	.9985	.9986	.9987	.9988	.9989	.9990	.9991	.9992	.9993
3.0	.9994	.9995	.9996	.9997	.9998	.9999	.9999	.9999	.9999	.9999
3.1	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999
3.2	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999
3.3	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999
3.4	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999

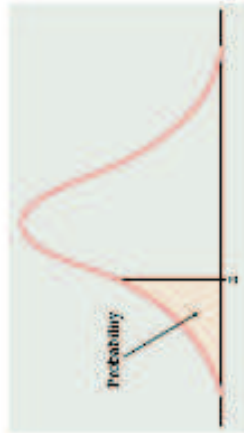


Table entry for z is the area under the standard normal curve to the left of z .

TABLE A. Standard normal probabilities

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0006	.0006	.0005
-3.1	.0010	.0009	.0009	.0008	.0008	.0008	.0008	.0008	.0008	.0007
-3.0	.0013	.0013	.0012	.0011	.0011	.0011	.0011	.0011	.0011	.0010
-2.9	.0016	.0015	.0015	.0014	.0014	.0014	.0014	.0014	.0014	.0013
-2.8	.0019	.0018	.0018	.0017	.0017	.0017	.0017	.0017	.0017	.0016
-2.7	.0023	.0022	.0022	.0021	.0021	.0021	.0021	.0021	.0021	.0020
-2.6	.0027	.0026	.0026	.0025	.0025	.0025	.0025	.0025	.0025	.0024
-2.5	.0031	.0030	.0030	.0029	.0029	.0029	.0029	.0029	.0029	.0028
-2.4	.0036	.0035	.0035	.0034	.0034	.0034	.0034	.0034	.0034	.0033
-2.3	.0040	.0039	.0039	.0038	.0038	.0038	.0038	.0038	.0038	.0037
-2.2	.0045	.0044	.0044	.0043	.0043	.0043	.0043	.0043	.0043	.0042
-2.1	.0050	.0049	.0049	.0048	.0048	.0048	.0048	.0048	.0048	.0047
-2.0	.0054	.0053	.0053	.0052	.0052	.0052	.0052	.0052	.0052	.0051
-1.9	.0059	.0058	.0058	.0057	.0057	.0057	.0057	.0057	.0057	.0056
-1.8	.0064	.0063	.0063	.0062	.0062	.0062	.0062	.0062	.0062	.0061
-1.7	.0069	.0068	.0068	.0067	.0067	.0067	.0067	.0067	.0067	.0066
-1.6	.0074	.0073	.0073	.0072	.0072	.0072	.0072	.0072	.0072	.0071
-1.5	.0078	.0077	.0077	.0076	.0076	.0076	.0076	.0076	.0076	.0075
-1.4	.0083	.0082	.0082	.0081	.0081	.0081	.0081	.0081	.0081	.0080
-1.3	.0088	.0087	.0087	.0086	.0086	.0086	.0086	.0086	.0086	.0085
-1.2	.0093	.0092	.0092	.0091	.0091	.0091	.0091	.0091	.0091	.0090
-1.1	.0098	.0097	.0097	.0096	.0096	.0096	.0096	.0096	.0096	.0095
-1.0	.0103	.0102	.0102	.0101	.0101	.0101	.0101	.0101	.0101	.0100
-0.9	.0106	.0105	.0105	.0104	.0104	.0104	.0104	.0104	.0104	.0103
-0.8	.0109	.0108	.0108	.0107	.0107	.0107	.0107	.0107	.0107	.0106
-0.7	.0113	.0112	.0112	.0111	.0111	.0111	.0111	.0111	.0111	.0110
-0.6	.0116	.0115	.0115	.0114	.0114	.0114	.0114	.0114	.0114	.0113
-0.5	.0119	.0118	.0118	.0117	.0117	.0117	.0117	.0117	.0117	.0116
-0.4	.0122	.0121	.0121	.0120	.0120	.0120	.0120	.0120	.0120	.0119
-0.3	.0125	.0124	.0124	.0123	.0123	.0123	.0123	.0123	.0123	.0122
-0.2	.0128	.0127	.0127	.0126	.0126	.0126	.0126	.0126	.0126	.0125
-0.1	.0131	.0130	.0130	.0129	.0129	.0129	.0129	.0129	.0129	.0128
0.0	.0134	.0133	.0133	.0132	.0132	.0132	.0132	.0132	.0132	.0131
0.1	.0136	.0135	.0135	.0134	.0134	.0134	.0134	.0134	.0134	.0133
0.2	.0139	.0138	.0138	.0137	.0137	.0137	.0137	.0137	.0137	.0136
0.3	.0141	.0140	.0140	.0139	.0139	.0139	.0139	.0139	.0139	.0138
0.4	.0144	.0143	.0143	.0142	.0142	.0142	.0142	.0142	.0142	.0141
0.5	.0146	.0145	.0145	.0144	.0144	.0144	.0144	.0144	.0144	.0143
0.6	.0147	.0147	.0147	.0146	.0146	.0146	.0146	.0146	.0146	.0145
0.7	.0148	.0148	.0148	.0147	.0147	.0147	.0147	.0147	.0147	.0146
0.8	.0149	.0149	.0149	.0148	.0148	.0148	.0148	.0148	.0148	.0147
0.9	.0149	.0149	.0149	.0148	.0148	.0148	.0148	.0148	.0148	.0147
1.0	.0149	.0149	.0149	.0148	.0148	.0148	.0148	.0148	.0148	.0147
1.1	.0149	.0149	.0149	.0148	.0148	.0148	.0148	.0148	.0148	.0147
1.2	.0149	.0149	.0149	.0148	.0148	.0148	.0148	.0148	.0148	.0147
1.3	.0149	.0149	.0149	.0148	.0148	.0148	.0148	.0148	.0148	.0147
1.4	.0149	.0149	.0149	.0148	.0148	.0148	.0148	.0148	.0148	.0147
1.5	.0149	.0149	.0149	.0148	.0148	.0148	.0148	.0148	.0148	.0147
1.6	.0149	.0149	.0149	.0148	.0148	.0148	.0148	.0148	.0148	.0147
1.7	.0149	.0149	.0149	.0148	.0148	.0148	.0148	.0148	.0148	.0147
1.8	.0149	.0149	.0149	.0148	.0148	.0148	.0148	.0148	.0148	.0147
1.9	.0149	.0149	.0149	.0148	.0148	.0148	.0148	.0148	.0148	.0147
2.0	.0149	.0149	.0149	.0148	.0148	.0148	.0148	.0148	.0148	.0147
2.1	.0149	.0149	.0149	.0148	.0148	.0148	.0148	.0148	.0148	.0147
2.2	.0149	.0149	.0149	.0148	.0148	.0148	.0148	.0148	.0148	.0147
2.3	.0149	.0149	.0149	.0148	.0148	.0148	.0148	.0148	.0148	.0147
2.4	.0149	.0149	.0149	.0148	.0148	.0148	.0148	.0148	.0148	.0147
2.5	.0149	.0149	.0149	.0148	.0148	.0148	.0148	.0148	.0148	.0147
2.6	.0149	.0149	.0149	.0148	.0148	.0148	.0148	.0148	.0148	.0147
2.7	.0149	.0149	.0149	.0148	.0148	.0148	.0148	.0148	.0148	.0147
2.8	.0149	.0149	.0149	.0148	.0148	.0148	.0148	.0148	.0148	.0147
2.9	.0149	.0149	.0149	.0148	.0148	.0148	.0148	.0148	.0148	.0147
3.0	.0149	.0149	.0149	.0148	.0148	.0148	.0148	.0148	.0148	.0147
3.1	.0149	.0149	.0149	.0148	.0148	.0148	.0148	.0148	.0148	.0147
3.2	.0149	.0149	.0149	.0148	.0148	.0148	.0148	.0148	.0148	.0147
3.3	.0149	.0149	.0149	.0148	.0148	.0148	.0148	.0148	.0148	.0147
3.4	.0149	.0149	.0149	.0148	.0148	.0148	.0148	.0148	.0148	.0147

FIGURE 14.4. Normal Dağılım Tablosu

Chapter 2

İstatistik

1 Giriş - Örneklem Teorisi (Sampling Theory)

İstatistik,

- verilerin toplanması,
- organize edilmesi,
- özetlenmesi,
- sunulması,
- tahlil edilmesi ve
- bu verilerden bir sonuca varılabilmesi için

kullanılan bilimsel metotlar topluluğudur.

Tanım 1.1. *Üzerinde çalışılan tüm gruba ya da istatistiksel sonuçların genelleştirileceği gruba kitle denir. Ya da üzerinde çalıştığımız birimlerin (insanlar, nesneler, olaylar, kayıtlar) oluşturduğu kümeye kitle denir.*

Tanım 1.2. *Bir kitlenin her üyesi için ölçülebilir olan bir değişkene, bu kitlenin bir karakteristiği denir.*

Tanım 1.3. *Bir kitlenin tamamlayıcı sayısal ölçüsüne parametre denir. Kitledeki gözlemlere dayandığından değeri genellikle bilinmemektedir.*

Örnek 1.4. *Bir incelemede Türkiye’de yaşayan yetişkinlerin boy uzunlukları ile ilgileniliyorsa kitle, Türkiye’deki tüm yetişkinler; karakteristik, her bir şahsın boy uzunluğu olacaktır. Parametre, kitlenin ortalama boy uzunluğu olabilir veya en uzun boy uzunluğu veya en kısa boy uzunluğu olabilir.*

Örnek 1.5. *1999 yılında üretilen arabaların hızları ile ilgileniyorsak, kitle 1999 de üretilen arabaların tümü, karakteristik her bir arabanın hızı olacaktır. Parametre kitledeki en çok hız yapan arabanın hızı olabilir. Ortalama hız olabilir.*

- Herhangi deneysel bir çalışmaya başlamadan önce kitlenin ve ölçülecek karakteristiklerin açıkça tanımlanmış olması gerekir. Uygulamada, araştırmacı karşılaştığı kitlenin tüm üyelerini inceleyemez. Örneğin, bir bölgenin tüm yüksek tansiyonlu hastaları kontrol edilemediği gibi her birine deneysel ilaçta verilemez.

Buradaki gibi araştırmacı kitlenin sonlu sayıdaki üyelerinin karakteristiklerinin ölçümleriyle yetinmek zorunda kalacaktır.

Tanım 1.6. *Bir kitlenin belli bir özelliğini incelemek üzere, kitlede belirli kurallara göre seçilen birimler topluluğuna örneklem (ya da örnek) denir. Örneklem kitle birimlerinin gözlenen bir alt kümesidir.*

Tanım 1.7. *Bir örneklemün tamamlayıcı sayısal olgusuna örneklem istatistiği denir.*

Örnek 1.8. *Kitle Ankara ili sınırları içindeki tüm yüksek tansiyonlu hastalar ise ölçümlerin yapıldığı 200 kişilik bir grup bir örneklem oluşturur. İstatistik; örneklemün kan basıncı ortalaması olabilir.*

Tanım 1.9. *Bir örneklemden bulunacak değerleri x ile göstereceğiz, x_1 , ilk değere karşılık; x_2 , ikinci değere karşılık, v.s. gelir. Örneklemde n gözlem varsa, onların değerleri $x_1; x_2, \dots, x_n$ dir.*

<>: Tüm kitle yerine neden önce örneklem seçeriz?

1.: Örneklem seçmek için zaman faktörü, kaynaklar ve uygulama amacıyla araştırmacının sabrı kitleyi ölçmeyi olanaksız yapabilir.

Örnek 1.10. *Özel bir televizyon programının beğenilip beğenilmediği konusunda yargıya varmak için ülkedeki her bir televizyon seyircisinin düşüncesini öğrenmek uzun zaman alacak ve pahalı olacaktır. O halde örnekleme dayanarak yargıya varmamız gerekecektir.*

Örnek 1.11. *Karşılaşılan bir örnek olarak nüfus sayımını verebiliriz. Burada kitlenin tümü ölçülmektedir. Çünkü, zaman ve kaynaklar bunu yapmak için uygundur.*

2.: Örneklemenin tek seçenek olmasıdır.

Örnek 1.12. *İnsanlıkla ilgili tıbbi deneylerde ahlaki düşüncelere dayanan üçüncü bir neden vardır. Şöyle ki bir hastanın tedavisinde yeni bir yöntem geliştirmede, bir operasyonun etkilerinin hastaya acı ve zarar vermeden bulunması gerekmektedir. Yeni yöntemler için dikkatle seçilmiş bir örneklem kullanılması en iyi yoldur.*

<>: Örneklemün istenilen oranın altında olmamasını sağlamak gerekmektedir. Örneklem kötü bir şekilde seçilmişse, matematiksel ve istatistiksel yöntemler bizi istenilen sonuca götürmemektedir. Örneklem tekniği çağımızın bilimsel araştırmalarında çok önemli yer tutmaktadır. Örneklem seçiminde en çok dikkat edilmesi gereken nokta örneklemün kitle için temsil edici olmasıdır. Örneğin, ülkemizdeki yetişkinlerin boy uzunlukları ortalaması hesaplanmak isteniyorsa ve biz basketbol oyuncularından oluşan bir örneklem seçmişsek, örneklem kitleyi temsil edici olmayacak, bulunan sonuçlar kitle için uygulanamayacaktır.

Tanım 1.13. *Bir kitle sonlu sayıda ya da sayılabilir sayıdaki birimlerden ya da elemanlardan oluşmuşsa kitleye kesikli kitle denir. Kitle kesikli değilse sürekli kitle denir.*

Örnek 1.14. *Belli bir A tipinde sigara içenlerin sayısı kesikli olmakla beraber, A tipinde sigara içme zamanının uzunluğu sürekli dir.*

<>: İstatistiksel örnekleme kuramı uygulamalı problemleri çözmek için kullanılan yöntemleri vermektedir. Kuram rasgele seçilmiş örneklemelere dayanır. Rasgele örneklem olasılık örneklemini olarak ta düşünülür. Yani, kitledeki her bir öğenin örnekleme seçilme şansı bilinmektedir. Rasgele örneklemin tanımı gerçekten seçim yönteminin tanımıdır.

Tanım 1.15. *Bir kitlenin her bir öğesinin seçilme şansı eşit ise örnekleme rasgele seçilmiş örneklem denir. Aynı nedenle, örneklem rasgele seçilmiş ve n birimden oluşuyorsa n öğeli tüm olanaklı örneklemeler eşit seçilme şansına sahiptirler. N kitle büyüklüğü, n bu kitleden seçilecek basit rasgele örneklemdaki birim sayısını göstermek üzere bu kitleden*

$$C_{N,n} = \frac{N!}{n!(N-n)!}$$

sayıda örneklem seçeriz. Kitle homojen yapıda ise genellikle basit rasgele örneklem seçilir.

Örnek 1.16. *B, C, D, E harflerinin oluşturduğu bir kitleden $n = 2$ birimlik tüm örneklemeler $C_{4,2} = 4!/2!*2! = 6$ olup*

$$BC, BD, BE, CD, CE, DE$$

dir. . B , 3 örneklemin içinde yer alır. Böylece B nin örnekleme girme şansı

$$P(B) = \frac{3}{6} = \frac{1}{2}$$

bulunur. Benzer şekilde $P(C) = P(D) = P(E) = 1/2$ olup kitlenin herbir öğesinin seçilme şansı eşittir. Ayrıca 6 örneklemin her biri aynı şansa sahiptir. (Yani $P(BC) = \dots = P(DE) = 1/6$ dir. Böylece basit rasgele örneklemin seçimine ilişkin özelliklerinin sağlandığı görülmektedir.

- Örneklem seçiminde aşağıdaki rasgele sayılar tablosundan da yararlanılır.

RASGELE SAYILAR TABLO X

93108	77033	68325	10160	38667	62441	87023	94372	06164	30700
28271	08589	83279	48838	60935	70541	53814	95588	05832	80235
21841	35545	11148	34775	17308	88034	97765	35959	52843	44895
22025	79554	19698	25255	50283	94037	57463	92925	12042	91414
09210	20779	02994	02258	86978	85092	54052	18354	20914	28460
90552	71129	03621	20517	16908	06668	29916	51537	93658	29525
01130	06995	20258	10351	99248	51660	38861	49668	74742	47181
22604	56719	21784	68788	38358	59827	19270	99287	81193	43366
06690	01800	34272	65497	94891	14537	94358	21587	95765	72605
59809	69982	71809	64984	48709	43991	24987	69246	86400	29559
56475	02726	58511	95405	70293	84971	06676	44075	32338	31980
02730	34870	83209	03138	07715	31557	55242	61308	26507	06186
74482	33990	13509	92588	10462	76546	46097	01825	20153	36271
19793	22487	94238	81054	95488	23617	15539	94335	73822	93481
19020	27856	60526	24144	98021	60564	46373	86928	52135	74919
69565	60635	65709	77887	42766	86698	14004	94577	27936	47220
69274	23208	61035	84263	15034	28717	76146	22021	23779	98562
83658	14204	09445	41081	49630	34215	89806	40930	97194	21747
78612	51102	66826	40430	54072	62164	68977	95583	11765	81072
14980	74158	78219	38985	60838	82836	42777	85321	90463	11813
63172	28010	29405	91554	75195	51183	65805	87525	35952	83204
71167	37984	52737	06869	38122	95322	41356	19391	96787	64410
78530	56410	19195	34434	83712	50397	80920	15464	81350	18673
98324	03774	07573	67864	06497	20758	83454	22756	83959	96347
55793	30055	08373	32652	02654	75980	02095	87545	88815	80086
05674	34471	61967	91266	38814	44728	32455	17057	08339	93997
15643	22245	07592	22078	73628	60902	41561	54608	41023	98345
66750	19609	70358	03622	64898	82220	69304	46235	97332	64539
42320	74314	50222	82339	51564	42285	50482	98501	02245	88990
73752	73818	15470	04914	24936	65514	56633	72030	30856	85183
97546	02188	46373	21486	28221	08155	23486	66134	88799	49496
32569	52162	38444	42004	78011	16909	94194	79732	47114	23919
36048	93973	82596	28739	86985	58144	65007	08786	14826	04896
40555	36702	38965	56042	80023	28169	04174	65533	52718	55255
33597	47071	55618	51796	71027	46690	08002	45066	02870	60012
22828	96380	35883	15910	17211	43258	14056	55438	98148	35384
00631	95925	19324	31497	88118	06283	84596	72091	53987	01477
75722	36478	07634	63114	27164	15467	03983	09141	60562	65725
80577	01771	61510	17099	28731	41426	18853	41523	14914	76661
10524	20900	65463	83680	05005	11611	64426	59065	06758	02892
93815	69446	75253	51915	97839	75427	90685	60352	96288	34248
81867	97119	93446	20862	46591	97677	42704	13718	44975	67145
64649	07689	16711	12169	15238	74106	60655	56289	74166	78561
55768	09210	52439	33355	57884	36791	00853	49969	74814	09270
38080	49460	48137	61589	42742	92035	21766	19435	92579	27683
22360	16362	05343	34613	24013	98831	17157	44089	07366	66196
40521	09057	00239	51284	71556	22605	41293	54854	39736	05113
19292	69862	59951	49644	53486	28244	20714	56030	39292	45166
79504	40078	06838	05509	68581	29400	85615	52314	83202	40313
64138	27983	84048	42631	58658	62243	82572	45211	37060	15017

Örnek 1.17. 100 fare arasından kontrol ve ilaç grubu olarak düşünülen 5'er deneklik iki gruba fareler nasıl seçilir?

Çözüm: Önce hayvanlar 1 den 100'e kadar numaralanır. Sonra rasgele sayılar tablosu 1.1 den faydalananarak herhangi bir yer başlangıç noktası olarak belirlenir. Örneğin 3.sütunu seçelim. Tablodaki 5 basamaklı sayıların ilk iki rakam seçilerek kontrol gurubundaki fareyi temsil etmiş olur. Sonraki satırdaki ilk iki basamak ise ilaç gurubundaki fareyi temsil edecektir. Bu yöntemle 10 fare seçilene kadar devam ederiz. Eğer 00 çekilmişse 100 ü temsil edecektir ve aynı sayıya rastlarsak bir sonrakine geçeriz. Bu yöntemle seçilen 10 fare aşağıdaki gibi olur:

Kontrol Grubu	68	11	02	20	34
İlaç Grubu	83	19	03	21	71

□

Tanım 1.18. Bir başka örneklem seçimi de sistematik örneklem seçimidir. Bir sistematik rasgele örneklem bir kitlenin her i nci elemanını içeren bir örneklemidir

Örnek 1.19. Örneğin aşağıdaki gibi 20 elemanlı bir kitle olsun:

$$A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, R, S, T, U.$$

Her $i = 5$. eleman örnekleme yer alsın. İlk 5 elemandan biri ile başlarız ve bunları 1,2,3,4,5 ile numaralandırırız. Rasgele sayılar tablosu 1.1 dan bir başlangıç noktası seçeriz. Örneğin 1.satır ve 1.sütundaki 93108 sayısı ile başlayalım. Bu sayıdaki ilk rakam 9, numaralandırdığımız 1,2,3,4,5 den biri olmadığından kullanamayız. Sonraki rakama geçeriz yani 3. 3 sayısı uygundur ve harf olarak C'ye karşılık gelir. C'den sonra 5. harf olan H, ve sonraki 5.harf olan M, ve sonraki 5. harf olan S örnekleme girer.

- Görüldüğü gibi basit ve sistematik rasgele örneklem seçimi yöntemleri birbirinden farklıdır.
- Kitleler bir liste halinde ya da dosyada kayıtlı olarak verilmişse bu kitlelerden sistematik rasgele örneklem seçilir.

Tanım 1.20. Bir başka örneklem seçimi de Tabakalı Rasgele Örneklem Seçimi dir. Kitle öncelikle tabakalandırılır ve her bir tabakadan rasgele örneklem seçilir. Tüm kitleden doğrudan seçim yerine, her bir tabakadan seçilecek örneklem büyüklüğüne önceden karar verilir ve her birinden eleman alınması sağlanır. Genel bir kural olarak, tabakalı örnekleme rasgele örneklemeden daha iyi sonuç verdiği söylenir. Tabakalar arasındaki farklar tabakalar içindeki farklardan büyük ise tabakalı örnekleme basit rasgele örnekleme göre daha iyi sonuç verecektir. Tabakalı rasgele örnekleme insan kitlelerine ilişkin örneklem alırken yaygın olarak kullanılır. Örneklem; yaş, eğitim düzeyi, cinsiyet, gelir düzeyi, coğrafi bölge vs. gibi karakteristiklere göre tabakalandırılır.

Tanım 1.21. Bir başka örneklem seçimi de Rasgele Küme Örneklemidir. Burada önce kümeler seçilir, sonra her bir kümenin elemanlarının tümü ya da bir kısmı örnekleme alınır. Örneğin: Geniş bir sanayi bölgesinde fabrikalarda çalışan işçilerin çalışma koşullarına ilişkin davranışları hakkında bilgi sahibi olmak istediğinizi kabul edelim. Her bir fabrika, işçilerin bir kümesini oluşturur. Böylece fabrikaların bir basit rasgele örneklemini seğiriz. Sonra bu fabrikalarda çalışanlardan rasgele örneklem seçer ve onlarla görüşme yaparız. Bu yönteme iki aşamalı (adımlı) örnekleme denir. İlk aşama kümelerin bir örnekleminin seçilmesidir, ikinci

aşama ise seçilen her bir kümedeki elemanların bir örnekleminin seçimidir.

Ekonomik olması nedeni ile sık sık kullanılan bir örnekleme yöntemidir. Küme örneklemleri homojen yapıdaki kitlelerde çok iyi çalışır. Fakat, kitleler çoğu zaman homojen değildir. Bununla birlikte önce tabakalama yapılırsa tabakalar nispeten homojen yapıda olur. Sonuç olarak farklı tabakalara sahip bir kitleden küme örnekleme yapılacaksa, önce tabakalama yapılmalı sonra her bir tabakadan küme örneklemleri seçilmelidir. Örneğin, bir şehir için gelir dağılımı hakkında bilgi sahibi olmak istiyorsak, önce düşük, orta ve yüksek gelir gruplarına göre tabakalama yapılır. Sonra küme örneklemesine geçilir.

2 Veri Düzenlemesi ve Analizi

2.1 Giriş ve Ön bilgiler

Tanım 2.1. Bir araştırmacı tarafından gözlemlerden elde edilen sonuçlara veri denir.

Veri genellikle dört farklı yolla toplanır.

- (1) Yayınlanmış kaynaklardan,
- (2) Tasarlanmış bir denemeden
- (3) Anket sonuçlarından
- (4) Gözlem sonuçlarının toplanmasından

Tanım 2.2. İstatistiksel araştırma yapılırken, belli bir topluluğa ya da topluluğu oluşturan birimlere ilişkin bilgi toplamaya "derleme (veri toplama)" denir. Başka bir deyişle derleme ele alınan özelliğin (ya da özelliklerin) topluluğu oluşturan birimlerdeki türlü görünümünün saptanmasıdır. Örneğin, belirli bir zamanda, belli bir okulun, belli bir sınıftaki öğrencilerin boyları, ağırlıkları, cinsiyetleri, kardeş sayısı, baba mesleği üzerinde araştırma yapılacaksa ele alınan özelliklere "boy, ağırlık, kardeş sayısı, baba mesleği" ilişkin bilgi toplandığında derleme yapılmış olur.

Veri toplama bir istatistiksel araştırmanın en önemli aşamasından biridir. Bu nedenle en uygun veri toplama tekniğinin seçimi de önemlidir. Veri toplama kapsamı bakımından "tamsayım" ve "örnekleme" olmak üzere ikiye ayrılır.

Tanım 2.3. Kitleyi oluşturan tüm birimlerden veri elde etmek demektir. Örneğin, nüfus sayımı bir tamsayım sayıdır.

Tanım 2.4. Bir kitleden örneklem seçme sürecine örnekleme denir.

Tanım 2.5. Toplanan verilerin yararlılık derecesi üç faktöre bağlıdır

- (1) Maliyet: Bir verinin yararlı olması için öncelikle sağlayacağı fayda, verinin elde edilmesi için yapılan harcamadan büyük olmalıdır
- (2) Doğruluk taşınmalıdır
- (3) Hızlıdır.

Tanım 2.6. (Betimsel amaçlı istatistik) Kitledeki tüm birimlerden ilgili değişken ya da değişkenler için veri toplandığında bunları kullanarak kitleyi özetlemeyi (betimlemeyi) amaçlar. Bu ise frekans dağılımı oluşturularak, grafikler çizilerek ya da parametreler (kitle ortalaması ve varyansı) hesaplanarak yapılır.

Tanım 2.7. (Tümevarımsal amaçlı istatistik) Kitle(ler) den rastgele seçilen örneklem(ler) den toplanan verileri kullanarak kitle(ler) nin parametrelerini tahmin etmeyi veya parametrelerle ilgili olan savların doğru olup olmadığının araştırılmasını amaçlar. Yalnız örneklem gözlemlendiğinde belirsizlikler içeren kitle hakkında sonuçlar çıkarılmak istenebilir. Örneklemdeki gözlemlerden kitle hakkındaki önermelere geçmek için geliştirilen süreçle istatistiksel sonuç çıkarılır.

2.2 Veri Düzenlemesi

Veri organizasyonu, özetlenmesi ve istatistiksel analizi için yöntemler düşünüldüğünde, iki tip veri ele al-
mamız gerekecektir.

Ağırlık, uzaklık, zaman veya bireylerin veya kusurların sayısı gibi değerlerden oluşan nümerik ölçümlere sayısal (nicel) veri denir. Bunlar doğal sayısal ölçek üzerinde kaydedilen ölçümlerdir. Nitelik belirleme amacına yönelik gözlemlerin sonucunda ise kategorik (nitel) veri elde edilir. Bunlar doğal sayısal ölçek üzerinde ölçülemeyen ve kategorilerin bir grubu içine sınıflandırılan ölçümlerdir. Örneğin Jeoloji ile ilgili bir araştırmada kaya türleri inceleniyorsa, nitelikler göz önüne alınarak sınıflandırma yapılır. İnceleme konusu kitlenin bir bireyden diğerine değişkenlik gösteren özellikler bize niteliksel veri verecektir.

Birimlerin farklı değerler alabildikleri nitelik veya niceliklerine değişken denir.

Değişkenleri aşağıdaki gibi sınıflandırabiliriz.

- (1) Ölçme ve sayma değişkenleri: (Nümerik değişkenler)
 - Sürekli değerler alan değişkenler (Yükseklik, ağırlık, uzunluk, zaman, sıcaklık, hacim vs.). Belli bir aralıktaki her değeri alabilen değişkenler.
 - Kesikli değer alan değişkenler (Satılan birimlerin sayısı, giriş hatalarının sayısı, reddetme sayısı, üretilen mal sayısı, ailedeki çocuk sayısı, farklı günlerde bir fabrikadaki kaza sayısı). Tam değerler alabilen değişkenler.
- (2) Sıralama değişkenleri: Bazi değişkenlere ait değerler ölçümle belirlenmemekle beraber büyüklükleri itibarıyla sıralanabilirler (1., 2., 3., ...vs. sıralama işlemi yapılır.)
- (3) Özellik belirten değişkenler: Özellikle kalite ve durum gibi nitel özellik belirten değişkenleri bu sınıfa koyabiliriz. Örneğin, siyah-beyaz, ölü-canlı, erkek-dişi, evli-bekar, öğrenim durumu-meslek gibi nitel bir özellik söz konusu olduğunda gözlenen durumlara dayanarak dağılım düzenlenir.

Örnek 2.8. 100 öğrencinin cinsiyetine göre dağılımı

Cinsiyet	Gözlem Sayısı	Oran
Kız	48	48/100
Erkek	52	52/100
Toplam	100	1

Örnek 2.9. 60 kobay renkleri itibari ile siyah, alaca ve gri olarak ayrılıyorsa aşağıdaki gibi bir dağılım elde ederiz.

Renk	Gözlem Sayısı
Siyah	4
Alaca	2
Gri	54
Toplam	60

2.3 Frekans (Sıklık) Dağılımı

Derlenen veri üzerinde herhangi bir işlem yapılmamışsa, bunlara "ham veri" ya da sınıflandırılmamış (gruplandırılmamış) veri denir. Birim sayısı az olan yığınların çeşitli özellikleri ham veriye dayanarak kolaylıkla belirtilir. Yığın çok sayıda birimden oluşuyorsa, bunları sınıflandırmakla yığının çeşitli özelliklerini belirlemek kolaylaşacaktır. Sınıflandırmanın en doğru yolu frekans tablosudur. Burada, gözlenen veri sınıflara ayrılır. Sonuçlanan tablo her bir sınıftaki gözlem sayısını verecektir.

Frekans tablosunu hazırlamak için aşağıdaki adımlar izlenir:

- I. Adım: Gözlemlerin sayısı belirlenir. n = Gözlem sayısını gösterir ($50 \leq n$ tercih edilir).
- II. Adım: En büyük değer (L) ve en küçük değer (S) bulunur. (L-S) farkı hesaplanır. R ile gösterilen bu farka değişim genişliği (range) denir.
- III. Adım: Sınıf sayısı bulunur. $\sqrt{n} < k$ (k=tamsayı), sınıf sayısı k ile gösterilir. Kesin bir yaklaşım olmamakla birlikte sınıf sayısı için $k = 1 + 3.3 * \log n$ formülü de kullanılabilir. Genel olarak $5 \leq k \leq 20$ seçilmesi benimsenir.
- IV. Adım: Sınıf genişliği: Ardışık iki sınıfın alt (ya da üst) sınıf limitleri arasındaki farka dağılım için sınıf genişliği denir. Baska bir sınıflamaya gerek duyulmadıkça sınıf genişliklerini eşit alma yoluna gidilir. Sınıf genişliği = h olmak üzere $\frac{R}{k} = \frac{L-S}{k} \leq h$ alınır. Özellikle tamsayı olan gözlemler için h nin tamsayı olması istenir.
- V. Adım: Sınıf limitleri: Frekans dağılımında sınıfları belirlemek için kullanılan sayılardır. En küçük gözlem değerine eşit ya da daha küçük olarak ilk sınıfın alt limiti seçilir. Bu değere ardışık olarak sınıf genişliğinin eklenmesiyle diğer sınıfların alt limitleri bulunur. Verinin sürekli ya da kesikli olmasına göre ilk sınıftan başlayarak üst limitler de aynı yolla bulunur.
- VI. Adım: Sınıf sınırları belirlenir. $\frac{i. \text{ sınıfın üst limiti} + \text{bir sonraki sınıfın alt limiti}}{2} = i. \text{ sınıfın üst sınıf sınırı}$
- VII. Adım: Sınıf frekansı: Her sınıf için -sınıf limitleri dahil -o sınıfa düşen gözlem sayısı (sınıf frekansı) saptanır. Frekanslar $f_1, f_2, f_3, \dots, f_k$ ile gösteririz.
- VIII. Adım: Sınıf orta noktası: Sınıf limitlerinin ya da sınıf sınırlarının ortalaması alınarak her sınıf için sınıf orta noktaları (SINIF ORTASI) bulunur. Sınıf ortalarını y_1, y_2, \dots, y_k ile gösteririz. İlk sınıfın sınıf ortası bulunduktan sonra ardışık olarak sınıf genişliğinin eklenmesiyle diğer sınıfların sınıf ortaları bulunur.
- IX. Adım: Frekanslar sütun olarak yazdıktan sonra toplamaları alınarak kontrol edilir. $f_1 + f_2 + f_3 + \dots + f_k = n$ olmalıdır.

- X. Adım: Eklemeli frekans: Ardışık olarak frekansları $f_1, f_1 + f_2, f_1 + f_2 + f_3, \dots$ biçiminde toplayarak eklemeli frekans sütunu oluşturulur.

Unutulmaması Gereken Önemli Notlar:

- (1) Sınıflar ayrık olmalıdır. Yani her veri yalnız bir sınıfta yer almalıdır.

$$4,00 - 4,50$$

$$4,50 - 5,00$$

gibi sınıflar varsa I. aralığın üst limiti, II.nin alt limiti aynı sayıdır. Bu iki sınıf ayrık değildir. O halde aralıklar:

$$4,00 - 4,49$$

$$4,50 - 4,99$$

ya da

$$4,00 - 4,50 \text{ den az}$$

$$4,50 - 5,00 \text{ den az}$$

biçiminde yazılmalıdır.

- (2) Sınıflar (aralıklar) veri kümesinde yer alan her değeri içine almalıdır.
- (3) Sınıf limitleri ve sınıf sayısı hiç bir sınıf boş kalmayacak şekilde (0 frekanslı sınıf olmayacak) seçilmelidir

Örnek 2.10. *Büyük bir şirketin 100 satış. elemanı vardır. Her bir satış elemanı tarafından gerçekleştirilen aylık satışlar aşağıda verilmiştir. Bu veriler için frekans tablosu düzenleyiniz.*

Milyon Lira				
23	16	14	20	27
19	17	17	16	17
26	14	9	11	14
11	17	13	19	17
20	17	20	16	16
11	24	21	27	5
17	20	8	16	17
16	16	14	22	13
14	27	19	16	20
16	15	9	17	8
19	14	8	19	27
22	21	0	9	3
20	14	6	11	12
7	20	9	13	20
10	16	10	19	13
15	15	14	13	25
14	9	16	8	16
7	8	13	5	13
9	16	19	14	29
18	14	18	13	10

Çözüm: 1.Adım: $n = 20 * 5 = 100$

2. Adım: (En büyük değer) $L = 29$, (En küçük değer) $S = 0$, (range) $R = L - S = 29$

3. Adım: $k \geq \sqrt{n} = \sqrt{100} \Rightarrow k \geq 10$, $k = 10$ alınabilir.

4.Adım: $h \geq \frac{R}{k} = \frac{29}{10} = 2.9$ tamsayı tercih edildiği için $h = 3$ alınabilir.

V.VI, VII, VIII, IX, ve X. adımlar; uygulandığında aşağıdaki tablo elde edilir.

Sınıf No	Sınıf Limitleri	Sınıf Sınırları	Gözlem Sayısı (f_i)	Sınıf Orta Noktası(y_i)	Ekllemeli Frekanslar
1	0 – 2	-0.5 ~ 2.5	1	$\frac{0+2}{2} = 1$	1
2	3 – 5	2.5 ~ 5.5	3	$\frac{3+5}{2} = 4$	1 + 3 = 4
3	6 – 8	5.5 ~ 8.5	8	$\frac{6+8}{2} = 7$	$\underbrace{1 + 3}_{=4} + 8 = 12$
4	9 – 11	8.5 ~ 11.5	13	$\frac{9+11}{2} = 10$	12 + 13 = 25
5	12 – 14	11.5 ~ 14.5	20	$\frac{12+14}{2} = 13$	25 + 20 = 45
6	15 – 17	14.5 ~ 17.5	25	$\frac{15+17}{2} = 16$	45 + 25 = 70
7	18 – 20	17.5 ~ 20.5	17	$\frac{18+20}{2} = 19$	70 + 17 = 87
8	21 – 23	20.5 ~ 23.5	5	$\frac{21+23}{2} = 22$	87 + 5 = 92
9	24 – 26	23.5 ~ 26.5	3	$\frac{24+26}{2} = 25$	92 + 3 = 95
10	27 – 29	26.5 ~ 29.5	5	$\frac{27+29}{2} = 28$	95 + 5 = 100

□

Uyarılar:

- (1) Çok sayıda sınıfın seriyi gereksiz yere uzatabileceği, az sayıda sınıfın bilgi yitirilmesine yol açabileceği düşünülerek uygun sayıda sınıf seçilmesine özen gösterilmelidir.
- (2) Gözlenen değerler büyüklük sırasına göre düzenlendiğinde bazen küçük değerler veya büyük değerlerin çok seyrek oldukları ve aralarındaki farkların büyük olmaları nedeniyle eşit sınıf aralıkları yerine zorunlu olarak değişik genişliğe sahip sınıflar düzenlenebilir. Bazı durumlarda da ilk ve/veya son sınıflar için açık uçlu aralıklar düşünülebilir.

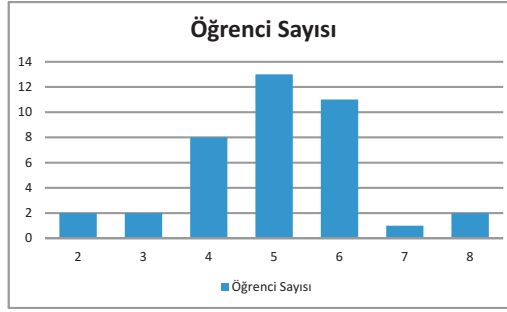
3 Grafiksel Gösterimler

3.1 Çizgisel Grafik

Tanım 3.1. Sınıflandırılmamış gözlemlerin dağılımının dik koordinat sistemindeki gösterimine "Çizgisel grafik" denir.

Örnek 3.2. Bir sınavda 39 öğrencinin aldığı notlar aşağıdaki gibi olsun. Bu notlar için çizgisel grafik çiziniz.

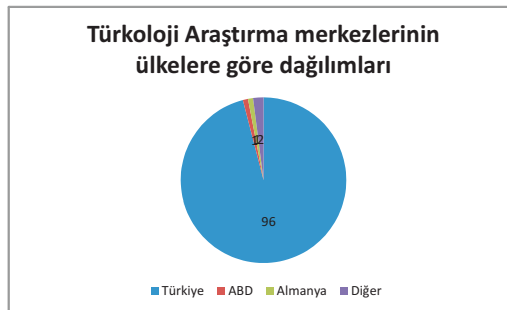
Notlar (X_i)	Öğrenci Sayısı (Y_i)
2	2
3	2
4	8
5	13
6	11
7	1
8	2



3.2 Dairesel Grafik

Örnek 3.3. Aşağıda Türkoloji Araştırma merkezlerinin ülkelere göre dağılımları verilmektedir. Bu dağılım için dairesel grafiği çiziniz.

Ülkeler	Merkezlerin Yüzdesi (%)
Türkiye	96
ABD	1
Almanya	1
Diğer	2

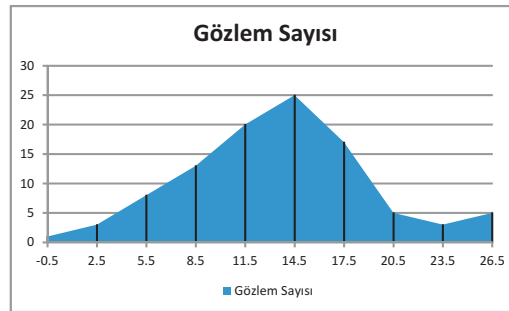


3.3 Poligon (Frekans Çokgeni)

Tanım 3.4. Bir dik koordinat sisteminde yatay ekseninde sınıf orta noktaları, dikey ekseninde sınıf frekansları işaretlenir. Bu sayı çiftlerine karşılık gelen noktaların doğru parçalarıyla birleştirilmesiyle oluşan grafiksel gösterime "Poligon ya da frekans çokgeni" denir.

Örnek 3.5.

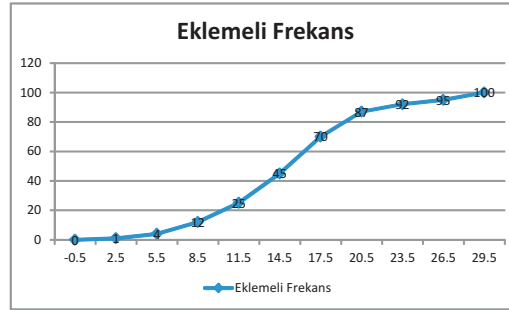
<i>Sınıf No</i>	<i>Sınıf Limitleri</i>	<i>Sınıf Sınırları</i>	<i>Gözlem Sayısı (f_i)</i>	<i>Sınıf Orta Noktası (y_i)</i>	<i>Eklemeli Frekanslar</i>
1	0 – 2	-0.5 ~ 2.5	1	$\frac{0+2}{2} = 1$	1
2	3 – 5	2.5 ~ 5.5	3	$\frac{3+5}{2} = 4$	4
3	6 – 8	5.5 ~ 8.5	8	$\frac{6+8}{2} = 7$	12
4	9 – 11	8.5 ~ 11.5	13	$\frac{9+11}{2} = 10$	25
5	12 – 14	11.5 ~ 14.5	20	$\frac{12+14}{2} = 13$	45
6	15 – 17	14.5 ~ 17.5	25	$\frac{15+17}{2} = 16$	70
7	18 – 20	17.5 ~ 20.5	17	$\frac{18+20}{2} = 19$	87
8	21 – 23	20.5 ~ 23.5	5	$\frac{21+23}{2} = 22$	92
9	24 – 26	23.5 ~ 26.5	3	$\frac{24+26}{2} = 25$	95
10	27 – 29	26.5 ~ 29.5	5	$\frac{27+29}{2} = 28$	100

**3.4 Eklemeli Frekans Çokgeni**

Tanım 3.6. Bir dik koordinat sisteminde yatay eksenle sınıf sınırları ve dikey eksenle eklemeli frekanslar işaretlenir. Her bir sınıftaki eklemeli frekanslara karşılık gelen noktaların birleştirilmesiyle oluşan grafiksel gösterime "Eklemeli frekans çokgeni" denir.

Örnek 3.7.

<i>Sınıf No</i>	<i>Sınıf Limitleri</i>	<i>Sınıf Sınırları</i>	<i>Gözlem Sayısı (f_i)</i>	<i>Sınıf Orta Noktası (y_i)</i>	<i>Eklemele Frekanslar</i>
1	0 – 2	-0.5 ~ 2.5	1	$\frac{0+2}{2} = 1$	1
2	3 – 5	2.5 ~ 5.5	3	$\frac{3+5}{2} = 4$	4
3	6 – 8	5.5 ~ 8.5	8	$\frac{6+8}{2} = 7$	12
4	9 – 11	8.5 ~ 11.5	13	$\frac{9+11}{2} = 10$	25
5	12 – 14	11.5 ~ 14.5	20	$\frac{12+14}{2} = 13$	45
6	15 – 17	14.5 ~ 17.5	25	$\frac{15+17}{2} = 16$	70
7	18 – 20	17.5 ~ 20.5	17	$\frac{18+20}{2} = 19$	87
8	21 – 23	20.5 ~ 23.5	5	$\frac{21+23}{2} = 22$	92
9	24 – 26	23.5 ~ 26.5	3	$\frac{24+26}{2} = 25$	95
10	27 – 29	26.5 ~ 29.5	5	$\frac{27+29}{2} = 28$	100



4 Merkezsel Eğilim Ölçüleri

Gözlenen verinin düzenlenerek çizelgelerle, grafiklerle sunulması çoğu kez yeterli olmaz. Genel durumu yansıtacak bir takım ölçülere gereksinim vardır. Öyle ölçüler ki, yalnızca verileri özlü bir biçimde belirtmekle kalmasınlar, yapılacak karşılaştırmalara genellemelere, yorumlara olanak sağlasınlar. Bu başlık altında nicel değişkenlere ilişkin ölçüleri inceleyeceğiz. Nicel dağılımlarda kullanılacak ölçüler dağılımın odaklaşma noktasını özetlemelidir. Bu tür ölçülere merkezi (merkezsel) eğilim ölçüleri denir. Bunlar sıra ile şöyledir:

- Aritmetik ortalama
- Medyan (Orta Deger)
- Mod (Tepe Deger)
- Geometrik Ortalama
- Harmonik Ortalama

4.1 Aritmetik Ortalama

Örnek 4.1. Gözlenen değerlerin tümü toplanarak gözlem sayısına bölündüğünde elde edilen değere aritmetik ortalama denir. Değişkeni X diye adlandırılıp, aldığı değerleri x_1, x_2, \dots, x_n ile gösterirsek aritmetik ortalama:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

Ya da \sum "toplam" gösterimi ile;

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

dir. Sonlu kitledeki eleman sayısı genellikle N ile örneklemdeki eleman sayısı n ile gösterilir. O halde kitle için aritmetik ortalama

$$\mu = \frac{\sum_{i=1}^N x_i}{N}$$

Bir değişkenin aldığı değerler x_1, x_2, \dots, x_k ve bu değerlerin her birinin tekrarlanma sayısı da f_1, f_2, \dots, f_k ise aritmetik ortalama :

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + \dots + x_k f_k}{f_1 + f_2 + \dots + f_k} = \frac{x_1 f_1 + x_2 f_2 + \dots + x_k f_k}{n}, n = \sum_{i=1}^k f_i$$

i -inci değer w_i olmak üzere, aritmetik ortalama;

$$\bar{x} = \frac{x_1 w_1 + x_2 w_2 + \dots + x_k w_k}{w_1 + w_2 + \dots + w_k}, n = \sum_{i=1}^k w_i$$

Sınıflara ayrılmış aritmetik ortalamayı hesaplamak için

n = toplam gözlemlerin sayısı

y_i = i . sınıfın orta değeri

k = sınıf sayısı

f_i = i . sınıfın frekansı

$$\bar{x} = \frac{y_1 f_1 + y_2 f_2 + \dots + y_k f_k}{f_1 + f_2 + \dots + f_k} = \frac{y_1 f_1 + y_2 f_2 + \dots + y_k f_k}{n}$$

Örnek 4.2. Bir öğrencinin sınavına girdiği derslerin önem dereceleri (ağırlıklar) farklı ise not ortalaması hesabında derslerin önem derecelerinin göz önüne alınması gerekir. Kabul edelim ki bir öğrenci için aşağıdaki

veriler vardır. Buna göre aritmetik ortalamayı hesaplayınız.

Not	Ağırlık	Puan	Ağırlıklı Puan
A	3	30	$3 * 30 = 90$
B	2	45	$2 * 45 = 90$
C	2	25	$2 * 25 = 50$
D	1	40	$1 * 40 = 40$
E	4	15	$4 * 15 = 60$
F	2	6	$2 * 6 = 12$
G	2	55	$2 * 55 = 110$
Toplam:			452

$$\bar{x} = \frac{452}{3 + 2 + 2 + 1 + 4 + 2 + 2} = 28.25$$

Örnek 4.3. Aşağıda verilen tablo için ayrılmış aritmetik ortalamayı hesaplayınız.

Sınıf No	Sınıf Limitleri	Sınıf Sınırları	Gözlem Sayısı (f_i)	Sınıf Orta Noktası(y_i)	Eklemeli Frekanslar
1	0 – 2	-0.5 ~ 2.5	1	$\frac{0+2}{2} = 1$	1
2	3 – 5	2.5 ~ 5.5	3	$\frac{3+5}{2} = 4$	4
3	6 – 8	5.5 ~ 8.5	8	$\frac{6+8}{2} = 7$	12
4	9 – 11	8.5 ~ 11.5	13	$\frac{9+11}{2} = 10$	25
5	12 – 14	11.5 ~ 14.5	20	$\frac{12+14}{2} = 13$	45
6	15 – 17	14.5 ~ 17.5	25	$\frac{15+17}{2} = 16$	70
7	18 – 20	17.5 ~ 20.5	17	$\frac{18+20}{2} = 19$	87
8	21 – 23	20.5 ~ 23.5	5	$\frac{21+23}{2} = 22$	92
9	24 – 26	23.5 ~ 26.5	3	$\frac{24+26}{2} = 25$	95
10	27 – 29	26.5 ~ 29.5	5	$\frac{27+29}{2} = 28$	100

Çözüm:

$$\begin{aligned}
 \bar{x} &= \frac{y_1 f_1 + y_2 f_2 + \dots + y_k f_k}{f_1 + f_2 + \dots + f_k} \\
 &= \frac{\left(\begin{aligned} &1 * 1 + 4 * 3 + 7 * 8 + 10 * 13 + 13 * 20 \\ &+ 16 * 25 + 19 * 17 + 22 * 5 + 25 * 3 + 28 * 5 \end{aligned} \right)}{100} \\
 &= \frac{1507}{100} = 15.07
 \end{aligned}$$

□

- Ortalamalar genellikle iki veri grubu arasındaki farkı belirlemek için kullanılır. Ayrıca standart belirlemek amacıyla da kullanılır. Nasıl kullanılırlarsa kullanılırsınlar sapan değerlerin etkisini gözleyebiliriz. Örneğin,

5.30, 0.10, 5.60, 5.80, 5.50

sayıları verilsin. Ortalama 4.46 dir. Burada 0.10 sapan değerdir. Bu değer ortalamayı aşağı çekmektedir. Bir norm/standart belirlemede ya da karşılaştırmada sapan değer çıkarıldıktan sonra bulunan ortalama daha anlamlı olacaktır, Sapan değer olmasaydı ortalama 5.55 olurdu. Eğer bir sapan değer hesaplama hatasından meydana geliyorsa ortalamayı hesaplamaadan önce bu değer düzeltilmelidir.

Teorem 4.4. İki grup gözleme sahip olalım. 1. grupta örneklemdaki eleman sayısı n ve gözlemsel değerler $x_{11}, x_{12}, \dots, x_{1n}$ olmak üzere aritmetik ortalaması

$$\bar{x}_1 = \frac{x_{11} + x_{12} + \dots + x_{1n}}{n}$$

1. grupta örneklemdaki eleman sayısı m ve gözlemsel değerler $x_{21}, x_{22}, \dots, x_{2m}$ olmak üzere aritmetik ortalaması

$$\bar{x}_2 = \frac{x_{21} + x_{22} + \dots + x_{2m}}{m}$$

olmak üzere $n + m$ gözlemin ortlaması

$$\bar{x} = \frac{n\bar{x}_1 + m\bar{x}_2}{n + m}$$

ile verilir. Bu sonuç ortalamaların ağırlıklı aritmetik ortalamasıdır. Bunu genleştirirsek, her birinde n_1, n_2, \dots, n_k gözlem bulunun k grubun aritmetik ortlamaları sırasıyla $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k$ olmak üzere ortalamaların ağırlıklı aritmetik ortalaması

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2 + \dots + n_k\bar{x}_k}{n_1 + n_2 + \dots + n_k}$$

ile hesaplanır.

Örnek 4.5. Bir firma tarafından alınan 200 küçük siparişin her birinde ortalama kâr 2 birim lira, 50 büyük siparişin her birinde ortalama kâr 4.5 birim liradır.

a) 250 siparişin tümü için her bir siparişteki ortalama kârı bulunuz.

b) Büyük siparişlerin ortalama kârı sabit tutularak 250 sipariş için ortalama kârın 3 birim lira olması istenirse küçük siparişlerdeki ortalama kâr ne olmalıdır?

Çözüm: a)

$$\bar{x} = \frac{200 * 2 + 50 * 4.5}{250} = 2.5$$

b)

$$3 = \bar{x} = \frac{200 * \bar{x}_1 + 50 * 4.5}{250} \Rightarrow \bar{x}_1 = \frac{3 * 250 - 50 * 4.5}{200} = 2.625$$

□

4.2 Medyan (Orta değer = Ortanca)

Bazı durumlarda bir yarısı kendisinden büyük diğer yarısı daha küçük olan değer gözlemler grubunun merkezi olarak uygun olabilir.

Tanım 4.6. Büyüklüklerine göre sınırlanmış gözlemler grubunun merkezi değerine medyan denir. Çift sayıda gözlem varsa orta değer (medyan) iki merkezi değer aritmetik ortalamasıdır. Matematiksel olarak x_1, x_2, \dots, x_n

gözlem değerleri büyüklüklerine göre artan sırada düzenlenmişse, M ile gösterilen medyan aşağıdaki gibi hesaplanır.

$$M = \begin{cases} x_{\frac{n+1}{2}} & , n \text{ tek ise} \\ (x_{\frac{n}{2}} + x_{\frac{n}{2}+1})/2 & , n \text{ çift ise} \end{cases}$$

Gruplandırılmış (sınıflara ayrılmış) gözlemler için medyan

$n =$ Gözlem sayısı

$L_m =$ Medyan sınıfın alt sınıf sınırı

$h =$ Medyan Sınıfının genişliği

$f_m =$ Medyan sınıfının frekansı

$n_m =$ Medyan sınıfından önceki sınıfların frekansları toplamı

$$M = L_m + \frac{(\frac{n}{2} - n_m) * h}{f_m}$$

ile hesaplanır. $f_1 + f_2 + \dots + f_i \geq \frac{n}{2}$ olan ilk sınıfa medyan sınıfı denir

Örnek 4.7. 2, 3, 2, 4, 4, 6, 6, 5, 8, 8, 9 sayıları için orta değeri bulunuz.

Çözüm: Öncelikle sayıları küçükten büyüğe doğru sıralayacağız. $n = 11$ tek sayıdır. Buna göre

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}
2	2	3	4	4	5	6	6	8	8	9

$$M = x_{\frac{11+1}{2}} = x_6 = 5$$

□

Örnek 4.8. 1, 2, 3, 3, 5, 5, 5, 6, 7, 7, 7, 8, 9, 9 biçiminde sıralanmış veriler için orta değeri bulunuz.

Çözüm: $n = 14$ çift sayıdır. Buna göre

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}
1	2	3	3	5	5	5	6	7	7	7	8	9	9

$$M = x_{\frac{14}{2}} + x_{\frac{14}{2}+1} = x_7 + x_8 = \frac{5+6}{2} = 5.5$$

□

Örnek 4.9. Aşağıdaki tablo için medyanı hesaplayınız.

Sınıf No	Sınıf Limitleri	Sınıf Sınırları	Gözlem Sayısı (f_i)	Sınıf Orta Noktası(y_i)	Eklemeli Frekanslar
1	0 – 2	-0.5 ~ 2.5	1	$\frac{0+2}{2} = 1$	1
2	3 – 5	2.5 ~ 5.5	3	$\frac{3+5}{2} = 4$	4
3	6 – 8	5.5 ~ 8.5	8	$\frac{6+8}{2} = 7$	12
4	9 – 11	8.5 ~ 11.5	13	$\frac{9+11}{2} = 10$	25
5	12 – 14	11.5 ~ 14.5	20	$\frac{12+14}{2} = 13$	45
6	15 – 17	14.5 ~ 17.5	25	$\frac{15+17}{2} = 16$	70
7	18 – 20	17.5 ~ 20.5	17	$\frac{18+20}{2} = 19$	87
8	21 – 23	20.5 ~ 23.5	5	$\frac{21+23}{2} = 22$	92
9	24 – 26	23.5 ~ 26.5	3	$\frac{24+26}{2} = 25$	95
10	27 – 29	26.5 ~ 29.5	5	$\frac{27+29}{2} = 28$	100

Çözüm:

$$n = 100, \frac{n}{2} = 50$$

$$f_1 + f_2 + \dots + f_5 = 45 \leq 50$$

$$f_1 + f_2 + \dots + f_6 = 70 \geq 50$$

olduğundan 6. sınıf medyan sınıfıdır. $L_6 = 14.5, n_6 = f_1 + f_2 + \dots + f_5 = 45, f_6 = 25, h = 3$ olduğundan

$$\begin{aligned}
 M &= L_m + \frac{\left(\frac{n}{2} - n_m\right) * h}{f_m} \\
 &= 14.5 + \frac{(50 - 45) * 3}{25} \\
 &= 15.1
 \end{aligned}$$

□

Örnek 4.10. Bir mağazada bir aylık süre içinde satış tutarları ve alış verişi yapan müşteri sayısı aşağıdaki gibi sınıflandırılmıştır. Sınıflara ayrılmış verileri kullanarak;

a) Satış tutarlarının aritmetik ortalamasını bulunuz.

b) Satış tutarlarının orta değerini bulunuz.

Satış Tutarı	Müşteri Sayısı	Sınıf Ortaları
0 – 99	200	49.5
100 – 199	75	149.5
200 – 299	40	249.5
300 – 399	25	349.5
400 – 499	10	449.5
500 – 999	5	749.5
1000 – 4999	15	2999.5
Toplam	370	

Çözüm: a)

$$\begin{aligned}\bar{x} &= \frac{200 * 49.5 + 75 * 149.5 + 40 * 249.5 + 25 * 349.5 + 10 * 449.5 + 5 * 749.5 + 15 * 2999.5}{370} \\ &= 251.53\end{aligned}$$

b)

sınırlar	Satış Tutarı	Müşteri Sayısı	Sınıf Ortaları
-0.5 – 99.5	0 – 99	200	49.5
99.5 – 199.5	100 – 199	75	149.5
199.5 – 299.5	200 – 299	40	249.5
299.5 – 399.5	300 – 399	25	349.5
399.5 – 499.5	400 – 499	10	449.5
499.5 – 999.5	500 – 999	5	749.5
999.5 – 4999.5	1000 – 4999	15	2999.5
	Toplam	370	

$$n = 370,$$

$\frac{n}{2} = \frac{370}{2} = 185$ olduğundan ilk satırdaki (200 müşteri sayılı) sınıfa düşecektir. İlk sınıf için alt sınır $L_1 = -0.5, n_1 = 0, h = 100, f_1 = 200$ olduğundan

$$\begin{aligned}M &= L_m + \frac{\left(\frac{n}{2} - n_m\right) * h}{f_m} \\ &= -0.5 + \frac{(185 - 0) * 100}{200} \\ &= 92.0\end{aligned}$$

□

4.3 Mod (Tepe Değer)

Tanım 4.11. Bir veri grubunda en çok tekrarlanan değere mod denir. Her değer yalnız bir kez elde edilmişse mod yoktur. Sınıflara ayrılmış gözlemlerde mod un hesaplanabilmesi için mod sınıfını saptamak gerekir. Mod

sınıfı ise en çok gözlem sayısına sahip olan sınıftır.

$$L_m = \text{Mod sınıfın alt sınıf sınırı}$$

$$\Delta_1 = \text{Mod sınıfın frekansı ile bir önceki sınıfın frekansı}$$

$$\Delta_2 = \text{Mod sınıfın frekansı ile bir sonraki sınıfın frekansı}$$

$$h = \text{Mod Sınıfının genişliği}$$

$$\text{Mod} = L_m + \frac{\Delta_1 * h}{\Delta_1 + \Delta_2}$$

Örnek 4.12. 3, 5, 6, 9, 9, 9, 10, 12 degerleri için mod 9 dur.

Örnek 4.13. 1, 3, 4, 4, 4, 5, 6, 8, 8, 8, 10, 12 degerleri için 4 ve 8 gibi iki mod değeri vardır.

Örnek 4.14. Aşağıdaki tablo için mod değerini hesaplayınız.

Sınıf No	Sınıf Limitleri	Sınıf Sınırları	Gözlem Sayısı (f_i)	Sınıf Orta Noktası(y_i)	Eklemeli Frekanslar
1	0 – 2	-0.5 ~ 2.5	1	$\frac{0+2}{2} = 1$	1
2	3 – 5	2.5 ~ 5.5	3	$\frac{3+5}{2} = 4$	4
3	6 – 8	5.5 ~ 8.5	8	$\frac{6+8}{2} = 7$	12
4	9 – 11	8.5 ~ 11.5	13	$\frac{9+11}{2} = 10$	25
5	12 – 14	11.5 ~ 14.5	20	$\frac{12+14}{2} = 13$	45
6	15 – 17	14.5 ~ 17.5	25	$\frac{15+17}{2} = 16$	70
7	18 – 20	17.5 ~ 20.5	17	$\frac{18+20}{2} = 19$	87
8	21 – 23	20.5 ~ 23.5	5	$\frac{21+23}{2} = 22$	92
9	24 – 26	23.5 ~ 26.5	3	$\frac{24+26}{2} = 25$	95
10	27 – 29	26.5 ~ 29.5	5	$\frac{27+29}{2} = 28$	100

Çözüm: Mod sınıfı 25 gözleme sahip 6.sınıftır.

$$L_6 = 14.5$$

$$\Delta_1 = 25 - 20 = 5$$

$$\Delta_2 = 25 - 17 = 8$$

$$h = 3$$

$$\begin{aligned} \text{Mod} &= L_m + \frac{\Delta_1 * h}{\Delta_1 + \Delta_2} \\ &= 14.5 + \frac{5 * 3}{5 + 8} \\ &= 15.654 \end{aligned}$$

□

Not: Frekans dağılımında mod hesaplanırken sınıf aralıkları eşit tutulmalıdır. Sınıf genişlikleri eşit değilse mod sınıfının hesaplanmasında dikkatli davranmak gereklidir. Çünkü en büyük frekans geniş aralıklı bir sınıfa

ait olabilir. Eğer sınıf genişlikleri eşitlenirse mod sınıfı değişebilir. Sınıf aralıkları eşit iken en büyük frekansa sahip sınıf mod sınıfıdır.

4.4 Aritmetik Ortalama, Orta Değer ve Mod'un özellikleri

4.4.1 Aritmetik Ortalama

- (1) Anormal değerler (uç değerler) aritmetik ortalamayı büyük ölçüde etkiler. Araştırmacıyı yanlış yargıya götürebilir.
- (2) x_1, x_2, \dots, x_n değerlerinin aritmetik ortalaması \bar{x} , herhangi bir değer x^* olmak üzere

$$\sum_{i=1}^n (x_i - \bar{x})^2 \leq \sum_{i=1}^n (x_i - x^*)^2$$

4.4.2 Orta Değer

- (1) Büyüklük sırasında karşılık gelen yer yönünden bir ortalamadır.
- (2) Birey sayısı ile etkilenir, anormal değerler tarafından etkilenmez.
- (3) $x_1 \leq x_2 \leq \dots \leq x_n$ değerlerinin orta değeri \hat{x} , herhangi bir değer x^* olmak üzere

$$\sum_{i=1}^n (x_i - \hat{x})^2 \leq \sum_{i=1}^n (x_i - x^*)^2$$

- (4) Büyüklüklerine göre sıralanmış değerlerin çoğunluğu merkeze yakın ise orta değer iyi bir merkezi eğilim ölçüsüdür.
- (5) Rasgele seçilen bir değer dizi içinde orta değerden büyük ya da küçük olması olasılığı 1/2 dir.

4.4.3 Mod

- (1) Anormal değerlerin etkisinden bağımsızdır.
- (2) Gözlem sayısının az olması halinde değerler tekrarlanmayabilir. Bu nedenle tepe değer belirlenemeyebilir.

4.5 Geometrik Ortalama

Tanım 4.15. n tane x_1, x_2, \dots, x_n değerlerinin çarpımının n . kökü geometrik ortalama olarak tanımlanır.

$$G.O = \sqrt[n]{x_1 * x_2 * \dots * x_n}$$

Ayrıca her iki yanın logaritması da alınırsa

$$\begin{aligned} \log(G.O) &= \frac{1}{n} \log(x_1 * x_2 * \dots * x_n) \\ &= \frac{1}{n} (\log x_1 + \log x_2 + \dots + \log x_n) \end{aligned}$$

elde edilir. Sınıflandırılmış frekans dağılımında

$$G.O = \sqrt[n]{x_1^{f_1} * x_2^{f_2} * \dots * x_n^{f_n}}$$

Sınıflandırılmış frekans dağılımında y_1, y_2, \dots, y_n sınıf ortaları olmak üzere

$$G.O = \sqrt[n]{y_1^{f_1} * y_2^{f_2} * \dots * y_n^{f_n}}$$

ile hesaplanır.

4.6 Harmonik Ortalama

Tanım 4.16. n tane x_1, x_2, \dots, x_n değerlerinin harmonik ortlaması

$$H.O = \frac{1}{\frac{1}{n} \left(\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n} \right)} = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$$

Not: Hortalama en az yararlı olandır. Tepe değer ise, örneğin bir gazeteden alınan halkın görüşlerine dayanan veriler kullanıldığı zaman, gerçekten yararlı bir istatistik, bununla birlikte bilimsel verilerin analizinde her zaman için ihmal edilebilmektedir. Geometrik ortalamanın yararlı olabildiği bir kaç durum vardır. En çok göze çarpan rasgele değişkenlerin analizinde onların ortalama ve varyansı ile ilgili olanıdır. Yalnız orta değer ve aritmetik ortalamayı "iyi" merkezi eğilim ölçüsü olarak düşüneceğiz. Orta değer kolayca hesaplanabilir ve bazı durumlarda aritmetik ortalamadan daha gerçekçi ve durumu daha iyi yansıtır. Fakat, matematiksel olarak kontrol etmek (ele almak) daha güçtür ve bu nedenle yargıya varmada daha az istenir. Aritmetik ortalama kolaylıkla hesaplanır ve matematiksel olarak çalışmamıza daha uygundur. Bu nedenlerle aritmetik ortalama pek çok durumlarda konum ya da "merkezi eğilim" ölçüsü olarak kullanılır.

5 Dağılım Ölçüleri

Merkezsel eğilim ölçüleri birimlerin ilgili değişken bakımından aldıkları değerler hakkında genel bir bilgi vermekle beraber, birimlerin ne ölçüde, çok ya da az farklı değerlere sahip olup olmadığını göstermezler. Birimlerin ilgili değişken bakımından ne ölçüde farklı olduklarını ifade etmek için dağılım ölçüleri kullanılır.

Tanım 5.1. x_1, x_2, \dots, x_N sonlu sayıdaki gözlem değerlerinin aritmetik ortlaması

$$\mu = \frac{\sum_{i=1}^N x_i}{N}$$

kitle varyansı σ^2 ise

$$\begin{aligned} \sigma^2 &= \frac{\sum_{i=1}^N (x_i - \mu)^2}{N} \\ &= \frac{\sum_{i=1}^N x_i^2}{N} - \mu^2 \end{aligned}$$

olarak hesaplanır. Kitle standart sapması σ , varyansın (pozitif) kareköküdür. N gözlemden oluşan bir kitlede x_1, x_2, \dots, x_k gözlem değerleri, f_1, f_2, \dots, f_k frekansları olmak üzere $N = \sum_{i=1}^k f_i$ olmak üzere ortalama

$$\mu = \frac{\sum_{i=1}^k x_i f_i}{N}$$

kitle varyansı σ^2 ise

$$\begin{aligned} \sigma^2 &= \frac{\sum_{i=1}^k (x_i - \mu)^2 f_i}{N} \\ &= \frac{\sum_{i=1}^k x_i^2 f_i}{N} - \mu^2 \end{aligned}$$

Tanım 5.2. n gözlemden oluşan bir örneklemin ortalaması \bar{x} olmak üzere S^2 ile gösterilen örneklem varyansı gruplandırılmamış veriler için

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

gruplandırılmış veriler için, y_1, y_2, \dots, y_n sınıf ortaları f_1, f_2, \dots, f_k frekansları olmak üzere

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n f_i (y_i - \bar{x})^2$$

Tanım 5.3. Örneklem varyansının (pozitif) kareköküne örneklem standart sapması denir ve S ile gösterilir

Örnek 5.4. 1, 2, 3, 4, 5 sayılarından örneklem varyansını ve standart sapmayı hesaplayınız.

Çözüm:

$$\begin{aligned} \bar{x} &= \frac{1+2+3+4+5}{5} = \frac{5*6}{5*2} = 3 \\ S^2 &= \frac{(1-3)^2 + (2-3)^2 + (3-3)^2 + (4-3)^2 + (5-3)^2}{4} \\ &= 2.5 \\ S &= \sqrt{2.5} = 1.5811 \end{aligned}$$

□

Teorem 5.5. n gözlemin varyansı gruplandırılmamış gözlemler için

$$\begin{aligned} S^2 &= \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - n * \bar{x}^2 \right) \\ &= \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n} \right) \end{aligned}$$

gruplandırılmış gözlemler için

$$\begin{aligned} S^2 &= \frac{1}{n-1} \left(\sum_{i=1}^k f_i * y_i^2 - n * \bar{x}^2 \right) \\ &= \frac{1}{n-1} \left(\sum_{i=1}^k f_i * y_i^2 - \frac{(\sum_{i=1}^k f_i * y_i)^2}{n} \right) \end{aligned}$$

İspat: ilk formülleri ispatlayacağız: gruplandırılmamış gözlemler için

$$\begin{aligned}
 S^2 &= \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \\
 &= \frac{1}{n-1} \sum_{i=1}^n (x_i^2 - 2x_i\bar{x} + \bar{x}^2) \\
 &= \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - 2\bar{x} \sum_{i=1}^n x_i + \bar{x}^2 \sum_{i=1}^n 1 \right) \\
 &\quad \left(\sum_{i=1}^n x_i = n * \bar{x} \right) \\
 &= \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - n * \bar{x}^2 \right)
 \end{aligned}$$

gruplandırılmış gözlemler için

$$\begin{aligned}
 S^2 &= \frac{1}{n-1} \sum_{i=1}^n f_i (y_i - \bar{x})^2 \\
 &= \frac{1}{n-1} \sum_{i=1}^n f_i (y_i^2 - 2f_i\bar{x} + \bar{x}^2) \\
 &= \frac{1}{n-1} \left(\sum_{i=1}^n f_i y_i^2 - 2\bar{x} \sum_{i=1}^n f_i^2 + \bar{x}^2 \sum_{i=1}^n f_i \right) \\
 &\quad \left(\sum_{i=1}^n f_i y_i = n\bar{x} \right) \\
 &= \frac{1}{n-1} \left(\sum_{i=1}^k f_i * y_i^2 - n * \bar{x}^2 \right)
 \end{aligned}$$

Örnek 5.6. Aşağıdaki tablo için standart sapmayı bulunuz.

Sınıf No	Sınıf Limitleri	Sınıf Sınırları	Gözlem Sayısı (f_i)	Sınıf Orta Noktası(y_i)	Eklemeli Frekanslar
1	0 – 2	-0.5 ~ 2.5	1	$\frac{0+2}{2} = 1$	1
2	3 – 5	2.5 ~ 5.5	3	$\frac{3+5}{2} = 4$	4
3	6 – 8	5.5 ~ 8.5	8	$\frac{6+8}{2} = 7$	12
4	9 – 11	8.5 ~ 11.5	13	$\frac{9+11}{2} = 10$	25
5	12 – 14	11.5 ~ 14.5	20	$\frac{12+14}{2} = 13$	45
6	15 – 17	14.5 ~ 17.5	25	$\frac{15+17}{2} = 16$	70
7	18 – 20	17.5 ~ 20.5	17	$\frac{18+20}{2} = 19$	87
8	21 – 23	20.5 ~ 23.5	5	$\frac{21+23}{2} = 22$	92
9	24 – 26	23.5 ~ 26.5	3	$\frac{24+26}{2} = 25$	95
10	27 – 29	26.5 ~ 29.5	5	$\frac{27+29}{2} = 28$	100

Çözüm:

$$S^2 = \frac{1}{n-1} \left(\sum_{i=1}^k f_i * y_i^2 - \frac{\left(\sum_{i=1}^k f_i * y_i \right)^2}{n} \right)$$

Gözlem Sayısı (f_i)	Sınıf Orta Noktası(y_i)	y_i^2	$f_i * y_i$	$f_i * y_i^2$
1	$\frac{0+2}{2} = 1$	1	$1 * 1 = 1$	$1 * 1 = 1$
3	$\frac{3+5}{2} = 4$	16	$1 * 4 = 4$	$3 * 16 = 48$
8	$\frac{6+8}{2} = 7$	49	$8 * 7 = 56$	$8 * 49 = 392$
13	$\frac{9+11}{2} = 10$	100	$13 * 10 = 130$	$13 * 100 = 1300$
20	$\frac{12+14}{2} = 13$	169	$20 * 13 = 260$	$20 * 169 = 3380$
25	$\frac{15+17}{2} = 16$	256	$25 * 16 = 400$	$25 * 256 = 6400$
17	$\frac{18+20}{2} = 19$	361	$17 * 19 = 323$	$17 * 361 = 6137$
5	$\frac{21+23}{2} = 22$	484	$5 * 22 = 110$	$5 * 484 = 2420$
3	$\frac{24+26}{2} = 25$	625	$3 * 25 = 75$	$3 * 625 = 1875$
5	$\frac{27+29}{2} = 28$	784	$5 * 28 = 140$	$5 * 784 = 3920$
			$\sum_{i=1}^k f_i * y_i = 1507$	Toplam: 25873

$$S^2 = \frac{1}{n-1} \left(\sum_{i=1}^k f_i * y_i^2 - \frac{\left(\sum_{i=1}^k f_i * y_i \right)^2}{n} \right)$$

$$= \frac{1}{100-1} \left(25873 - \frac{1507^2}{100} \right)$$

□

Teorem 5.7. (1) a ve b sabitler x_1, x_2, \dots, x_n gruplandırılmamış gözlem değerleri $z_i = ax_i + b$ olmak üzere

$$S_z^2 = a^2 S_x^2$$

(2) a ve b sabitler y_1, y_2, \dots, y_n gruplandırılmış gözlem değerlerinin orta değeri $z_i = ay_i + b$ olmak üzere $S_z^2 = a^2 S_x^2$

İspat: (1)

$$S_z^2 = \frac{1}{n-1} \sum_{i=1}^n (z_i - \bar{z})^2$$

$$z_i = ax_i + b \Rightarrow \bar{z} = \frac{1}{n} \sum_{i=1}^n z_i = \frac{1}{n} \sum_{i=1}^n (ax_i + b) = a\bar{x} + b$$

$$S_z^2 = \frac{1}{n-1} \sum_{i=1}^n (ax_i + b - (a\bar{x} + b))^2 = \frac{a^2}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

$$= a^2 S_x^2$$

(2)

$$\begin{aligned}
S_z^2 &= \frac{1}{n-1} \sum_{i=1}^n f_i (z_i - \bar{z})^2 \\
&= \frac{1}{n-1} \sum_{i=1}^n f_i (ax_i + b_i - (a\bar{x} + b))^2 \\
&= \frac{a^2}{n-1} \sum_{i=1}^n f_i (x_i - \bar{x})^2 = a^2 S_x^2
\end{aligned}$$

Örnek 5.8. 1, 2, 3, 4, 5 sayılarından örneklem varyansı 2.5 olduğuna göre 5,7,9,11,13 sayılarının varyansı nedir?

Çözüm: $5=1*2+3, 7=2*2+3, 9=3*2+3, 11=4*2+3, 13=5*2+3$ olduğundan $a=2, b=3$ ve yukarıdaki teoremi uygularsak

$$S_z^2 = 4 * S_x^2 = 4 * 2.5 = 10$$

□

Teorem 5.9. n ve m birikimli örneklerden hesaplanan iki bağımsız varyans sırasıyla S_1^2, S_2^2 olmak üzere $n+m$ örneklemin birleşmiş varyansı

$$S^2 = \frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}$$

Tanım 5.10. n gözlemden oluşan bir örneklemin ortalaması \bar{x} olmak üzere gruplandırılmamış veriler için ortalama mutlak sapma

$$OMS = \frac{1}{n} \sum_{i=1}^n |x_i - \bar{x}|$$

gruplandırılmış veriler için, y_1, y_2, \dots, y_n sınıf ortaları f_1, f_2, \dots, f_k frekansları olmak üzere ortalama mutlak sapma

$$S^2 = \frac{1}{n} \sum_{i=1}^n f_i |y_i - \bar{x}|$$

ile hesaplanır.

Örnek 5.11. 1, 2, 3, 4, 5 sayılarının ortalama mutlak sapmasını hesaplayınız.

İspat:

$$\begin{aligned}
\bar{x} &= \frac{1+2+3+4+5}{5} = \frac{5*6}{5*2} = 3 \\
OMS &= \frac{|1-3| + |2-3| + |3-3| + |4-3| + |5-3|}{5} \\
&= 1.2
\end{aligned}$$

Tanım 5.12. Orta deger fikrinin genişletilmesi olarak, gözlemler tekrar eşit frekanslı iki gruba bölünebilir. Böylece gözlemlerin %25 i Q_1 altında, %25 i Q_3 ün üzerinde olmak üzere iki yeni Q_1 ve Q_3 değerleri tanımlanabilir. Q_1 ve Q_3 sayılarına dörtte birlikler (çeyrekler) denir $d = Q_3 - Q_1$ farkına verilerin iç çeyrekler arası değişim genişliği denir. Bir grup gözlemde Q_1 ve Q_3 öyle sayılardır ki gözlemlerin %25 i Q_1 in altında, %25 i Q_3 ün üzerindedir. Orta deger Q_1 ve Q_3 ile birlikte gözlemleri eşit frekanslı dört gruba bölerler. Frekans dağılımı için

çizilen eğrinin altında, veri kümesi 4 adet %25 lik gruba ayrılmıştır. Gruplandırılmış gözlemler için Q_1 ve Q_3 sayıları aşağıdaki gibi bulunur:

$$\begin{aligned}
 n_1 &= \text{İlk dörtte birliğin bulunduğu sınıftan önceki sınıfların toplam frekansı} \\
 J_1 &= \frac{n}{4} - n_1 \\
 f_{Q_1} &= \text{İlk dörtte birliğin bulunduğu sınıfın frekansı} \\
 h &= \text{Sınıf genişliği} \\
 L_1 &= \text{İlk dörtte birliğin bulunduğu sınıfın alt sınıf sınırı} \\
 Q_1 &= L_1 + \frac{J_1 * h}{f_{Q_1}} \\
 n_3 &= \text{Üst çeyreğin bulunduğu sınıftan önceki sınıfların toplam frekansı} \\
 L_3 &= \text{Üst çeyreğin bulunduğu sınıfın alt sınıf sınırı} \\
 f_{Q_3} &= \text{Üst çeyreğin bulunduğu sınıfın frekansı} \\
 J_3 &= \frac{3n}{4} - n_3 \\
 Q_3 &= L_3 + \frac{J_3 * h}{f_{Q_3}}
 \end{aligned}$$

Örnek 5.13. 1, 2, 3, 4, 4, 6, 7, 8, 9, 10, 10 sayıları için Q_1 ve Q_3 sayılarını bulunuz.

Çözüm: Orta değer 6 dır. Toplam olarak 11 sayı mevcuttur. $11*25/100 = 2.75$ üste yuvarlandığında 3. sayı 3 Q_1 sayısıdır. $11*75/100 = 8.25$ üste yuvarlandığında 9. sayı olan 9 Q_3 sayısıdır. \square

Örnek 5.14. Aşağıdaki tablo için Q_1 ve Q_3 sayılarını bulunuz.

Sınıf No	Sınıf Limitleri	Sınıf Sınırları	Gözlem Sayısı (f_i)	Sınıf Orta Noktası(y_i)	Eklemeli Frekanslar
1	0 – 2	-0.5 ~ 2.5	1	$\frac{0+2}{2} = 1$	1
2	3 – 5	2.5 ~ 5.5	3	$\frac{3+5}{2} = 4$	4
3	6 – 8	5.5 ~ 8.5	8	$\frac{6+8}{2} = 7$	12
4	9 – 11	8.5 ~ 11.5	13	$\frac{9+11}{2} = 10$	25
5	12 – 14	11.5 ~ 14.5	20	$\frac{12+14}{2} = 13$	45
6	15 – 17	14.5 ~ 17.5	25	$\frac{15+17}{2} = 16$	70
7	18 – 20	17.5 ~ 20.5	17	$\frac{18+20}{2} = 19$	87
8	21 – 23	20.5 ~ 23.5	5	$\frac{21+23}{2} = 22$	92
9	24 – 26	23.5 ~ 26.5	3	$\frac{24+26}{2} = 25$	95
10	27 – 29	26.5 ~ 29.5	5	$\frac{27+29}{2} = 28$	100

Çözüm: $n = 100, n/4 = 25$

$$f_1 + f_2 + f_3 + f_4 = 25 \geq 25$$

olduğundan 4.sınıf Q_1 in olduğu sınıftır.

$$L_1 = 8.5,$$

$$n_1 = 1 + 3 + 8 = 12$$

$$f_{Q_1} = 13$$

$$h = 3$$

$$J_1 = \frac{n}{4} - n_1 = 25 - 12 = 13$$

$$Q_1 = L_1 + \frac{J_1 * h}{f_{Q_1}} = 8.5 + \frac{13 * 3}{13} = 11.5$$

$$3n/4 = 75$$

$$f_1 + f_2 + f_3 + f_4 + f_5 + f_6 + f_7 = 87 \geq 75$$

olduğundan 7.sınıf Q_3 in olduğu sınıftır.

$$L_3 = 17.5,$$

$$n_3 = f_1 + f_2 + f_3 + f_4 + f_5 + f_6 = 70$$

$$f_{Q_3} = 17$$

$$h = 3$$

$$J_3 = \frac{3n}{4} - n_3 = 75 - 70 = 5$$

$$Q_3 = L_3 + \frac{J_3 * h}{f_{Q_3}} = 17.5 + \frac{5 * 3}{17} = 18.382$$

□

5.1 Çarpıklık Ölçütü

Tanım 5.15. Veri kümesi için karşılık gelen çarpıklık ölçütü $(x_i - \bar{x})^3$ kübik sapmaların ortalaması ile hesaplanır ve m_3 ile gösterilir \bar{x} gözlem değerlerinin ortalama değeri olmak üzere çarpıklık ölçütü

$$m_3 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^3$$

formülü ile hesaplanır.

Örnek 5.16. -3,5,40 değerleri için çarpıklık ölçütünü bulunuz.

Çözüm:

$$\begin{aligned} \bar{x} &= \frac{-3 + 5 + 40}{3} = 14 \\ m_3 &= \frac{(-3 - 14)^3 + (5 - 14)^3 + (40 - 14)^3}{2} = 5967 \end{aligned}$$

□

Tanım 5.17. x_1, x_2, \dots, x_n gözlem değerleri için \bar{x} gözlem değerlerinin ortalama değeri ve S standart sapması olsun. x_i lere karşılık gelen

$$y_i = \frac{x_i - \bar{x}}{S}$$

değerine standartlaştırılmış gözlem değeri

$$\tilde{m}_3 = \frac{m_3}{S^3}$$

değerine de standartlaştırılmış çarpıklık ölçütü denir.

5.2 Sivrilik (basıklık) ölçütü

Tanım 5.18. Veri kümesi için sivrilik ölçütü m_4 ile gösterilir \bar{x} gözlem değerlerinin ortalama değeri olmak üzere sivrilik ölçütü

$$m_4 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^4$$

formülü ile hesaplanır. x_1, x_2, \dots, x_n gözlem değerleri için \bar{x} gözlem değerlerinin ortalama değeri ve S standart sapması olsun.

$$\tilde{m}_4 = \frac{m_4}{S^4}$$

değerine de standartlaştırılmış sivrilik ölçütü denir.

6 Örneklem Dağılımları ve Tahmin Etme

İstatistiğin en önemli problemlerinden biri örneklemden bulunan sonuçları kitleye genelleştirmektir. Bu bölümde bu genelleştirmenin nasıl yapıldığını göreceğiz.

Tanım 6.1. Bir istatistiğin olasılık dağılımına örneklem dağılımı denir.

Örnek 6.2. Bir rasgele değişken ortalaması μ ve varyansı σ^2 olan bir normal dağılıma sahipse, bu dağılımdan alınmış bir x_1, x_2, \dots, x_n rasgele örneklemini kullanarak μ ve σ^2 'yi tahmin etmeyi isteyebiliriz.

Örnek 6.3. Belli bir şehirdeki doğum kayıtları doğmuş 1000 bebeğin 510'unun erkek olduğunu gösterirse, p nin tahmini değeri yani doğacak bir çocuğun erkek olma olasılığı 0.510'dur.

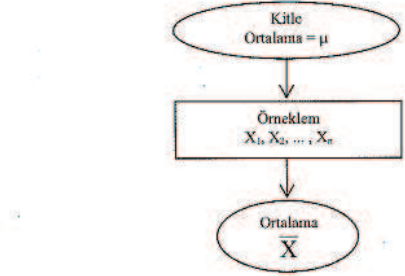
Tanım 6.4. Örneklem ortalaması \bar{X} , kitle ortalaması μ olmak üzere kitle parametresini tahmin etmek için kitle yerine örneklemin kullanılmasından oluşan hataya, yani $\bar{X} - \mu$ farkına örneklem hatası denir. Örneklem kitlenin tümünü kapsadığında $\bar{X} = \mu$

İstatistiksel yöntemleri kullanarak kitle hakkında sonuç çıkarmak için rasgele örneklem seçilecektir. Örneğin belli tipte bir küçük hesap makinesi için üretilmiş lityum pillerin oluşturduğu büyük bir kitleyi düşünelim. Bu kitlenin ne olasılık dağılımı nede ortalama ve varyansı hakkında bilgi sahibiyiz. Bu kitleden n birimlik pilden oluşan rasgele örneklem seçmek istiyoruz. Herhangi bir pilin seçimi diğerinden bağımsızdır.

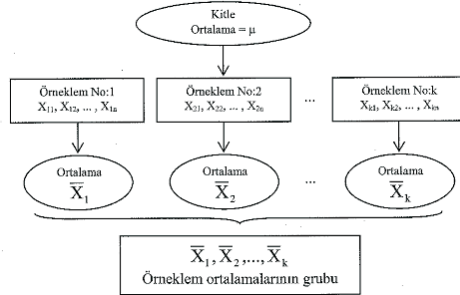
Kitledeki pillerin dayanma süresi X rasgele değişkeni olsun. $i = 1, 2, \dots, n$ olmak üzere inceleme yapmak için seçilecek i -inci pilin dayanma süresi X_i bir rasgele değişkendir. Ayrıca rasgele değişkenler birbirlerinden bağımsızdırlar ve özdeş dağılıma sahiptirler. Seçim sonrasında n pilin dayanma süreleri gözlemlendiğinde x_1, x_2, \dots, x_n sayılan elde edilecektir. Bu sayılar X_1, X_2, \dots, X_n rasgele değişkenlerinin gözlenen değerleri olur.

O halde "rasgele örneklem" ifadesini inceleme yapmak üzere seçilen n nesne, seçilecek nesnelerle ilgili rasgele değişkenler veya değişkenlerin sayısal değerleri için kullanacağız.

Gerçek örneklem sürecinde verilerin bir tek örneklemini seçip ortalamayı hesaplarız.



Karmaşık örnekleme sürecinde ise μ ortalamalı bir kitleden n birimlik k adet örneklem seçilerek ortalamaları hesaplanır.



7 Örneklem Ortalaması ve Varyansın Bazı Özellikleri

Teorem 7.1. X_1, X_2, \dots, X_n rasgele değişkenler ve $\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$, ve her birinin beklenen değeri μ ise $E[\bar{X}] = \mu$

İspat:

$$\begin{aligned}
 E[\bar{X}] &= E\left[\frac{X_1 + X_2 + \dots + X_n}{n}\right] \\
 &= \frac{1}{n} (E[X_1] + E[X_2] + \dots + E[X_n]) \\
 &= \frac{1}{n} (\mu + \mu + \dots + \mu) = \frac{1}{n} n\mu = \mu
 \end{aligned}$$

Örnek 7.2. İki zar atıldığında üstte gelen yüzlerdeki noktaların sayısı sırasıyla X_1 ve X_2 ise $E[\bar{X}] = 7/2$ olduğunu gösteriniz.

Çözüm:

$$\begin{aligned}
 E[X_1] &= E[X_2] = \frac{7}{2} \\
 E[\bar{X}] &= E\left[\frac{X_1 + X_2}{2}\right] \\
 &= \frac{1}{2}(E[X_1] + E[X_2]) \\
 &= \frac{1}{2}\left(\frac{7}{2} + \frac{7}{2}\right) = \frac{7}{2}
 \end{aligned}$$

□

Örnek 7.3. Kitle 2, 4, 6, 8, 10 olarak verilsin. Yerine koymaksızın bu kitleden 2 birimlik örneklem seçelim. Örneklem ortalamasının beklenen değerini bulunuz.

Çözüm: Kitle ortalaması

$$\mu = \frac{1}{5}(2 + 4 + 6 + 8 + 10) = 6$$

5 kitleden 2 li seçim $C_{5,2} = \frac{5!}{2!*3!} = 10$ farklı biçimde gerçekleşir.

Örneklem	Ortalaması	Olasılığı
2, 4	$(2 + 4)/2 = 3$	$\frac{1}{10}$
2, 6	$(2 + 6)/2 = 4$	$\frac{1}{10}$
2, 8	$(2 + 8)/2 = 5$	$\frac{1}{10}$
2, 10	$(2 + 10)/2 = 6$	$\frac{1}{10}$
4, 6	$(4 + 6)/2 = 5$	$\frac{1}{10}$
4, 8	$(4 + 8)/2 = 6$	$\frac{1}{10}$
4, 10	$(4 + 10)/2 = 7$	$\frac{1}{10}$
6, 8	$(6 + 8)/2 = 7$	$\frac{1}{10}$
6, 10	$(6 + 10)/2 = 8$	$\frac{1}{10}$
8, 10	$(8 + 10)/2 = 9$	$\frac{1}{10}$

$\bar{X} = x$	3	4	5	6	7	8	9
$f(x)P(\bar{X} = x)$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{1}{10}$	$\frac{1}{10}$

$$\begin{aligned}
 E[\bar{X}] &= 3 * \frac{1}{10} + 4 * \frac{1}{10} + 5 * \frac{2}{10} + 6 * \frac{2}{10} + 7 * \frac{2}{10} + 8 * \frac{1}{10} + 9 * \frac{1}{10} \\
 &= 6
 \end{aligned}$$

□

Teorem 7.4. X_1, X_2, \dots, X_n rasgele değişkenler ve $E[X_i] = \mu, \text{Var}(X_i) = \sigma^2$ ise $\text{Var}(\bar{X}) = \frac{\sigma^2}{n}$ dir.

İspat:

$$\begin{aligned}
 Var(\bar{X}) &= E[(\bar{X} - \mu)^2] \\
 &= E\left[\left(\frac{X_1 + X_2 + \dots + X_n}{n} - \mu\right)^2\right] \\
 &= E\left[\left(\frac{X_1 + X_2 + \dots + X_n}{n} - \frac{n\mu}{n}\right)^2\right] \\
 &= E\left[\left(\frac{(X_1 - \mu) + (X_2 - \mu) + \dots + (X_n - \mu)}{n}\right)^2\right] \\
 &= E\left[\frac{(X_1 - \mu)^2}{n^2} + \dots + \frac{(X_n - \mu)^2}{n^2} + 2\left(\frac{(X_1 - \mu)(X_2 - \mu)}{n^2} + \dots + \frac{(X_{n-1} - \mu)(X_n - \mu)}{n^2}\right)\right] \\
 &= \frac{1}{n^2} \left(E[(X_1 - \mu)^2] + \dots + E[(X_n - \mu)^2]\right) \\
 &\quad + \frac{2}{n^2} \left(\underbrace{E[(X_1 - \mu)(X_2 - \mu)]}_{=0, (Bağımsız)} + \dots + \underbrace{E[(X_{n-1} - \mu)(X_n - \mu)]}_{=0, (Bağımsız)}\right) \\
 &= \frac{1}{n^2} (Var(X_1) + \dots + Var(X_n)) = \frac{1}{n^2} (\sigma^2 + \dots + \sigma^2) = \frac{\sigma^2}{n}
 \end{aligned}$$

Örnek 7.5. İki zar atıldığında üstte gelen yüzlerdeki noktaların sayısı sırasıyla X_1 ve X_2 ise $Var(\bar{X})$ değerini hesaplayınız.

Çözüm:

x	$f(x) = P(X = x)$	$x * f(x)$	$x^2 * f(x)$
1	$\frac{1}{6}$	$\frac{1}{6}$	$1^2 * \frac{1}{6} = \frac{1}{6}$
2	$\frac{1}{6}$	$\frac{2}{6}$	$2^2 * \frac{1}{6} = \frac{4}{6}$
3	$\frac{1}{6}$	$\frac{3}{6}$	$3^2 * \frac{1}{6} = \frac{9}{6}$
4	$\frac{1}{6}$	$\frac{4}{6}$	$4^2 * \frac{1}{6} = \frac{16}{6}$
5	$\frac{1}{6}$	$\frac{5}{6}$	$5^2 * \frac{1}{6} = \frac{25}{6}$
6	$\frac{1}{6}$	$\frac{6}{6}$	$6^2 * \frac{1}{6} = \frac{36}{6}$
		$Toplam : \mu_x = E[X] = \frac{7}{2}$	$Toplam : E[X^2] = \frac{91}{6}$

$$\sigma_X^2 = Var(X) = E[X^2] - \mu_X^2 = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}$$

$$Var(X_1) = Var(X_2) = \sigma^2 = \frac{35}{12}$$

$$Var(\bar{X}) = Var\left(\frac{X_1 + X_2}{2}\right) = \frac{\sigma^2}{2} = \frac{\frac{35}{12}}{2} = \frac{35}{24}$$

□

Örnek 7.6. İki para 3 kez atılsın. Turaların ortalaması \bar{X} ise $E[\bar{X}]$ ve $Var(\bar{X})$ değerlerini hesaplayınız.

1 para için ebeklenen değer

S	YYY	YYT	YTY	TTY	YTT	TYT	TTY	TTT
X	0	1	1	1	2	2	2	3
$P(X = x)$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

$$f(0) = P(X = 0) = \frac{1}{8}, f(1) = P(X = 1) = \frac{3}{8}$$

$$f(2) = P(X = 2) = \frac{3}{8}, f(3) = P(X = 3) = \frac{1}{8}$$

$$E[X] = \sum_x x f(x) = 0 * \frac{1}{8} + 1 * \frac{3}{8} + 2 * \frac{3}{8} + 3 * \frac{1}{8} = \frac{3}{2}$$

$$E[X_1] = E[X_2] = \frac{3}{2} \Rightarrow E[\bar{X}] = \frac{3}{2}$$

bir para için varyans

X	0	1	2	3
$f(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$
$(x - \mu_x)^2$	$(-\frac{3}{2})^2 = \frac{9}{4}$	$(1 - \frac{3}{2})^2 = \frac{1}{4}$	$(2 - \frac{3}{2})^2 = \frac{1}{4}$	$(3 - \frac{3}{2})^2 = \frac{9}{4}$
$(x - \mu_x)^2 f(x)$	$\frac{9}{4} * \frac{1}{8} = \frac{9}{32}$	$\frac{1}{4} * \frac{3}{8} = \frac{3}{32}$	$\frac{1}{4} * \frac{3}{8} = \frac{3}{32}$	$\frac{9}{4} * \frac{1}{8} = \frac{9}{32}$

$$\sigma_X^2 = Var(X) = \sum_x (x - \mu_x)^2 f(x) = \frac{9}{32} + \frac{3}{32} + \frac{3}{32} + \frac{9}{32} = \frac{3}{4}$$

$$Var(\bar{X}) = Var\left(\frac{X_1 + X_2}{2}\right) = \frac{\sigma^2}{2} = \frac{\frac{3}{4}}{2} = \frac{3}{8}$$

Teorem 7.7. *N herhangi sonlu sayıda kitleden oluşan bir kitlenin büyüklüğünü. n bu kitleden yerine koymaksızın seçilen bir örneklemin büyüklüğünü gösterirse, n birimlik tüm olanaklı örneklem için ortalamaların ortalaması μ_x . . ortalamaların varyansı σ_x^2 olmak üzere*

$$\begin{aligned} \mu_x &= E[\bar{X}] = \mu \\ Var(\bar{X}) &= \sigma_x^2 = \frac{\sigma^2}{n} \frac{N - n}{N - 1} \end{aligned}$$

ile hesaplanır.

Teorem 7.8. *N herhangi sonlu sayıda kitleden oluşan bir kitlenin büyüklüğünü. n bu kitleden yerine koyma yolu ile seçilen bir örneklemin büyüklüğünü gösterirse, n birimlik tüm olanaklı örneklem için ortalamaların ortalaması μ_x . . ortalamaların varyansı σ_x^2 olmak üzere*

$$\begin{aligned} \mu_x &= E[\bar{X}] = \mu \\ Var(\bar{X}) &= \sigma_x^2 = \frac{\sigma^2}{n} \end{aligned}$$

ile hesaplanır.

Örnek 7.9. 7, 10, 9, 11, 13, 12, 15 kitlesinden $n = 2$ olan rasgele örneklemeleri yerine koymaksızın seçtiğimizde, $E[\bar{X}]$ ve $Var(\bar{X})$ değerlerini hesaplayınız.

Çözüm: 1.Yol: 7 kitleden 2 li seçim $C_{7,2} = \frac{7!}{2!*5!} = 21$ farklı biçimde gerçekleşir.

Örneklem	Ortalaması	Olasılığı
7, 10	8.5	$\frac{1}{21}$
7, 9	8	$\frac{1}{21}$
7, 11	9	$\frac{1}{21}$
7, 13	10	$\frac{1}{21}$
7, 12	9.5	$\frac{1}{21}$
7, 15	11	$\frac{1}{21}$
10, 9	9.5	$\frac{1}{21}$
10, 11	10.5	$\frac{1}{21}$
10, 13	11.5	$\frac{1}{21}$
10, 12	11	$\frac{1}{21}$
10, 15	12.5	$\frac{1}{21}$
9, 11	10	$\frac{1}{21}$
9, 13	11	$\frac{1}{21}$
9, 12	10.5	$\frac{1}{21}$
9, 15	12	$\frac{1}{21}$
11, 13	12	$\frac{1}{21}$
11, 12	11.5	$\frac{1}{21}$
11, 15	13	$\frac{1}{21}$
13, 12	12.5	$\frac{1}{21}$
13, 15	14	$\frac{1}{21}$
12, 15	13.5	$\frac{1}{21}$

$\bar{X} = x$	$P(\bar{X} = x)$
8	$\frac{1}{21}$
8.5	$\frac{1}{21}$
9	$\frac{1}{21}$
9.5	$\frac{2}{21}$
10	$\frac{2}{21}$
10.5	$\frac{2}{21}$
11	$\frac{3}{21}$
11.5	$\frac{2}{21}$
12	$\frac{2}{21}$
12.5	$\frac{2}{21}$
13	$\frac{1}{21}$
13.5	$\frac{1}{21}$
14	$\frac{1}{21}$

$$\begin{aligned}
E[\bar{X}] &= \frac{1}{21} * 8 + \frac{1}{21} * 8.5 + \frac{1}{21} * 9 + \frac{2}{21} * 9.5 + \frac{2}{21} * 10 \\
&+ \frac{2}{21} * 10.5 + \frac{3}{21} * 11 + \frac{2}{21} * 11.5 + \frac{2}{21} * 12 + \frac{2}{21} * 12.5 \\
&+ \frac{1}{21} * 13 + \frac{1}{21} * 13.5 + \frac{1}{21} * 14 = 11
\end{aligned}$$

$\bar{X} = x$	$P(\bar{X} = x)$	$(\bar{X} - E[\bar{X}])^2$
8	$\frac{1}{21}$	$(8 - 11)^2 = 9$
8.5	$\frac{1}{21}$	$(8.5 - 11)^2 = 6.25$
9	$\frac{1}{21}$	$(9 - 11)^2 = 4$
9.5	$\frac{2}{21}$	$(9.5 - 11)^2 = 2.25$
10	$\frac{2}{21}$	$(10 - 11)^2 = 1$
10.5	$\frac{2}{21}$	$(10.5 - 11)^2 = 0.25$
11	$\frac{3}{21}$	$(11 - 11)^2 = 0$
11.5	$\frac{2}{21}$	$(11.5 - 11)^2 = 0.25$
12	$\frac{2}{21}$	$(12 - 11)^2 = 1$
12.5	$\frac{2}{21}$	$(12.5 - 11)^2 = 2.25$
13	$\frac{1}{21}$	$(13 - 11)^2 = 4$
13.5	$\frac{1}{21}$	$(13.5 - 11)^2 = 6.25$
14	$\frac{1}{21}$	$(14 - 11)^2 = 9$

$$\begin{aligned}
Var(\bar{X}) &= \frac{1}{21} * 9 + \frac{1}{21} * 6.25 + \frac{1}{21} * 4 + \frac{2}{21} * 2.25 + \frac{2}{21} * 1 \\
&+ \frac{2}{21} * 0.25 + \frac{3}{21} * 0 + \frac{2}{21} * 0.25 + \frac{2}{21} * 1 + \frac{2}{21} * 2.25 \\
&+ \frac{1}{21} * 4 + \frac{1}{21} * 6.25 + \frac{1}{21} * 9 = 2.5
\end{aligned}$$

2. Yol (Yukarıdaki Teoremi kullanarak)

$$\begin{aligned}
\mu &= \frac{1}{7} (7 + 10 + 9 + 11 + 13 + 12 + 15) = 11 \\
E[\bar{X}] &= \mu = 11 \\
\sigma^2 &= \frac{1}{7} \left((7 - 11)^2 + (10 - 11)^2 + (9 - 11)^2 + (11 - 11)^2 + (13 - 11)^2 + (12 - 11)^2 + (15 - 11)^2 \right) \\
&= 6 \\
Var(\bar{X}) &= \sigma_x^2 = \frac{\sigma^2}{n} \frac{N - n}{N - 1} \\
&= \frac{6}{2} * \frac{7 - 2}{7 - 1} = 2.5
\end{aligned}$$

□

Tanım 7.10. Örneklem ortalamasının standart sapması σ_x ye ortalamamanın standart hatası denir.

Yukarıdaki örnekte görüldüğü üzere $\sigma^2 = 6, \sigma_x^2 = 2.5$ olup $\sigma = \sqrt{6}, \sigma_x = \sqrt{2.5}$ dır yani örneklem ortalamasının standart sapması dah küçüktür. Buna göre

- (1) $\sigma_x < \sigma$
- (2) Daha büyük n değerleri için daha küçük σ_x değerleri elde edilir.
- (3) $\sigma_x = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$ formülünde $\sqrt{\frac{N-n}{N-1}}$ değerine sonlu kitle çarpanı denir.
- (4) N oldukça büyük ve n oldukça küçük bir değer ise $\sigma_x \simeq \frac{\sigma}{\sqrt{n}}$ olarak alınabilir.

Teorem 7.11. X_1, X_2, \dots, X_n rasgele değişkenler ve $E[X_i] = \mu, \text{Var}(X_i) = \sigma^2$ olan bir normal dağılıma sahip ise $E[\bar{X}] = \mu_{\bar{x}} = \mu, \text{Var}(\bar{X}) = \sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$ dir.

Teorem 7.12. X_1, X_2, \dots, X_n rasgele değişkenler ve $E[X_i] = \mu, \text{Var}(X_i) = \sigma^2$ olan bir normal dağılıma sahip olmak üzere

$$Z = \frac{\bar{X} - \mu}{\sigma_{\bar{x}}} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

ile tanımlanan Z değişkeni n nin büyük değerleri için ($n > 30$) yaklaşık olarak standart normal dağılıma sahiptir.

Örnek 7.13. Bir sanayi kuruluşunda çalışanların gündelikleri ortalama $\mu = 800.000$ TL ve $\sigma = 90.000$ TL standart sapmalıdır. Rasgele seçilen 81 işçinin gündeliklerinin ortalamasının 810.000 TL ve 825.000 TL arasında bulunması olasılığı nedir? $P(810.000 < \bar{X} < 825.000) = ?$

Çözüm: $n=81 > 30$ olduğundan normal dağılım düşüneceğiz. Öncelikle kitlenin varyansının hesaplayalım:

$$\begin{aligned}\sigma_x &= \frac{\sigma}{\sqrt{n}} = \frac{90000}{\sqrt{81}} = 10\,000 \\ Z &= \frac{\bar{X} - \mu}{\sigma_{\bar{x}}}\end{aligned}$$

değişkenini kullanarak

$$\begin{aligned}z_1 &= \frac{810000 - 800000}{10\,000} = 1 \\ z_2 &= \frac{825000 - 800000}{10\,000} = 2.5 \\ P(810.000 < \bar{X} < 825.000) &= P(1 < \bar{Z} < 2.5) \\ &= P(\bar{Z} < 2.5) - P(\bar{Z} \leq 1) \\ &= 0.9938 - 0.8413 = 0.1525\end{aligned}$$

□