2. Frekans Dağılımları

2.1 Genel tanımlar

- **Ham veriler (Liste).** Eğer derlenen veriler ilgilenilen konunun dışından başka bir yönde, örneğin; gözlem sırasına göre sıralanmışsa, bu sıralamaya "liste" adı verilir. Listeler ham yani işlenmemiş verilerdir. Açıktır ki, derlenen verilerden ihtiyaç duyulan bilgilerin bir liste yardımıyla elde edilmesi veri sayısı arttıkça giderek zorlaşır. Çünkü, her aşamada listedeki sonuçların tekrar tekrar gözden geçirilmesi gerekir.
- **Örnek.** Aşağıda 80 öğrencinin Matematik notlarının listesi verilmiştir.

68	84	75	82	68	90	62	88	76	93
73	79	88	73	60	93	71	59	85	75
61	65	75	87	74	62	95	78	63	72
66	78	82	75	94	77	69	74	68	60
96	78	89	61	75	95	60	79	83	71
79	62	67	97	78	85	76	65	71	75
65	80	73	57	88	78	62	76	53	74
86	67	73	81	72	63	76	75	85	77

- **Sıralanış (Seri).** İşlenmemiş ham nümerik verilerin yani listelerin artan veya azalan sırada yeniden yazılmasıdır. En küçük değer ile en büyük değer arasındaki farka serinin açıklığı denir. Yukarıdaki örnekte en büyük değer 97 ve en küçük değer 53 olduğundan açıklık $97 - 53 = 44$ tür.

2.1 Genel tanımlar

- **Frekans Dağılımı (veya frekans serisi).** Verilerin daha kolay kavranması açısından, gözlem değerlerinin yanına gözlem değerinin kaç kez tekrarlandığı kaydedilerek oluşturulan seriye "frekans serisi" veya "frekans dağılımı" tekrarlara da "frekans" adı verilir, f ile gösterilir.
- **Örnek.** Bir sınıfta bulunan 100 öğrencinin boy uzunlukları aşağıdaki gibi verildiğinde bir frekans serisi verilmiş olur.

Boy uzunluğu	Öğrenci Sayısı (f)
60–62	5
63–65	18
66–68	42
69–71	27
72–74	8
	Toplam 100

2.1 Genel tanımlar

- **Sınıf aralıkları ve sınıf sınırları.** Bir seride değeri a ile b arasında olan tüm elemanlar $a - b$ ile gösterilir. Bu gösterim $a - b$ sınıfı olarak adlandırılır. a ya sınıfın alt sınırı b ye ise sınıfın üst sınırı denir. Eğer sınıflar

Sınıflar	Frekans
$a - b$	f_1
$b - c$	f_2
$c - d$	f_3
$\vdots$	$\vdots$

şeklinde verilmişse dikkat edilmesi gereken nokta alt sınır bir sınıfa dahil üst sınır bir sınıfa dahil değildir. Yani yukarıdaki örnekte $a - b$ sınıfının üst sınırı olan b dir. Değeri tam olarak b olan elemanlar bu sınıfta değil bir sonraki $b - c$ sınıfındadır.

2.1 Genel tanımlar

- **Sınıf uzunluğu (büyüklüğü veya genişliği).** Bir sınıfın üst sınırı ile alt sınırı arasındaki farka bu sınıfın uzunluğu (büyüklüğü veya genişliği) denir, c ile gösterilir.
- **Sınıf numarası (sınıf orta noktası veya sınıf sırası).** Bir sınıfın alt ve üst sınırlarının toplamının yarısıdır. Bu sınıftaki tüm elemanların değerinin bu sınıfın numarasına eşit olduğu kabul edilir.
- **Örnek.**

Sınıflar	Frekans	Sınıf Numarası	Sınıf Uzunluğu (c)
0 — 10	3	5	10
10 — 26	5	18	16
26 — 30	8	28	4
30 — 50	12	40	20
50 — 60	7	55	10
60 — 90	4	75	30

2.2 Histogramlar ve Frekans Poligonları

- Frekans dağılımları genel olarak iki tür grafikle gösterilir. Bunlar histogram ve frekans poligonu adı verilen grafiklerdir.
- Histogram denilen grafikler x eksenine dik olan, alanı ilgili sınıfın frekansına ve tabanı da ilgili sınıfın uzunluğuna (c değerine) eşit, birbirine bitişik dikdörtgenlerden oluşan bir grafik gösterimdir. Bir histogram çizilmeden önce, sözü edilen dikdörtgenlerin yüksekliklerinin ayarlanması gerekir. Bunun için frekanslar sınıf aralığına bölünerek, dikdörtgenlerin alanları ilgili sınıfların frekanslarına eşit hale getirilir.
- **Örnek.** Aşağıdaki frekans dağılımının histogramını çiziniz.

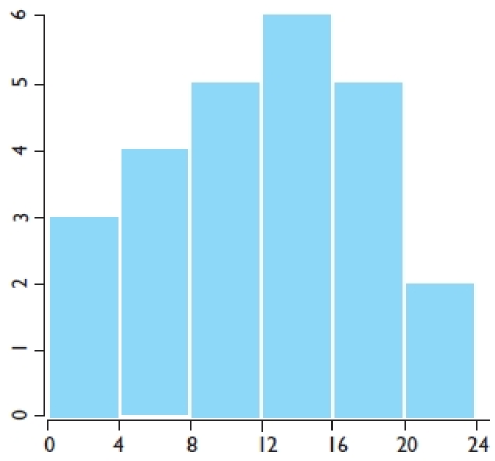
Sınıflar	frekanslar
0 – 4	12
4 – 8	16
8 – 12	20
12 – 16	24
16 – 20	20
20 – 24	8
	100

- **Çözüm.** Önce histogramı oluşturan dikdörtgenlerin taban ve yükseklik uzunlukları belirlenir.

Sınıflar	f	Sınıf Aralıkları c	Ayarlanmış Frekanslar f/c.
0 – 4	12	4	12 / 4 = 3.0
4 – 8	16	4	16 / 4 = 4.0
8 – 12	20	4	20 / 4 = 5.0
12 – 16	24	4	24 / 4 = 6.0
16 – 20	20	4	20 / 4 = 5.0
20 – 24	8	4	8 / 4 = 2.0
	100		

.

Sınıf aralıklarının dikdörtgenlerin alt taban uzunlukları ve ayarlanmış frekansların yani f/c değerlerinin de dikdörtgenlerin yükseklikleri olduğu düşünülürse istenen histogram aşağıdaki gibi olur:

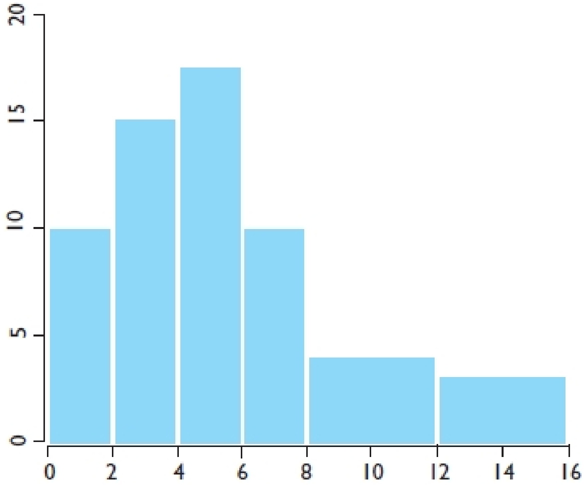


Örnek. Aşağıdaki serinin histogramını çiziniz.

Sınıflar	f
0 – 2	20
2 – 4	30
4 – 6	36
6 – 8	20
8 – 12	16
12 – 16	12
	128

Çözüm.

Sınıflar	f	c	f/c
0 – 2	20	2	10.0
2 – 4	30	2	15.0
4 – 6	36	2	18.0
6 – 8	20	2	10.0
8 – 12	16	4	4.0
12 – 16	12	4	3.0
	134		

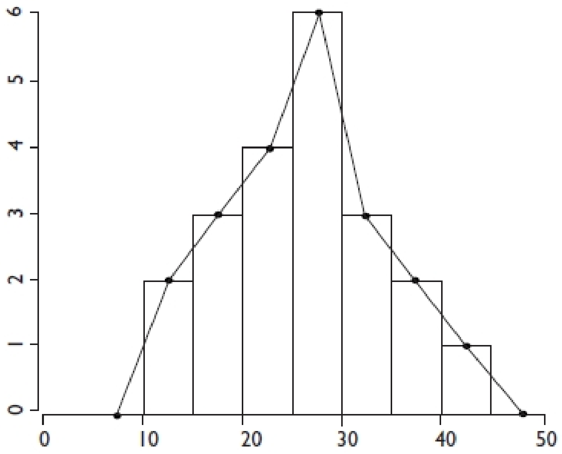


- Frekans poligonu, histogramın tepe orta noktalarının birleştirilmesiyle elde edilen, sınıflandırılmış serilere ilişkin, diğer bir grafik türüdür. Histogramın tepe orta noktaları, ilgili sınıflara ilişkin değişkenlerin değerlerini ifade ettiğinden, frekans poligonu, değişkenlerin değerlerine göre oluşturulmuş bir grafiktir.
- **Örnek.** Aşağıdaki serinin frekans poligonunu çiziniz.

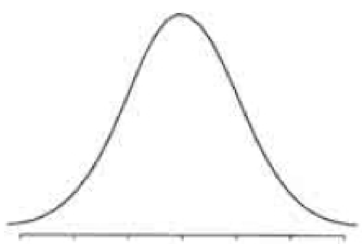
Sınıflar	frekanslar
10 – 15	10
15 – 20	15
20 – 25	20
25 – 30	30
30 – 35	15
35 – 40	10
40 – 45	5
	105

Çözüm.

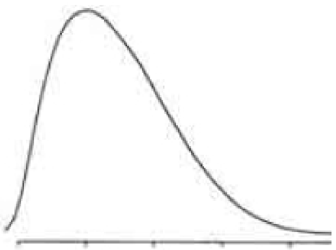
Sınıflar	f	X	c	f / c
10 – 15	10	12.5	5	2
15 – 20	15	17.5	5	3
20 – 25	20	22.5	5	4
25 – 30	30	27.5	5	6
30 – 35	15	32.5	5	3
35 – 40	10	37.5	5	2
40 – 45	5	42.5	5	1
	105			



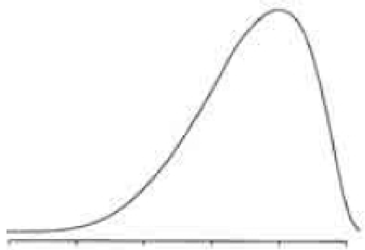
- Histogram ve frekans poligonu sürekli değişkenler için uygun grafiklerdir. Eğer gözlem sayısı artarken sınıf aralığı sonsuz küçültülürse, frekans poligonu bir frekans eğrisi şekline dönüşür.Uygulamada sıkça karşılaşılan frekans eğrileri aşağıdaki gibidir:



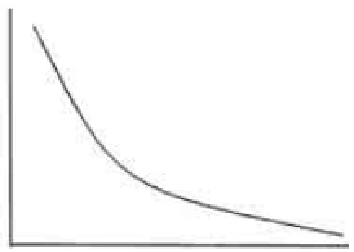
(a) Simetrik Tek Modlu Eğri



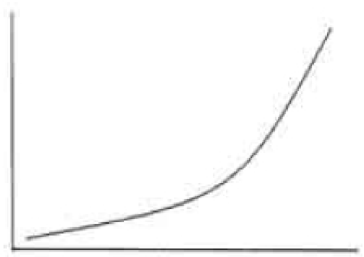
(b) Sağa Eğik Eğri



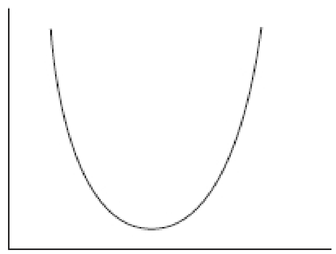
(c) Sola Eğik Eğri



(d) Ters J Eğrisi



(e) J Eğrisi



(f) U Eğrisi

2.3 Birikimli Frekans Dağılımları

- Bir frekans dağılımında, her sınıfın frekansına bir önceki sınıfın frekansı eklenerek oluşturulan seriye "birikimli seri", bu tür oluşturulan frekanslara da "birikimli frekanslar" adı verilir. Birikimli seriler, küçükten büyüğe ya da büyükten küçüğe doğru oluşturulabilirler. Eğer birikimli seriler küçükten büyüğe doğru oluşturulmuşsa "-den az", büyükten küçüğe doğru oluşturulmuşsa "-den çok" olarak isimlendirilirler
- **Örnek.** Aşağıda verilen frekans serisi için -den az ve -den çok serilerini oluşturunuz.

X	f
5	3
10	5
15	8
20	6
25	3
30	5
	30

- **Çözüm.**

X	f	(-den az)	(-den çok)
5	3	3	30
10	5	8	27
15	8	16	22
20	6	22	14
25	3	25	8
30	5	30	5
	30		

- **Örnek.** Aşağıda verilen frekans serisi için

Sınıflar	Frekans
0 – 10	3
10 – 20	5
20 – 30	8
30 – 40	12
40 – 50	7
50 – 60	4
60 – 70	1

  - -den az ve -den çok birikimli serilerini bulunuz.
  - gözlem değeri 40 tan az kaç gözlem vardır?
  - gözlem değeri 30 ve 30 dan çok kaç gözlem vardır?
  - gözlem değeri 10 ve 10 dan fazla 60 tan az kaç gözlem vardır?

- **Çözüm.**

Sınıflar	Frekans	-den az	-den çok
$0-10$	3	3	40
$10-20$	5	8	37
$20-30$	8	16	32
$30-40$	12	28	24
$40-50$	7	35	12
$50-60$	4	39	5
$60-70$	1	40	1

  - -den az ve -den çok birikimli serileri yukarıdaki gibidir.
  - gözlem değeri 40 tan az 28 gözlem vardır.
  - gözlem değeri 30 ve 30 dan çok 24 gözlem vardır.
  - gözlem değeri 10 ve 10 dan fazla 60 tan az $5+8+12+7+4=36$ gözlem vardır.

3. Ortalama, Medyan, Mod ve Diğer Merkezi Eğilim Ölçüleri

3.1 Toplam Gösterimi

- $X_1 + X_2 + \cdots + X_N$ toplamı kısaca $\sum_{j=1}^{N} X_j$ şeklinde gösterilir.
- $\sum_{j=1}^{N} X_j Y_j = X_1 Y_1 + X_2 Y_2 + \cdots + X_N Y_N$
- $\sum_{j=1}^{N} aX_j = aX_1 + aX_2 + \cdots + aX_N = a\left(X_1 + X_2 + \cdots + X_N\right) = a\sum_{j=1}^{N} X_j$
- a, b ve c sabitler olmak üzere
$\sum_{j=1}^{N} \left(aX_j + bY_j - cZ_j\right) = a\sum_{j=1}^{N} X_j + b\sum_{j=1}^{N} Y_j - c\sum_{j=1}^{N} Z_j$

3.2 Aritmetik Ortalama

Ortalama bir seriyi temsil etmeye yarayan ölçüdür öyle ki ortalama seriyi oluşturan elemanların en çok hangi eleman etrafında yoğunlaştığını gösterir. Bu yüzden ortalamalara **merkezi eğilim ölçüleri** de denir. Ortalama ile bir seri özetlenir ve iki seri arasında kıyaslama yapılır. Ana çizgileriyle ortalamalar, duyarlı ve duyarlı olmayan ortalamalar olmak üzere, iki ana başlık altında incelenebilir. Duyarlı ortalamalar, serideki tüm gözlem değerlerinden etkilenen ortalamalardır. Burada duyarlı ortalamalardan sadece aritmetik, geometrik, harmonik ve kareli (kök ortalama kare) ortalamalar ele alınacaktır. Duyarlı olmayan ortalamalardan ise medyan ve mod incelenecektir.
Ortalamanın merkezi bir değer olması ortalamanın

$$X_{\text{min}} \leq \text{Ortalama} \leq X_{\text{max}}$$

şeklinde olmasını gerektirir.
En kolay hesaplanan ve en çok kullanılan ortalama, aritmetik ortalamadır. Eğer ne tür olduğu belirtilmeden bir ortalamadan söz ediliyorsa, muhtemelen kastedilen aritmetik ortalamadır. Aritmetik ortalamaların bir çok avantajının yanında en önemli dezavantajı serideki aşırı değerlerden doğrudan etkileniyor olmasıdır. Yani seride çok büyük değerler varsa bu durumda ortalama seriyi tam anlamıyla temsil etmeyecektir.

- N tane $X_1, X_2, ..., X_N$ sayılarının aritmetik ortalaması kısacaca $\overline{X}$ ile gösterilir ve

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_N}{N} = \frac{\sum_{j=1}^{N} X_j}{N}$$

ile bulunur.
- **Örnek.** $8, 3, 5, 12, 10$ sayılarının aritmetik ortalamasını bulunuz.
- **Çözüm.**

$$\overline{X} = \frac{8+3+5+12+10}{5} = 7.6$$

- **Örnek.** Aşağıda 30 internet kullanıcısının haftalık internete bağlanma süreleri (saat olarak) verilmiştir.Bu 30 değerin ortalamasını bulunuz.

3	4	4	5	5	5	5	5	5	6
6	6	6	7	7	7	7	7	8	8
9	10	10	10	10	10	10	12	55	60

- **Çözüm.** Aritmetik ortalama

$$\overline{X} = \frac{3 + 2\cdot 4 + 6\cdot 5 + 4\cdot 6 + 5\cdot 7 + 2\cdot 8 + 9 + 6\cdot 10 + 12 + 55 + 60}{30} = 10.4$$

tür. Dikkat edilirse ortalama 10.4 tüm saat süreleri için tipik bir değer değildir. Yukarıdaki 30 değerden 21 i yani büyük bir kısmı tek basamaklıdır.

- Eğer $X_1, X_2, ..., X_N$ sayılarının frekansları sırasıyla $f_1, f_2, ..., f_n$ ise bu durumda aritmetik ortalama

$$\overline{X} = \frac{f_1X_1 + f_2X_2 + \cdots + f_nX_N}{f_1 + f_2 + \cdots + f_n} = \frac{\sum_{j=1}^{N} f_jX_j}{\sum_{j=1}^{N} f_j}$$

ile bulunur.

- **Örnek.**

Gözlem değerleri (**X**)	Frekans (**f**)
5	3
10	5
15	8
20	12
25	7
30	4

şeklinde verilen frekans serisinin aritmetik ortalamasını bulunuz.

- **Çözüm.**

$$\overline{X} = \frac{5 \cdot 3 + 10 \cdot 5 + 15 \cdot 8 + 20 \cdot 12 + 25 \cdot 7 + 30 \cdot 4}{3 + 5 + 8 + 12 + 7 + 4} = 18.46$$

- Sınıflandırılmış serilerde aritmetik ortalama bulunurken sınıf numaraları (yani sınıf orta noktaları) bulunur ve bu değerler ile herbir sınıfın frekansı çarpılır, elde edilen bu değerler toplanır ve bu toplam değer toplam frekansa bölünür.
- **Örnek.**

Sınıflar	**Frekans** (f)
$0-10$	3
$10-20$	5
$20-30$	8
$30-40$	12
$40-50$	7
$50-60$	4
$60-70$	1

şeklinde verilen sınıflandırılmış serinin aritmetik ortalamasını bulunuz.

- **Çözüm.**Önce sınıf numaraları (yani sınıf orta noktaları) bulunur.

Sınıflar	**Frekans** (f)	**Sınıf Numarası**
$0-10$	3	5
$10-20$	5	15
$20-30$	8	25
$30-40$	12	35
$40-50$	7	45
$50-60$	4	55
$60-70$	1	65

Bu durumda aritmetik ortalama

$$\overline{X} = \frac{3 \cdot 5 + 5 \cdot 15 + 8 \cdot 25 + 12 \cdot 35 + 7 \cdot 45 + 4 \cdot 55 + 1 \cdot 65}{3 + 5 + 8 + 12 + 7 + 4 + 1} = 32.75$$

şeklinde bulunur.

3.3 Ağırlıklı Aritmetik Ortalama

- Bazen $X_1, X_2, ..., X_N$ sayıları arasında önem farkı ve ağırlık farkı olabilir. Bu sayılar sırasıyla $w_1, w_2, ..., w_N$ ağırlıklarına sahip olsunlar. Bu durumda aritmetik ortalama ağırlıklı aritmetik ortalama adını alır ve

$$\overline{X} = \frac{w_1X_1 + w_2X_2 + \cdots + w_NX_N}{w_1 + w_2 + \cdots + w_N} = \frac{\sum wX}{\sum w}$$

ile bulunur. Dikkat edilirse ağırlıklı aritmetik ortalama hesaplanırken $w_1, w_2, ..., w_N$ değerleri frekans değerleri gibi düşünülerek işlem yapılmıştır.

- **Örnek.** Bir öğrenci bir dersin iki vizesinden 70 ve 90, final sınavından ise 85 almıştır. Final sınavı ara sınavlara göre 3 kez ağırlğa sahip olduğuna göre bu öğrencinin ortalama notunu bulunuz.
- **Çözüm.** Ağırlıklı aritmetik ortalama formülünden

$$\overline{X} = \frac{1 \cdot 70 + 1 \cdot 90 + 3 \cdot 85}{1 + 1 + 3} = 83$$

- **Örnek.** Bir öğrencinin matematik, fizik, ingilizce ve sağlık bilimleri derslerinden aldığı final notları 82, 86, 90 ve 70 tir. Eğer bu derslere ait krediler sırasıyla 3, 5, 3 ve 1 ise ortalama notu hesaplayınız.
- **Çözüm.** Ağırlıklı aritmetik ortalama formülünden

$$\overline{X} = \frac{3 \cdot 82 + 5 \cdot 86 + 3 \cdot 90 + 1 \cdot 70}{3 + 5 + 3 + 1} = 84.67$$

- **Örnek.** 80 kişilik çalışanı olan bir şirkette, çalışanlardan 60'ı saatte 10\$ ve 20'si saatte 13\$ kazanmaktadır. Saatteki ortalama kazancı hesaplayınız.
- **Çözüm.**

$$\overline{X} = \frac{60 \cdot 10 + 20 \cdot 13}{60 + 20} = 10.75\$$$

3.4 Aritmetik Ortalamanın Özellikleri

1. Terimlerin aritmetik ortalamadan cebirsel sapmalarının toplamı sıfırdır.
2. Terimlerin herhangi bir a sayısından sapmalarının karelerinin toplamı ancak ve ancak $a = \overline{X}$ ise minimumdur. Yani terimlerin herhangi bir a sayısından sapmalarının karelerinin toplamı $a = \overline{X}$ ise minimumdur.
3. Eğer f_1 adet sayının aritmetik ortalaması m_1, f_2 adet sayının aritmetik ortalaması m_2, ..., f_n adet sayının aritmetik ortalaması m_n ise bütün sayıların aritmetik ortalaması

$$\overline{X} = \frac{f_1 \cdot m_1 + f_2 \cdot m_2 + \cdots + f_n \cdot m_n}{f_1 + f_2 + \cdots + f_n}$$

dir. Yani $\overline{X}$, bütün aritmetik ortalamaların ağırlıklı aritmetik ortalamasıdır.

4. Bir serinin tüm terimlerine aynı sayı eklenirse (çıkarılırsa) aritmetik ortalama eklenen (çıkarılan) sayı kadar artar (azalır).
5. Bir serinin tüm terimleri aynı sayı ile çarpılırsa (bölünürse) aritmetik ortalama çarpılan (bölünen) sayı ile orantılı olarak büyür (küçülür).

Örnek. Yukarıdaki özellikleri aşağıdaki seri üzerinde açıklayalım.

X
3
4
5
7
11

Serinin aritmetik ortalaması $\overline{X} = \dfrac{3+4+5+7+11}{5} = 6$ dır.

- Terimlerin aritmetik ortalamadan cebirsel sapmalarına bakalım.

$X_j - \overline{X}$
$3 - 6 = -3$
$4 - 6 = -2$
$5 - 6 = -1$
$7 - 6 = 1$
$11 - 6 = 5$
$\sum = 0$

Demek ki terimlerin aritmetik ortamadan cebirsel sapmaları toplamı sıfırdır.

- $a = 6$ ve $a = 7$ için terimlerin a sayısından sapmalarının karelerinin toplamını inceleyelim.

$(X_j - 6)^2$	$(X_j - 7)^2$
$(3-6)^2 = 9$	$(3-7)^2 = 16$
$(4-6)^2 = 4$	$(4-7)^2 = 9$
$(5-6)^2 = 1$	$(5-7)^2 = 4$
$(7-6)^2 = 1$	$(7-7)^2 = 0$
$(11-6)^2 = 25$	$(11-7)^2 = 16$
$\sum = 40$	$\sum = 45$

Görüldüğü gibi terimlerin $a = 6 = \overline{X}$ sayısından sapmalarının karelerinin toplamı daha küçüktür.

- Bütün terimlere 3 ekleyerek yeni bir seri oluşturalım.

X
6
7
8
10
14

Bu serinin aritmetik ortalaması $\overline{X} = \dfrac{6+7+8+10+14}{5} = 9$ dur. Daha önceki serinin ortalaması 6 iken 3 artmıştır.

- Serinin tüm terimleri 3 ile çarpalım. Aşağıdaki seri oluşur:

X
9
12
15
21
33

Bu serinin airtmetik ortalaması $\overline{X} = \dfrac{9+12+15+21+33}{5} = 18$ dir ve en başta verilen serinin aritmetik ortalamasının tam 3 katıdır.

3.5 Geometrik Ortalama

- N tane $X_1, X_2, ..., X_N$ sayısı pozitif olmak üzere bu sayıların geometrik ortalaması
$$G = \sqrt[n]{X_1 \cdot X_2 \cdots X_N}$$
şeklinde tanımlanır.
- **Örnek.** $2, 4, 8$ sayılarının geometrik ortalamasını bulunuz.
- **Çözüm.** Üç sayı olduğundan $n = 3$ tür. Buna göre $G = \sqrt[3]{2 \cdot 4 \cdot 8} = 4$ olur.
- **Örnek.** $3, 5, 6, 6, 7, 10, 12$ sayılarının geometrik ortalamasını bulunuz.
- **Çözüm.** Yedi sayı olduğundan $n = 7$ dır. Buna göre $G = \sqrt[7]{3 \cdot 5 \cdot 6^2 \cdot 7 \cdot 10 \cdot 12} = 6.43$ olur.

- **Örnek.** Bir yıl boyunca bir litre sütün fiyatının bir ekmek fiyatına oranı 3, bir yıl sonra ise 2 dir. Buna göre
  - İki yıl için sütün fiyatının ekmeğin fiyatına oranının aritmetik ortalamasını bulunuz.
  - İki yıl için ekmeğin fiyatının sütün fiyatına oranının aritmetik ortalamasını bulunuz.
  - Ortalama oran için aritmetik ortalamayı kullanmak uygun mudur?
  - Ortalama oran için geometrik ortalamayı kullanmak uygun mudur?
- **Çözüm.**
  - $\frac{3+2}{2} = 2.5$
  - $\frac{\frac{1}{3}+\frac{1}{2}}{2} = 0.42$
  - Eğer burada aritmetik ortalama kullanmak uygun olsaydı yukarıdaki oranlar karşılıklı olmalıydı. Oysa $\frac{1}{0.42} = 2.38 \neq 2.5$ Demek ki aritmetik ortalama kullanmak uygun değildir.
  - İki yıl için geometrik ortalamalar $\sqrt{2 \cdot 3} = \sqrt{6}$ ve $\sqrt{\frac{1}{3} \cdot \frac{1}{2}} = \frac{1}{\sqrt{6}}$ dır. Bu ortalamalar karşılıklı olduğundan geometrik ortalama kullanmak uygundur.

- **Örnek.** Bir kültür içinde bulunan bakteri sayısı 3 gün içinde 1000 den 4000 e çıkmıştır. Günlük ortalama artış oranı kaçtır?
- **Çözüm.** Artış oranı r olsun.
  - Birinci günün sonunda $1000 + 1000r = 1000(1+r)$
  - İkinci günün sonunda $1000(1+r) + 1000(1+r)r = 1000(1+r)^2$
  - Üçüncü günün sonunda $1000(1+r)^3$ olur.
  - $1000(1+r)^3 = 4000$ den $1 + r = \sqrt[3]{4}$, $r = 0.59$ olur. Yani günlük artış oranı %59 dur.

3.6 Harmonik Ortalama

- N tane $X_1, X_2, ..., X_N$ sayılarının harmonik ortalaması

$$H = \frac{N}{\frac{1}{X_1} + \frac{1}{X_2} + \cdots + \frac{1}{X_N}} = \frac{N}{\sum_{j=1}^{N} \frac{1}{X_j}}$$

  şeklinde tanımlanır.
- **Örnek.** 2, 4, 8 sayılarının harmonik ortalamasını bulunuz.
- **Çözüm.** $H = \frac{3}{\frac{1}{2}+\frac{1}{4}+\frac{1}{8}} = 3.43$ olur.
- Sınıflandırılmış serilerde harmonik ortalama hesaplanırken sınıf orta noktaları dikkate alınır.

- **Örnek.** Aşağıdaki serinin harmonik ortalamasını bulunuz.

Sınıflar	f
0 - 4	1
4 - 8	4
8 - 12	8
12 - 16	5
16 - 20	2
	20

- **Çözüm.** Sınıf orta noktaları dikkate alınır. Bu durumda harmonik ortalama

$$H = \frac{20}{\frac{1}{2} + \frac{1}{6} \times 4 + \frac{1}{10} \times 8 + \frac{1}{14} \times 5 + \frac{1}{18} \times 2} = 8.21$$

  bulunur.
- Aritmetik, geometrik ve harmonik ortalama arasında daima

$$H \leq G \leq \overline{X}$$

  ilişkisi vardır. Eşitlik durumları sayıların tümünün birbirine eşit olduğu durumda gerçekleşir.

3.7 Kareli Ortalama veya Kök Ortalama Kare

- $X_1, X_2, ..., X_N$ sayılarının kareli ortalaması (kök ortalama kare olarak ta adlandırılır) ve

$$K = \sqrt{\frac{X_1^2 + X_2^2 + \cdots + X_N^2}{N}} = \sqrt{\frac{\sum_{j=1}^{N} X_j^2}{N}}$$

ile bulunur.

- **Örnek.** $1, 3, 4, 5, 7$ sayılarının kareli ortalamasını bulunuz.
- **Çözüm.**

$$K = \sqrt{\frac{1^2 + 3^2 + 4^2 + 5^2 + 7^2}{5}} = 4.47$$

- **Örnek.** Aşağıdaki serinin kareli ortalamasını hesaplayınız.

Sınıflar	f
0 - 4	1
4 - 8	4
8 - 12	8
12 - 16	5
16 - 20	2
	20

- **Çözüm.** Sınıf orta noktaları kullanılarak

$$K = \sqrt{\frac{2^2 \cdot 1 + 6^2 \cdot 4 + 10^2 \cdot 8 + 14^2 \cdot 5 + 18^2 \cdot 2}{20}} = 11.35$$

3.8 Medyan

- **Tanım.** Bir istatistik serisinde tam ortaya düşen dolayısıyla seriyi iki eşit kısma bölen gözlem değerine medyan denir. Medyan uygulamada, ilgilenilen seride aşırı değerlerin varlığı durumunda uygun sonuçlar veren bir merkezi eğilim ölçüsüdür.
- **Örnek.** $3, 4, 4, 5, 6, 8, 8, 8, 10$ sayılarının medyanı 6 dır.
- **Örnek.** $5, 5, 7, 9, 11, 12, 15, 18$ sayılarının medyanı $\dfrac{9+11}{2} = 10$ dur.
- **Örnek.** Aşağıdaki 30 değerin Medyanını bulunuz.

3	4	4	5	5	5	5	5	5	6
6	6	6	7	7	7	7	7	8	8
9	10	10	10	10	10	10	12	55	60

- **Çözüm.** 30 tane değer olduğuna göre ortadaki elemanlar 15. ve 16. elemanlardır. Bunların aritmetik ortalaması $\dfrac{7+7}{2} = 7$ Medyandır.

- Sınıflandırılmış serilerde Medyan hesaplanırken önce **medyan sınıfı** bulunur. Medyan sınıfı $\dfrac{N}{2}$. elemanın olduğu sınıftır. Bu sınıf yardımıyla aşağıdaki değerler bulunur:
  - L_1 =Medyan sınıfının alt sınırı
  - $(\sum f)_1$ =Medyan sınıfından önceki sınıfların frekansları toplamı
  - f_{med} =medyan sınıfının frekansı
  - c =Medyan sınıfının uzunluğu
- Bulunan bu değerler

$$Med = L_1 + \left(\frac{\frac{N}{2} - (\sum f)_1}{f_{med}} \right) c$$

formülünde yazılır. Bu değer medyan sınıfında olmalıdır.

- **Örnek.** Aşağıdaki serinin medyanını hesaplayınız.

Sınıflar	f
10 - 14	3
14 - 18	4
18 - 22	8
22 - 26	6
26 - 30	1
	22

- **Çözüm.** $N = 22$ olduğundan $\frac{22}{2} = 11$. elemanın olduğu sınıf yani 3. sınıf medyan sınıfıdır. Bu sınıf yardımıyla $L_1 = 18$, $(\sum f)_1 = 3 + 4 = 7$, $f_{med} = 8$ ve $c = 22 - 18 = 4$ olduğundan

$$Med = L_1 + \left(\frac{\frac{N}{2} - (\sum f)_1}{f_{med}} \right) c = 18 + \left(\frac{11 - 7}{8} \right) 4 = 20$$

bulunur.

- **Örnek.** Aşağıdaki serinin medyanını hesaplayınız.

Sınıflar	f
0 - 4	4
4 - 8	6
8 - 12	10
12 - 16	8
16 - 20	6
20 - 24	3
24 - 28	3
	40

- **Çözüm.** $N = 40$ olduğundan $\frac{40}{2} = 20$. elemanın olduğu sınıf yani 3. sınıf medyan sınıfıdır. Bu sınıf yardımıyla $L_1 = 8$, $(\sum f)_1 = 4 + 6 = 10$, $f_{med} = 10$ ve $c = 12 - 8 = 4$ olduğundan

$$Med = L_1 + \left(\frac{\frac{N}{2} - (\sum f)_1}{f_{med}} \right) c = 8 + \left(\frac{20 - 10}{10} \right) 4 = 12$$

bulunur.

3.9 Mod

- **Tanım.** Bir istatistik serisinin modu en büyük frekansa sahip sınıfın değeridir. Yani bir seride en çok tekrarlanan değere mod adı verilir. Mod bazen olmayabilir veya birden fazla olabilir. Eğer tüm frekansalar 1 ise bu durumda mod yoktur.
- **Örnek.** $10, 12, 14, 20$ serisinin modu yoktur.
- **Örnek.** $2, 2, 5, 7, 9, 9, 9, 10, 10, 11, 12$ serisinin modu 9 dur.
- **Örnek.** $2, 3, 4, 4, 4, 5, 5, 7, 7, 7, 9$ serisinin iki modu vardır. Bunlar 4 ve 7 dir.
- **Örnek.** Aşağıdaki serinin modunu bulunuz.

x	f
10	2
13	5
14	8
16	7
19	3
	25

- **Çözüm.** Frekansı en büyük olan değer 14 olduğundan mod 14 tür.

- Sınıflandırılmış serilerde mod hesaplanırken önce **mod sınıfı** bulunur. Mod sınıfı en büyük frekansa karşılık gelen sınıftır. Bu sınıf yardımıyla aşağıdaki değerler bulunur:
  - L_1 =Mod sınıfının alt sınırı
  - Δ_1 =Mod sınıfı ile bir üst sınıfın frekansları farkı
  - Δ_2 =Mod sınıfı ile bir alt sınıfın frekansları farkı
  - c =Medyan sınıfının uzunluğu
- Bulunan bu değerler

$$Mod = L_1 + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2} \right) c$$

formülünde yazılır. Bu değer mod sınıfında olmalıdır.

- **Örnek.** Aşağıdaki serinin modunu hesaplayınız.

Sınıflar	f
10 - 14	2
14 - 18	4
(18 - 22)	7
22 - 26	5
26 - 30	3
	21

- **Çözüm.** Frekansı en büyük olan sınıf $18-22$ sınıfıdır. Bundan dolayı bu sınıf mod sınıfıdır. Bu sınıf yardımıyla $L_1 = 18$, $\Delta_1 = 7-4 = 3$, $\Delta_2 = 7-5 = 2$, $c = 22-18 = 4$ formülde yazılırsa

$$Mod = L_1 + \left(\frac{\Delta_1}{\Delta_1+\Delta_2}\right) c = 18 + \left(\frac{3}{3+2}\right) 4 = 20.4$$

- **Örnek.** Aşağıdaki seri veriliyor. Serinin modu 9 olduğuna göre $10-13$ sınıfının frekansını bulunuz.

Sınıflar	Frekans
$1-4$	$-$
$4-7$	6
$7-10$	10
$10-13$	x
$13-16$	$-$

- **Çözüm.** Serinin modu 9 olduğuna göre ve bu değer $7-10$ aralığında olduğundan bu sınıf mod sınıfıdır. Bu sınıf yardımıyla $L_1 = 7$, $\Delta_1 = 10-6 = 4$, $\Delta_2 = 10-x$, $c = 10-7 = 3$ değerleri formülde yazılırsa

$$\begin{aligned} 9 &= 7 + \left(\frac{4}{4+10-x}\right) 3 \\ 2 &= \left(\frac{4}{14-x}\right) 3 \\ x &= 8 \end{aligned}$$

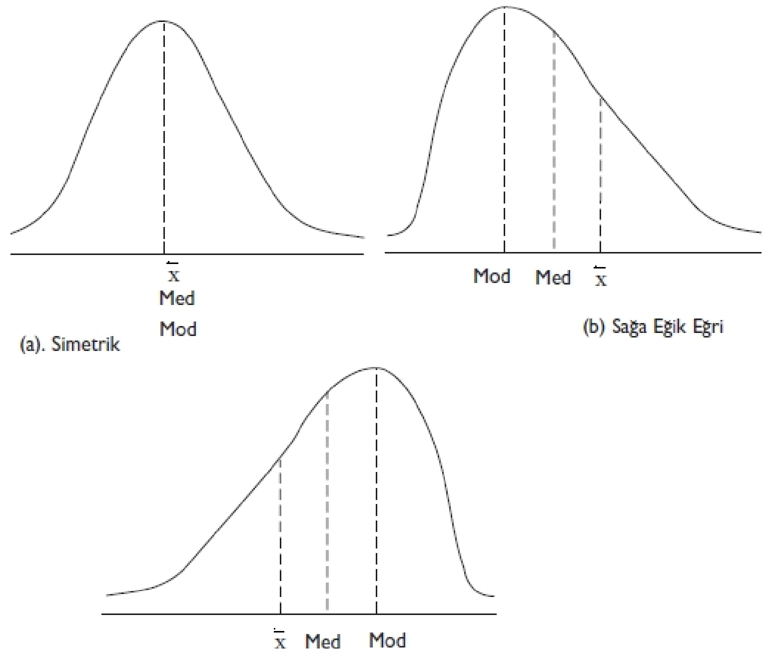
bulunur.

3.10 Ortalamaların Genel Özellikleri

1. Geometrik ortalama terimlerin anlık ve anormal artışlarına karşılık aritmetik ortalamaya göre daha az duyarlıdır. Bu yüzden gerçeği daha iyi yansıtır.
2. Bazı terimler negatif veya sıfır ise geometrik ortalama kullanılmaz.
3. Serideki terimlerden biri bile sıfır ise harmonik ortalama sıfır olur.
4. Terimlerin bazıları negatif ise kareli ortalama kullanılır.
5. Diğer ortalamaların aksine sınıflandırılmış serilerin medyan hesabında sınıf aralıklarının eşit olması gerekmez.
6. Ortalamalar içinde mod en temsili olanıdır.
7. Orta derecede çarpık (yani simetrik olmayan) tek modlu frekans eğrilerinde aşağıdaki denemeli ilişki vardır:

$$\overline{X} - \text{mod} = 3\left(\overline{X} - \text{med}\right)$$

- Sağa ve sola çarpık frekans eğrilerinde medyan daima ortalama ve mod arasındadır. Simetrik eğrilerde ise üçü de üst üstedir.



3.11 Çeyreklikler, Ondabirlikler ve Yüzdebirlikler

- **Tanım.** Büyüklük sırasına göre düzenlenmiş bir veri kümesini iki eşit parçaya bölen değer medyandı. Buna benzer olarak dağılımı dört eşit parçaya bölen Q_1, Q_2, Q_3 değerlerine sırasıyla birinci, ikinci ve üçüncü çeyreklikler (veya kartiller) denir. Bu değerlerden Q_2 medyana eşittir. Seriyi 10 eşit parçaya bölen $D_1, D_2, ..., D_9$ değerlerine ondabirlikler (desiller), 100 eşit parçaya bölen $P_1, P_2, ..., P_{99}$ değerlerinde de yüzdebirlikler (santiller) adı verilir. Beşinci ondabirlik D_5 ile ellinci yüzdebirlik P_{50} medyana karşılık gelir. Ayrıca 25. ve 75. yüzde birlikler birinci ve üçüncü çeyrekliklere karşılık gelir.
- Basit serilerde birinci çeyreklik Q_1, $\frac{N+2}{4}$. terim, üçüncü çeyreklik Q_3 ise $\frac{3N+2}{4}$.terimdir. Bu terimler tam sayı veya buçuklu ise medyanda olduğu gibi işlem yapılır. Eğer buçuklu olmayan bir ondalıklı sayı ise en yakın tamsayıya yuvarlanarak işlem yapılır.

- **Örnek.** Aşağıdaki basit serinin çeyrekliklerini bulunuz.

X
11
22
34
46
57

- **Çözüm.** $N = 5$ olduğuna göre $Q_1 = \frac{N+2}{4} = \frac{7}{4} = 1.75 \approx 2$. terim, $Q_3 = \frac{3N+2}{4} = \frac{17}{4} = 4.25 \approx 4$. terimdir. Demek ki $Q_1 = 22$, $Q_3 = 46$ dır. Q_2 ise medyandır.

- **Örnek.** Aşağıdaki basit serinin çeyrekliklerini bulunuz.

X
12
23
36
49

- **Çözüm.** $N = 4$ olduğuna göre $Q_1 = \frac{N+2}{4} = \frac{6}{4} = 1.5$ olduğundan 1. ve 2. terimlerin aritmetik ortalaması birinci çeyreklik, $Q_3 = \frac{3N+2}{4} = \frac{14}{4} = 3.5$ olduğundan 3. ve 4. terimlerin aritmetik ortalaması üçüncü çeyrekliktir. Demek ki $Q_1 = \frac{12+23}{2} = 17.5$ ve , $Q_3 = \frac{36+49}{2} = 42.5$ dir. Q_2 ise medyandır.

- Sınıflandırılmış serilerde çeyreklikler bulunurken $\frac{N}{4}$. ve $\frac{3N}{4}$. elemanların bulunduğu sınıflar bulunur. Bu sınıflar çeyrekliklerin sınıflarıdır. Bu sınıflar yardımıyla aşağıdaki değerler bulunur:
  - L_1 =Çeyreklik sınıfının alt sınırı
  - $(\sum f)_1$ =Çeyreklik sınıfından önceki sınıfların frekansları toplamı
  - f_{Q_1}, f_{Q_3} =Çeyreklik sınıflarının frekansı
  - c =Çeyreklik sınıflarının uzunluğu
- Bulunan bu değerler

$$
\begin{aligned}
Q_1 &= L_1 + \left(\frac{\frac{N}{4} - (\sum f)_1}{f_{Q_1}} \right) c \\
Q_3 &= L_1 + \left(\frac{\frac{3N}{4} - (\sum f)_1}{f_{Q_3}} \right) c
\end{aligned}
$$

formüllerinde yazılır. Bu değerler çeyreklik sınıflarında olmalıdır.

- **Örnek.** Aşağıdaki sınıflandırılmış serinin çeyrekliklerini bulunuz.

Sınıflar	f
$0-2$	4
$2-4$	3
$4-6$	1
$6-8$	2

- **Çözüm.** $\dfrac{N}{4} = \dfrac{10}{4} = 2.5$'inci terimin olduğu sınıf yani $0-2$ sınıfı birinci çeyreklik sınıfıdır. Bu sınıf yardımıyla
  - $L_1 = 0$
  - $(\sum f)_1 = 0$
  - $f_{Q_1} = 4$
  - $c = 2 - 0 = 2$
- Bulunan bu değerler formülde yazılırsa ilk çeyreklik

$$Q_1 = L_1 + \left(\frac{\frac{N}{4} - (\sum f)_1}{f_{Q_1}} \right) c = 0 + \left(\frac{2.5 - 0}{4} \right) 2 = 1.25$$

olarak bulunur.

- $\dfrac{3N}{4} = \dfrac{30}{4} = 7.5$'inci elemanın olduğu $4-6$ sınıfı üçüncü çeyreklik sınıfıdır. Bu sınıf yardımıyla
  - $L_1 = 4$
  - $(\sum f)_1 = 4 + 3 = 7$
  - $f_{Q_3} = 1$
  - $c = 6 - 4 = 2$
- Bulunan bu değerler formülde yazılırsa üçüncü çeyreklik

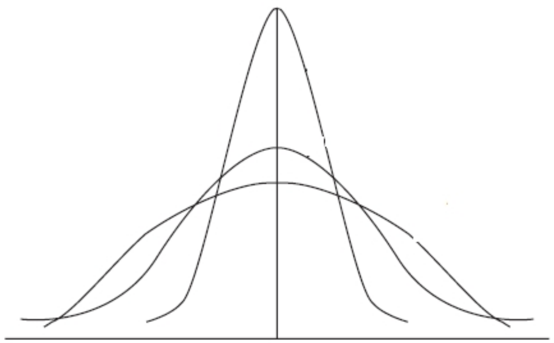
$$Q_3 = L_1 + \left(\frac{\frac{3N}{4} - (\sum f)_1}{f_{Q_3}} \right) c = 4 + \left(\frac{7.5 - 7}{1} \right) 2 = 5$$

olarak bulunur.
- İkinci çeyreklik ise medyandır.

4. Standart Sapma ve Diğer Yayılım Ölçüleri

4.1 Yayılım veya Değişim

- **Tanım.** Sayısal verilerin bir ortalama etrafında yayılma eğiliminin derecesine o verinin yayılımı veya değişimi denir.
- Ortalamalar serileri özetlemek için gerekli olmakla birlikte terimlerin birbirine yakınlık derecesini temsil etmede yeterli değildir. İki serinin ortalaması aynı olduğu halde ikisi de gerçekte çok farklı bir anlama gelebilir.
- Değişkenlik bir serideki terimlerin ortalamadan ne ölçüde uzaklaştıklarını gösteren bir ölçüdür. Değişkenlik arttıkça serinin x ekseninde kapladığı alan genişler.



4.2 Açıklık

- **Tanım.** Bir serideki maksimum ve minimum değerler arasındaki farka açıklık denir.
- **Örnek.** Aşağıdaki serinin açıklığını bulunuz.

X
2
3
17
28

- **Çözüm.** Serideki en büyük değer 28, en küçük değer 2 olduğundan serininin açıklığı $28 - 2 = 26$ dır.

4.3 Ortalama Sapma

- $X_1, X_2, ..., X_N$ değerlerinin ortalama sapması OS ile gösterilir ve

$$OS = \frac{\sum_{j=1}^{N} |X_j - \overline{X}|}{N}$$

olarak tanımlanır.

- **Örnek.** $2, 3, 6, 8, 11$ serisinin ortalama sapmasını hesaplayınız.
- **Çözüm.** $\overline{X} = \frac{2+3+6+8+11}{5} = 6$ dır. Buna göre

$$OS = \frac{\sum_{j=1}^{5} |X_j - 6|}{5} = \frac{|2-6| + |3-6| + |6-6| + |8-6| + |11-6|}{5} = 2.8$$

olarak bulunur.

- **Örnek.** Aşağıdaki serinin ortalama sapmasını hesaplayınız.

Sınıflar	f
$0-2$	3
$2-4$	7
$4-6$	4
$6-8$	2

- **Çözüm.** Sınıf orta noktaları dikkate alınır.

Sınıflar	f	X
$0-2$	3	1
$2-4$	7	3
$4-6$	4	5
$6-8$	2	7

- $\overline{X} = \dfrac{3\cdot 1 + 7\cdot 3 + 4\cdot 5 + 2\cdot 7}{16} = 3.62$ dir. Buna göre

$$OS = \frac{|1-3.62|\cdot 3 + |3-3.62|\cdot 7 + |5-3.62|\cdot 4 + |7-3.62|\cdot 2}{16} = 1.53$$

olarak bulunur.

4.4 Yarı Çeyreklikler Arası Açıklık

- Serinin genişliğinin serinin iki ucunda yer alan aşırı değerlerden etkilenmesi sakıncasını gidermek için kullanılan bir yayılım ölçüsüdür ve

$$Q = \frac{Q_3 - Q_1}{2}$$

şeklinde tanımlanır. Burada Q_1 ve Q_3 sırasıyla birinci ve üçüncü çeyreklerdir.

- **Örnek.** Aşağıdaki serinin yarı çeyreklikler arası açıklığını bulunuz.

X	f
1	2
5	7
7	9
12	2

- **Çözüm.** $Q_1 = 5$ ve $Q_3 = 7$ olduğundan $Q = \dfrac{Q_3 - Q_1}{2} = \dfrac{7-5}{2} = 1$ olur.

- **Örnek.** Aşağıdaki serinin yarı çeyreklikler arası açıklığını bulunuz.

Sınıflar	f
$0-2$	3
$2-4$	7
$4-6$	4
$6-8$	2

- **Çözüm.** $Q_1 = 2.29$ ve $Q_3 = 5$ olduğundan $Q = \dfrac{5-2.29}{2} = 1.36$ olur.

4.5 Standart Sapma

- $X_1, X_2, ..., X_N$ değerlerinin standart sapması σ (sigma) ile gösterilir ve

$$\sigma = \sqrt{\frac{\sum_{j=1}^{N} \left(X_j - \overline{X}\right)^2}{N}}$$

ile hesaplanır.

- **Örnek.** Aşağıdaki serinin standart sapmasını hesaplayınız.

x
1
4
5
7
9
10

- **Çözüm.** $\overline{X} = 6$ olduğundan standart sapma

$$\sigma = \sqrt{\frac{(1-6)^2 + (4-6)^2 + (5-6)^2 + (7-6)^2 + (9-6)^2 + (10-6)^2}{6}} = 3.06$$

olarak bulunur.

- **Örnek.** Aşağıdaki sınıflandırılmış serinin standart sapmasını hesaplayınız.

Sınıflar	f
0 - 2	2
2 - 4	4
4 - 6	6
6 - 8	5
8 - 10	3
	20

4.6 Varyans

- **Çözüm.** Sınıf orta noktaları dikkate alınmalıdır. Buna göre, $\overline{X} = \frac{106}{20} = 5.3$ tür. Böylece standart sapma

$$\sigma = \sqrt{\frac{(1-5.3)^2 \cdot 2 + (3-5.3)^2 \cdot 4 + (5-5.3)^2 \cdot 6 + (7-5.3)^2 \cdot 5 + (9}{20}} = 2.39$$

olur.

- **Tanım.** Bir veri kümesinin varyansı bu veri kümesinin standart sapmasının karesi olarak tanımlanır ve σ^2 ile gösterilir.
- **Örnek.** Bir önceki örnekte verilen serinin varyansını hesaplayınız.
- **Çözüm.** Bir önceki örnekte standart sapma $\sigma = 2.39$ olarak bulunmuştu. Buna göre varyans $\sigma^2 = 2.39^2 = 5.71$ olacaktır.

4.7 Standart Sapmanın Özellikleri

1. Standart sapma formülünde yerine $\overline{X}$ yerine herhangi bir a ortalaması alınarak başka sapmalar bulunabilir. Ancak bu değerler içinde en küçüğü $a = \overline{X}$ olandır.
2. Bir serinin tüm terimlerine aynı sayıyı eklersek veya çıkarırsak standart sapma değişmez.
3. Bir serinin tüm terimleri c ile çarpılırsa oluşan yeni serinin standart sapması ilk serinin standart sapmasının c katı olur.
4. Aritmetik ortalamaları eşit olan iki seri verilsin. N_1 sayıdan oluşan birinci serinin varyansı σ_1^2, N_2 sayıdan oluşan ikinci serinin varyansı σ_2^2 ise tüm elemanlardan oluşan serinin varyansı
$$\sigma^2 = \frac{N_1 \cdot \sigma_1^2 + N_2 \cdot \sigma_2^2}{N_1 + N_2}$$
ile bulunur. Bu özellik aritmetik ortalamaları eşit ikiden fazla seri için de geçerlidir.
5. Kareli ortalama, aritmetik ortalama ve standart sapma arasında
$$K^2 - \overline{X}^2 = \sigma^2$$
bağıntısı vardır.

- **Örnek.** Bir serinin aritmetik ortalaması $\overline{X} = 100$, standart sapması $\sigma = 12$ olduğuna göre kareli ortalaması kaçtır?
- **Çözüm.**

$$\begin{aligned} K^2 - \overline{X}^2 &= \sigma^2 \\ K^2 - 100^2 &= 12^2 \\ K &= \sqrt{100^2 + 12^2} = 100.72 \end{aligned}$$

olarak bulunur.

4.8 Değişim Katsayısı

- Buraya kadar ele alınan değişkenlik ölçüleri, mutlak değişkenlik ölçüleridir. Bu nedenle farklı ölçü birimlerine göre oluşturulan serilerin değişkenlikleri, bu ölçülerle karşılaştırılamaz. Ayrıca mutlak değişkenlik ölçüleri, seriyi oluşturan gözlem değerlerinin büyüklüklerinin de etkisi altındadır.
- Aşağıdaki iki seriyi göz önüne alalım.

X	Y
10	43
11	48
13	58
14	63
17	73
65	**285**

Buradaki seriler için $\overline{X} = 13, \sigma_x = 2.45, \overline{Y} = 57, \sigma_y = 10.68$ dir. $\sigma_y > \sigma_x$ olması Y serisindeki gözlem değerlerinin X serisindeki gözlem değerlerinden büyük olmasından kaynaklanabilir. Bundan dolayı Y serisinin değişkenliği daha fazladır diyemeyiz. Bunun için farklı bir ölçü birimine gerek vardır.

- **Tanım.** Standart sapması σ ve aritmetik ortalaması $\overline{X}$ olan bir serinin değişim katsayısı yüzde olarak

$$DK = \frac{\sigma}{\overline{X}} \cdot 100$$

olarak tanımlanır.
- Farklı sayıda gözlem değeri içeren veya farklı ölçü birimlerine göre oluşturulmuş serilerin değişkenliklerinin kıyaslanmasında değişim katsayısı kullanılır.
- **Örnek.** Aşağıdaki X ve Y serilerinin değişkenlik katsayılarını bularak hangi serideki değişkenliğin daha fazla olduğunu bulunuz.

X	Y
10	43
11	48
13	58
14	63
17	73
65	**285**

- **Çözüm.** $\overline{X} = 13, \sigma_x = 2.45, \overline{Y} = 57, \sigma_y = 10.68$ olduğundan X serisinin değişkenlik katsayısı

$$DK = \frac{2.45}{13} \cdot 100 = 18.85$$

ve Y serisinin değişkenlik katsayısı

$$DK = \frac{10.68}{57} \cdot 100 = 18.74$$

dir. Demek ki X serisinin değişkenliği daha fazladır.

- **Örnek.** Aşağıdaki seri veriliyor:

Sınıflar	f
0 - 4	1
4 - 8	4
8 - 12	8
12 - 16	5
16 - 20	2
	20

  - Serinin açıklığını,
  - Ortalama sapmasını,
  - Yarı çeyreklikler arasındaki açıklığını,
  - Standart sapmasını,
  - Varyansını,
  - Değişim katsayısını bulunuz.

- **Çözüm.**
  - Serinin açıklığı en büyük değer ile en küçük değerin farkı olan $20 - 0 = 20$ dir.
  - $$\overline{X} = \frac{2 \cdot 1 + 6 \cdot 4 + 10 \cdot 8 + 14 \cdot 5 + 18 \cdot 2}{20} = 10.6$$
    olduğundan ortalama sapma
    $$\begin{aligned} OS &= \frac{|2 - 10.6| \cdot 1 + |6 - 10.6| \cdot 4 + |10 - 10.6| \cdot 8 + |14 - 10.6| \cdot 5 + |18 - \text{:}}{20} \\ &= 3.18 \end{aligned}$$
  - Birinci çeyreklik $Q_1 = 8$ ve üçüncü çeyreklik $Q_3 = 13.6$ olduğundan yarı çeyreklikler arasındaki açıklık $Q = \frac{Q_3 - Q_1}{2} = \frac{13.6 - 8}{2} = 2.8$ olur.
  - Standart sapma
    $$\begin{aligned} \sigma &= \sqrt{\frac{(2 - 10.6)^2 \cdot 1 + (6 - 10.6)^2 \cdot 4 + (10 - 10.6)^2 \cdot 8 + (14 - 10.6)^2 \cdot 5 +}{20}} \\ &= 4.05 \end{aligned}$$
  - Varyans $\sigma^2 = 4.05^2 = 16.40$
  - Değişim katsayısı
    $$DK = \frac{\sigma}{\overline{X}} \cdot 100 = \frac{4.05}{10.6} \cdot 100 = 38.20$$

4.9 Sheppard Varyans Düzeltmesi

- Sınıflandırılmış serilerde standart sapma hesaplanırken sınıf orta noktaları alınarak bir yuvarlama yapılmış olur. Bu yuvarlama sonucunda belli bir hata oluşur. Bu hatayı düzeltmek için
  $$\begin{aligned} \text{Düzeltilmiş varyans} &= \sigma^2 - \frac{c^2}{12} \\ \text{Düzeltilmiş standart sapma} &= \sqrt{\sigma^2 - \frac{c^2}{12}} \end{aligned}$$
  formüllerini kullanırız. Bu formüldeki $\frac{c^2}{12}$ terimine Sheppard varyans düzeltmesi denir. Burada c sınıf aralığıdır.
- **Örnek.** Zekâ bölümü (IQ)
  $$ZB = \frac{\text{Zekâ Yaşı}}{\text{Doğum Yaşı}} \cdot 100$$
  olarak tanımlanır. Örneğin 8 yaşındaki bir çocuk 12 yaşındaki bir çocuğun zekâsına sahipse zekâ bölümü
  $$ZB = \frac{12}{8} \cdot 100 = 150$$
  dir.

- **Örnek.** Aşağıdaki tabloda bir ilkokuldaki 480 çocuğun zekâ bölümü değerleri verilmiştir.

Sınıf orta Değeri	Frekans
70	4
74	9
78	16
82	28
86	45
90	66
94	85
98	72
102	54
106	38
110	27
114	18
118	11
122	5
126	2

  - Varyansı,
  - Düzeltilmiş varyansı,
  - Düzeltilmiş standart sapmayı bulunuz.

- **Çözüm.** $\overline{X} = 95.97$, $\sigma = 10.47$ ve $c = 4$ tür. Buna göre
  - Varyans $\sigma^2 = 10.47^2 = 109.62$
  - Düzeltilmiş varyans $\sigma^2 - \frac{c^2}{12} = 109.62 - \frac{4^2}{12} = 108.29$ dir.
  - Düzeltilmiş standart sapma ise $\sqrt{108.29} = 10.41$ olur.

4.10 Yayılım Ölçüleri Arasındaki Denemeli İlişkiler

- Çarpıklığı orta derecede olan dağılımlar için ortalama sapma (OS) ve standart sapma (σ) ile Yarı çeyrek açıklık (Q) ve standart sapma arasında (σ) arasında sırasıyla aşağıdaki denemeli ilişkiler (yaklaşık eşitlikler) vardır:

$$\begin{aligned} OS &= \frac{4}{5}\sigma = 0.8\sigma \\ Q &= \frac{2}{3}\sigma = 0.67\sigma \end{aligned}$$

- **Örnek.** XYZ üniversitesindeki 100 öğrencinin boy uzunlukları (inç cinsinden) aşağıdaki gibidir. Yukarıdaki denemeli formüllerin geçerliliğini tartışınız.

Boy (inç)	Öğrenci Sayısı
60 – 63	5
63 – 66	18
66 – 69	42
69 – 72	27
72 – 75	8

- **Çözüm.** $OS = 2.26, \sigma = 2.92, Q = 1.98$ olduğundan

$$\begin{aligned} \frac{OS}{\sigma} &= \frac{2.26}{2.92} = 0.77 \\ \frac{Q}{\sigma} &= \frac{1.98}{2.92} = 0.68 \end{aligned}$$

dir. Her iki değer de denemeli formüllere uygundur.

4.11 Standartlaştırılmış Değişken; Standart Puanlar

- Bir X gözlem değerinin ortalamadan sapması $X - \overline{X}$ olarak tanımlanmıştı. Eğer bu sapma standart sapmaya bölünürse standart puan (z) (ya da standartlaştırılmış değişken)

$$z = \frac{X - \overline{X}}{\sigma}$$

ile bulunur. Bu değer kullanılan birimlerden bağımsızdır ve dağılımların karşılaştırılmasında kullanılır.

- **Örnek.** Bir öğrenci ortalama notun 76 ve standart sapmanın 10 olduğu matematik final sınavından 84 almıştır. Aynı öğrenci ortalama notun 82 ve standart sapmanın 16 olduğu fizik final sınavından ise 90 almıştır. Hangi derste daha iyi durumdadır?
- **Çözüm.** Standart puanlar sırasıyla

$$\begin{aligned} z_{matematik} &= \frac{84 - 76}{10} = 0.8 \\ z_{fizik} &= \frac{90 - 82}{16} = 0.5 \end{aligned}$$

olduğundan bu öğrenci matematikte daha iyi durumdadır.

5. Momentler, Çarpıklık ve Basıklık

5.1 Momentler

- **Tanım.** Terimlerin sıfırdan cebirsel sapmalarının $\left(X_j - 0\right)$ veya aritmetik ortalamadan cebirsel sapmalarının $\left(X_j - \overline{X}\right)$ çeşitli kuvvetlerinin aritmetik ortalamalarının tamamına **moment** denir.
- **Tanım.**$X_1, X_2, ..., X_N$ değerlerinin r. momenti

$$\overline{X^r} = \frac{X_1^r + X_2^r + \cdots + X_N^r}{N} = \frac{\sum_{j=1}^{N} X_j^r}{N}$$

olarak tanımlanır. $r = 1$ ise 1. moment elde edilir ve bu değer aritmetik ortalamadır.

- **Tanım.** $\overline{X}$ etrafındaki r. moment ise

$$m_r = \frac{\left(X_1 - \overline{X}\right)^r + \left(X_2 - \overline{X}\right)^r + \cdots + \left(X_N - \overline{X}\right)^r}{N} = \frac{\sum_{j=1}^{N} \left(X_j - \overline{X}\right)^r}{N}$$

şeklinde tanımlanır. Özel olarak $r = 1$ ise $m_1 = 0$ ve $r = 2$ ise $m_2 = \sigma^2$ (varyans) olur.

- **Tanım.** Herhangi bir A merkezi etrafındaki r. moment ise

$$m'_r = \frac{(X_1 - A)^r + (X_2 - A)^r + \cdots + (X_N - A)^r}{N} = \frac{\sum_{j=1}^{N} (X_j - A)^r}{N}$$

şeklinde tanımlanır. Özel olarak $r = 1$ ise $m'_1 = \overline{X} - A$ olur. Ayrıca $A = 0$ ise $m'_r = \overline{X^r}$, $A = \overline{X}$ ise $m'_r = m_r$ dir.

- **Örnek.** 2, 3, 7, 8, 10 sayı kümesinin
  - Birinci,
  - İkinci,
  - Üçüncü momentlerini bulunuz.
- **Çözüm.**
  - Birinci moment (aritmetik ortalamaya eşittir)

$$\overline{X^1} = \overline{X} = \frac{2 + 3 + 7 + 8 + 10}{5} = 6$$

  - İkinci moment

$$\overline{X^2} = \frac{2^2 + 3^2 + 7^2 + 8^2 + 10^2}{5} = 45.2$$

  - Üçüncü moment

$$\overline{X^3} = \frac{2^3 + 3^3 + 7^3 + 8^3 + 10^3}{5} = 378$$

olarak bulunur.

- **Örnek.** 2, 3, 7, 8, 10 sayı kümesinin aritmetik ortalama etrafındaki
  - Birinci,
  - İkinci,
  - Üçüncü momentlerini bulunuz.
- **Çözüm.** $\overline{X} = 6$ olduğundan
  - Aritmetik ortalama etrafındaki birinci moment

$$m_1 = \frac{(2-6) + (3-6) + (7-6) + (8-6) + (10-6)}{5} = 0$$

dır. m_1 daima sıfıra eşittir.

  - Aritmetik ortalama etrafındaki ikinci moment

$$m_2 = \frac{(2-6)^2 + (3-6)^2 + (7-6)^2 + (8-6)^2 + (10-6)^2}{5} = 9.2$$

dir. Bu değer varyansa eşittir. Yani $m_2 = \sigma^2$ dir.

  - Aritmetik ortalama etrafındaki üçüncü moment

$$m_3 = \frac{(2-6)^3 + (3-6)^3 + (7-6)^3 + (8-6)^3 + (10-6)^3}{5} = -3.6$$

- **Örnek.** 2, 3, 7, 8, 10 sayı kümesinin 4 merkezi etrafındaki
  - Birinci,
  - İkinci,
  - Üçüncü momentlerini bulunuz.
- **Çözüm.**
  - 4 merkezi etrafındaki birinci moment
  $$m_1' = \frac{(2-4)+(3-4)+(7-4)+(8-4)+(10-4)}{5} = 2$$
  - 4 merkezi etrafındaki ikinci moment
  $$m_2' = \frac{(2-4)^2+(3-4)^2+(7-4)^2+(8-4)^2+(10-4)^2}{5} = 13.2$$
  - 4 merkezi etrafındaki üçüncü moment
  $$m_3' = \frac{(2-4)^3+(3-4)^3+(7-4)^3+(8-4)^3+(10-4)^3}{5} = 59.6$$

  olarak bulunur.

5.2 Gruplandırılmış Verilerde Momentler

- Gruplandırılmış serilerde moment hesaplanırken yine sınıf orta noktaları dikkate alınır.
- **Örnek.** *XYZ* üniversitesindeki 100 öğrencinin boy uzunlukları (inç cinsinden) aşağıdaki gibidir. Bu boy dağılımı için aritmetik ortalama etrafındaki ilk üç momenti bulunuz.

Boy (inç)	Öğrenci Sayısı
60 – 63	5
63 – 66	18
66 – 69	42
69 – 72	27
72 – 75	8

- **Çözüm.** Sınıflandırılmış serilerde momentler bulunurken sınıf orta noktaları dikkate alınır.

Boy (inç)	Sınıf Orta Noktaları	Öğrenci Sayısı
60 – 63	61.5	5
63 – 66	64.5	18
66 – 69	67.5	42
69 – 72	70.5	27
72 – 75	73.5	8

Buna göre aritmetik ortalama

$$\overline{X} = \frac{61.5\cdot 5+64.5\cdot 18+67.5\cdot 42+70.5\cdot 27+73.5\cdot 8}{100} = 67.95$$

tir.

- Aritmetik ortalama etrafındaki ilk üç moment
  - $$m_1 = \frac{(61.5-67.95)\cdot 5+(64.5-67.95)\cdot 18+(67.5-67.95)\cdot 42+(70.5-}{100} \\ = 0$$
  - $$m_2 = \frac{(61.5-67.95)^2\cdot 5+(64.5-67.95)^2\cdot 18+(67.5-67.95)^2\cdot 42+(70.5}{100} \\ = 8.53$$
  - $$m_3 = \frac{(61.5-67.95)^3\cdot 5+(64.5-67.95)^3\cdot 18+(67.5-67.95)^3\cdot 42+(70.5}{100} \\ = -2.69$$

  şeklinde elde edilir.

5.3 Momentler Arasındaki İlişkiler

- $\overline{X}$ etrafındaki r. moment m_r ve herhangi bir A sayısı etrafındaki moment r. moment m'_r olmak üzere bu momentler arasında aşağıdaki ilişkiler vardır:

$$\begin{aligned} m_2 &= m'_2 - (m'_1)^2 \\ m_3 &= m'_3 - 3m'_1 m'_2 + 2(m'_1)^3 \\ m_4 &= m'_4 - 4m'_1 m'_3 + 6(m'_1)^2 m'_2 - 3(m'_1)^4 \\ &\vdots \end{aligned}$$

Burada $m'_1 = \overline{X} - A$ dır. Ayrıca daima $m_1 = 0$ dır. Bu eşitlikler **Köning teoremi** olarak bilinir.
- **Örnek.** Bir serinin bir A sayısı etrafındaki ilk üç momenti $6, 51$ ve 506 olduğuna göre ortalama etrafındaki ilk üç momenti bulunuz.
- **Çözüm.** $m'_1 = 6, m'_2 = 51$ ve $m'_3 = 506$ olarak verilmiş. Buna göre

$$\begin{aligned} m_1 &= 0 \\ m_2 &= m'_2 - (m'_1)^2 = 51 - 6^2 = 15 \\ m_3 &= m'_3 - 3m'_1 m'_2 + 2(m'_1)^3 = 506 - 3 \cdot 6 \cdot 51 + 2 \cdot 6^3 = 20 \end{aligned}$$

bulunur.

5.4 Momentler İçin Sheppard Düzeltmeleri

- Sınıflandırılmış serilerde sınıf orta noktaları dikkate alındığından momentler bulunurken gerçek değerden sapmalar olur. Bundan dolayı düzeltmeler uygulanır. Sheppard düzeltmesi denilen bu düzeltmeler ikinci ve dördüncü momentlere uygulanır. Birinci ve üçüncü momentlere uygulamaya gerek yoktur. c sınıf uzunluğu olmak üzere düzeltilmiş momentler

$$\begin{aligned} m_2 &= m_2 - \frac{1}{12}c^2 \\ m_4 &= m_4 - \frac{1}{2}c^2 m_2 + \frac{7}{240}c^4 \end{aligned}$$

şeklinmdedir.
- **Örnek.** XYZ üniversitesindeki 100 öğrencinin boy uzunlukları ile ilgi olan ve daha önce verilen aşağıdaki seri seri için ortalama etrafındaki ilk üç momente ait Sheppard düzeltmesini uygulayınız.

Boy (inç)	Öğrenci Sayısı
60 – 63	5
63 – 66	18
66 – 69	42
69 – 72	27

- **Çözüm.** Ortalama etrafındaki ilk üç moment

$$\begin{aligned} m_1 &= 0 \\ m_2 &= 8.53 \\ m_3 &= -2.69 \end{aligned}$$

şeklinde bulunmuştu. m_1 ve m_3 momentlerinin düzeltilmesine gerek yoktur. m_2 momentine düzeltme uygularsak

$$m_2 = m_2 - \frac{1}{12}c^2 = 8.53 - \frac{1}{12}3^2 = 7.78$$

bulunur.

5.5 Boyutsuz (Birimsiz) Biçimde Momentler

- Ölçü birimlerinin oluşturduğu problemi ortadan kaldırmak amacıyla ortalama etrafındaki boyutsuz (yani birimsiz) momentler $\sigma = \sqrt{m_2}$ olduğundan

$$\alpha_r = \frac{m_r}{\sigma^r} = \frac{m_r}{\sqrt{(m_2)^r}}$$

şeklinde tanımlanır. Ayrıca $m_1 = 0$ ve $m_2 = \sigma^2$ olacağından ilk iki boyutsuz moment $\alpha_1 = 0, \alpha_2 = 1$ olacaktır.
- **Örnek.** XYZ üniversitesindeki 100 öğrencinin boy dağılımı için α_3 birimsiz momentini bulunuz.
- **Çözüm.** $m_2 = 8.53$ ve $m_3 = -2.69$ olduğundan

$$\alpha_3 = \frac{m_3}{\sqrt{(m_2)^3}} = \frac{-2.69}{\sqrt{8.53^3}} = -0.11$$

olur.

5.6 Çarpıklık

- **Tanım.** Çarpıklık bir dağılımın simetrik olmayışı veya simetriklikten ayrılma derecesidir. Eğer bir dağılımın frekans eğrisinin merkezi maksimumun sağındaki kuyruk solundakinden daha uzun ise bu dağılımın sağa çarpık veya pozitif çarpıklığa sahip olduğu söylenir. Aksi halde ise sola çarpık veya negatif çarpıklığa sahip olduğu söylenir.
- Seriler **simetrik** veya **asimetrik** olduğu gibi **normal**, **sivri** ve **basık** serilerden birine de benzeyebilir. İşte bir serinin bunlardan hangisine uyduğunu belirlemek için aşağıdaki ölçülere ihtiyaç duyarız.
- Daha önce belirtildiği gibi bir seride medyan daima aritmetik ortalama ile mod arasındadır. Buna göre

$$
\begin{aligned}
\overline{X} &> Med > Mod \text{ ise seri sağa çarpık, çarpıklık pozitif,} \\
\overline{X} &= Med = Mod \text{ ise seri simetrik olabilir, çarpıklık 0,} \\
\overline{X} &< Med < Mod \text{ ise seri sola çarpık,çarpıklık negatif}
\end{aligned}
$$

şekildedir.
- Seri simetrik ise üç ortalama birbirine eşittir yani $\overline{X} = Med = Mod$ şeklindedir. Ancak tersi her zaman doğru değildir. $\overline{X} = Med = Mod$ olduğu halde seri simetrik olmayabilir.

- Genel olarak ortalamalara dayalı olarak çarpıklık

$$\text{Çarpıklık} = \frac{\overline{X} - Mod}{\sigma}$$

şeklinde tanımlanır. Daha önce verilen denemeli ilişki gereği $\overline{X} - Mod = 3\left(\overline{X} - Med\right)$ olduğundan mod kullanılmadan çarpıklık

$$\text{Çarpıklık} = \frac{3\left(\overline{X} - Med\right)}{\sigma}$$

olarak verilebilir. Bu eşitliklere sırasıyla **Pearson birinci ve ikinci çarpıklık katsayısı** denir. Ancak bu formüllerden ikincisinin ilki kadar sağlıklı sonuç vermediğine dikkat edilmelidir.

- **Örnek**. Bir A serisinde $\overline{X} = 3.62, \sigma = 1.83$ başka bir B serisinde $\overline{X} = 5, \sigma = 1.87$ dir. Diğer taraftan A serisinde $Mod = 3.14$, $Med = 3.43$, B serisinde ise $Mod = 5.4$, $Med = 5.25$ şeklindedir. Buna göre Pearson birinci ve ikinci çarpıklık katsayılarını bularak ve A ve B serilerinin çarpıklığını yorumlayınız.
- **Çözüm.**
  - A serisinin Pearson birinci ve ikinci çarpıklık katsayıları

$$
\begin{aligned}
\text{Çarpıklık}_A &= \frac{\overline{X} - Mod}{\sigma} = \frac{3.62 - 3.14}{1.83} = 0.26 > 0 \\
\text{Çarpıklık}_A &= \frac{3\left(\overline{X} - Med\right)}{\sigma} = \frac{3\,(3.62 - 3.43)}{1.83} = 0.31 > 0
\end{aligned}
$$

  - B serisinin Pearson birinci ve ikinci çarpıklık katsayıları

$$
\begin{aligned}
\text{Çarpıklık}_B &= \frac{\overline{X} - Mod}{\sigma} = \frac{5 - 5.4}{1.87} = -0.21 < 0 \\
\text{Çarpıklık}_B &= \frac{3\left(\overline{X} - Med\right)}{\sigma} = \frac{3\,(5 - 5.25)}{1.83} = -0.4 < 0
\end{aligned}
$$

  şeklindedir. Buna göre, A serisinin çarpıklığı B serisinden fazladır. Ayrıca A serisi sağa eğik olduğu halde B serisi sola eğiktir. Burada, Mod değerine bakarak karşılaştırma yapmak Med değerine bakarak karşılaştırma yapmaya göre daha sağlıklı olduğundan Mod değerlerine baktık.

- Çarpıklık ölçüleri çeyrekliklere göre de verilebilir. Eğer

$$
\begin{aligned}
Q_3 - Q_2 &> Q_2 - Q_1 \text{ ise seri sağa çarpık} \\
Q_3 - Q_2 &= Q_2 - Q_1 \text{ ise seri simetrik olabilir} \\
Q_3 - Q_2 &< Q_2 - Q_1 \text{ ise seri sola çarpık}
\end{aligned}
$$

şekildedir. Seri simetrik ise $Q_3 - Q_2 = Q_2 - Q_1$ eşitliği gerçekleşir. Ancak bu eşitliğin olması serinin simetrik olduğunu göstermez.
- Yukarıda verilen farklara dayalı çarpıklık ölçüleri

$$\text{Çarpıklık} = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_2) + (Q_2 - Q_1)} = \frac{Q_1 - 2Q_2 + Q_3}{Q_3 - Q_1}$$

şeklinde tanımlanır. Bu eşitlik **Bowley çarpıklık katsayısı** olarak bilinir.

- **Örnek.** Bir A serisinde $Q_1 = 2.29$, $Q_2 = 3.43$, $Q_3 = 5$ ve bir B serisinde ise $Q_1 = 4.25$, $Q_2 = 5.25$, $Q_3 = 6.4$ olarak bulunuyor. Buna göre bu iki serinin çarpıklığını karşılaştırınız.
- **Çözüm.** Yukarıda verilen üç çeyreklik formülde yazılırsa

$$\begin{aligned}\text{Çarpıklık}_A &= \frac{Q_1 - 2Q_2 + Q_3}{Q_3 - Q_1} = 0.16 > 0 \\ \text{Çarpıklık}_B &= \frac{Q_1 - 2Q_2 + Q_3}{Q_3 - Q_1} = 0.07 > 0\end{aligned}$$

bulunur. Demek ki her iki seri de sağa çarpık (eğik), ancak A serisinin çarpıklığı B serisinin çarpıklığından fazladır.

- Başka bir çarpıklık ölçüsü de ortalama etrafındaki üçüncü momente dayalı olarak verilen çarpıklık ölçüsüdür. Buna göre

$$\text{Moment çarpıklık katsayısı} = \alpha_3 = \frac{m_3}{\sigma^3} = \frac{m_3}{\sqrt{(m_2)^3}}$$

şeklinde verilir. Seri simetrik ise $\alpha_3 = 0$ dır. Ancak tersi doğru olmayabilir. Buna karşılık $\alpha_3 > 0$ ise seri sağa eğik, $\alpha_3 < 0$ ise seri sola eğiktir. Bu çarpıklık ölçüsü en sağlıklı çarpıklık ölçüsüdür. $\alpha_3 > 0.5$ ve $\alpha_3 < -0.5$ olan serilerin çarpıklığı kuvvetli kabul edilir.
- **Örnek.** Bir seride $m_2 = 2.29$, $m_3 = 0.61$ ve bir B serisinde ise serinin moment çarpıklık katsayısını bulunuz.
- **Çözüm.**

$$\begin{aligned}\text{Moment çarpıklık katsayısı} &= \alpha_3 = \frac{m_3}{\sigma^3} = \frac{m_3}{\sqrt{(m_2)^3}} = \frac{0.61}{\sqrt{2.29^3}} \\ &= 0.18\end{aligned}$$

olacaktır. $0.18 > 0$ olsuğundan seri sağa eğik fakat çarpıklığı zayıftır.

5.7 Basıklık

- **Tanım.** Basıklık bir dağılımın sivrilik derecesidir ve çoğunlukla normal dağılıma göre ele alınır.
- Ortalama etrafındaki dördüncü moment kullanılarak bir basıklık ölçüsü boyutsuz olarak

$$\text{Moment basıklık ölçüsü} = \alpha_4 = \frac{m_4}{\sigma^4} = \frac{m_4}{\sqrt{(m_2)^4}} = \frac{m_4}{(m_2)^2}$$

şeklinde verilir. Basıklığı $\alpha_4 = 3$ ve çarpıklığı $\alpha_3 = 0$ olan seriler **normal seri** olarak kabul edilir. $\alpha_4 = 3$ iken $\alpha_4 > 3$ ise seri sivri, $\alpha_4 < 3$ ise seri basıktır.

- **Örnek.** Bir A serisinde $m_2 = 2.29$, $m_4 = 7.92$ ve bir B serisinde $m_2 = 3.17$, $m_4 = 31.47$ olarak veriliyor. Bu iki serinin moment basıklık ölçüsünü bulunuz ve yorumlayınız .
- **Çözüm.** Ortalama etrafındaki dördüncü moment kullanılarak bir basıklık ölçüsü boyutsuz olarak

$$\begin{aligned}\text{Moment basıklık ölçüsü}_A &= \alpha_4 = \frac{m_4}{(m_2)^2} = \frac{7.92}{2.29^2} = 1.51 \\ \text{Moment basıklık ölçüsü}_B &= \alpha_4 = \frac{m_4}{(m_2)^2} = \frac{31.47}{3.17^2} = 3.13\end{aligned}$$

olur. $1.51 < 3$ olduğundan A serisi basık, $3.13 > 3$ olduğundan B serisi ise sivsiridir.

6. Temel Olasılık Teorisi

6.1 Faktöriyel

- **Tanım.** n pozitif bir doğal sayı olmak üzere $n!$ ile gösterilen n faktöriyel

$$n! = 1 \cdot 2 \cdot 3 \cdots (n-1) \cdot n$$

şeklinde tanımlanır. Özel olarak $0! = 1! = 1$ olarak kabul edilir. Buna göre, $0! = 1! = 1$, $2! = 2 = 2$, $3! = 6$, $4! = 24$, $5! = 120$, ... olur.

- **Örnek.**

$$\frac{9! + 8!}{8! - 7!}$$

ifadesini sadeleştiriniz.

- **Çözüm.** $9! = 9 \cdot 8!$ ve $8! = 8 \cdot 7!$ olduğundan

$$\begin{aligned}\frac{9!+8!}{8!-7!} &= \frac{9 \cdot 8! + 8!}{8 \cdot 7! - 7!} = \frac{8! \cdot (9+1)}{7! \cdot (8-1)} \\ &= \frac{8 \cdot 7! \cdot (9+1)}{7! \cdot (8-1)} = \frac{8 \cdot 10}{7} = \frac{80}{7}\end{aligned}$$

olur.

6.2 Saymanın Temel İlkesi

- **Tanım.** Bir A olayı $A_1, A_2, ..., A_r$ gibi r farklı olayın ard arda gerçekleşmesiyle meydana gelsin. Eğer A_1 olayı n_1 farklı yolla, A_2 olayı n_2 farklı yolla,...,A_r olayı n_r farklı yolla meydana geliyorsa A olayı $n_1 \cdot n_2 \cdots n_r$ farklı yolla meydana gelir. Buna **saymanın temel ilkesi** denir.
- **Örnek.** 4 tişörtü 3 pantolonu olan biri 1 tişört ve 1 pantolon ikilisini kaç farklı şekilde seçebilir?
- **Çözüm.** Saymanın temel ilkesi gereği $4 \cdot 3 = 12$ farklı yolla seçebilir.
- **Örnek.** 20 kişilik bir sınıftan önce bir başkan arkasından bir başkan yardımcısı, onun arkasından da bir onur kolu başkanı kaç farklı şekilde seçebilir?
- **Çözüm.** Saymanın temel ilkesi gereği bu üç kişi $20 \cdot 19 \cdot 18 = 6840$ farklı yolla seçilebilir.
- **Örnek.** 4 mektup 6 posta kutusuna atılacaktır.
  - Kaç farklı şekilde atılabilir?
  - Bir posta kutusuna 1 den fazla mektup atılmayacaksa kaç farklı şekilde atılabilir?
- **Çözüm.**
  - $6 \cdot 6 \cdot 6 \cdot 6 = 1296$
  - $6 \cdot 5 \cdot 4 \cdot 3 = 360$

- **Örnek.** 0,1,2,3,...,9 rakamları kullanılarak 4 basamaklı
  - Rakamları tekrarlı veya tekrarsız kaç sayı yazılabilir?
  - Rakamları tekrarsız kaç sayı yazılabilir?
  - Son basamağı 0 olan rakamları tekrarsız kaç sayı yazılabilir?
- **Çözüm.**
  - İlk basamağa 0 gelemeyeceğinden $9 \cdot 10 \cdot 10 \cdot 10 = 9000$
  - $9 \cdot 9 \cdot 8 \cdot 7 = 4536$
  - $9 \cdot 8 \cdot 7 \cdot 1 = 504$
- **Örnek.** Bir test sınavı her biri 5 seçenekli 20 sorudan oluşmaktadır. Bu testin cevap anahtarı
  - kaç farklı şekilde oluşturulabilir?
  - ardışık herhangi iki sorunun doğru cevabı aynı şık olmamak koşuluyla kaç farklı şekilde oluşturulabilir?
- **Çözüm.**
  - $5 \cdot 5 \cdots 5 = 5^{20}$
  - $5 \cdot 4 \cdot 4 \cdot 4 \cdots 4 = 20 \cdot 4^{18}$

6.3 Permütasyon (Sıralama,Diziliş)

- **Tanım.** Birbirinden farklı n tane nesneden r tanesinin dizilişlerinden her birine bu n elemanın r li bir permütasyonu (dizilişi,sıralanışı) denir. n farklı elemanın tüm r li permütasyonları ${}_nP_r$, $P(n,r)$, ya da $P_{n,r}$ ile gösterilir ve
$$P(n,r) = \frac{n!}{(n-r)!}$$
bağıntısı ile bulunur.
- **Örnek.** a, b, c harflerinin tüm permütasyonlarını bulunuz.
- **Çözüm.**

1 elemanlı permütasyonlar	:	a, b, c
2 elemanlı permütasyonlar	:	ab, ba, ac, ca, bc, cb
3 elemanlı permütasyonlar	:	$abc, acb, bac, bca, cab, cba$

- **Örnek.** 5 farklı renkteki bilye yanyana kaç farklı şekilde sıralanabilir?
- **Çözüm.** $n = 5$ ve $r = 5$ olduğundan
$$P(5,5) = \frac{5!}{(5-5)!} = \frac{5!}{0!} = 120$$
olur. Genel olarak n farklı nesne yan yana $n!$ sayıda sıralanabilir.

- **Örnek.** 4 kişilik bir koltuğa 10 kişi kaç farklı şekilde oturabilir?
- **Çözüm.** $n = 10$ ve $r = 4$ olduğundan
$$\begin{aligned} P(10,4) &= \frac{10!}{(10-4)!} = \frac{10!}{6!} = \frac{10\cdot 9\cdot 8\cdot 7\cdot 6!}{6!} \\ &= 10\cdot 9\cdot 8\cdot 7 = 5040 \end{aligned}$$
olarak bulunur.
- **Örnek.** 12345 sayısının rakamları kullanılarak rakamları farklı 3 basamaklı kaç sayı yazılabilir?
- **Çözüm.** $n = 5$ ve $r = 3$ olduğundan tüm 3 basamaklı sayılar
$$P(5,3) = 5\cdot 4\cdot 3 = 60$$
tanedir.

6.4 Tekrarlı Permütasyon

- n_1 sayıda nesne kendi aralarında özdeş, bunlardan farklı n_2 sayıda nesne kendi aralarında özdeş,..., bunlardan farklı n_k sayıda nesne kendi aralarında özdeş ve $n_1 + n_2 + \cdots + n_k = n$ olmak üzere bu nesnelerin tümü yan yana
$$\frac{n!}{n_1!\cdot n_2!\cdots n_k!}$$
farklı şekilde sıralanabilir. Buna tekrarlı permütasyon denir.
- **Örnek.** "istatistik" kelimesinin tüm harfleri yan yana kaç farklı şekilde sıralanabilir?
- **Çözüm.** 2 tane s, 3 tane t, 1 tane a, 3 tane i ve 1 tane k olduğundan ve toplam 10 harf olduğundan tüm sıralanışlar
$$\frac{10!}{2!\cdot 3!\cdot 1!\cdot 3!\cdot 1!} = 50400$$
kadar olacaktır.
- **Örnek.** 5 kırmızı, 2 beyaz, 3 mavi bilye yan yana dizilecektir. Aynı renkte olanlar ayırt edilemiyorsa, kaç farklı diziliş mümkündür?
- **Çözüm.** Toplam 10 bilye olduğundan
$$\frac{10!}{5!\cdot 2!\cdot 3!} = 2520$$

6.5 Dairesel Permütasyon

- n farklı nesne bir çember etrafına $(n-1)!$ kadar farklı şekilde sıralanabilir. Buna dairesel permütasyon denir.
- **Örnek.** Yuvarlak bir masa etrafına 7 kişi
  - kaç farklı şekilde oturabilir?
  - belirli iki kişi daima yan yana olacak biçimde kaç farklı şekilde oturabilir?
  - belirli iki kişi hiçbir zaman yan yana olmayacak ise kaç farklı şekilde oturabilir?
- **Çözüm.**
  - $n = 7$ olduğundan hiç bir sınırlama olmayan dairesel dizilişler $(n-1)! = 6! = 720$ tanedir.
  - Belirli iki kişi bir kişi gibi düşünülürse problem 6 kişinin dairesel olarak sıralanma problemine dönüşür. Ayrı belirli iki kişi kendi aralarında da sıralanacaklarından tüm sıralanışlar $5!\cdot 2! = 240$ kadardır.
  - Tüm durumlar 720, belirli iki kişinin yanyana olduğu durumlar 240 kardardı. O halde belirli iki kişininin yan yana olmadığı durumların sayısı $720 - 240 = 480$ olur.

6.6 Kombinasyon (Seçim, Gruplama)

- **Tanım.** n farklı nesne arasından seçilen r farklı elemana n nin r li bir kombinasyonu denir. Ya da n elemanlı bir kümenin r elemanlı alt kümelerinden her birine n nin r li bir kombinasyonu denir.
- **Örnek.** a, b, c harflerinin tüm kombinasyonlarını bulunuz.
- **Çözüm.**

1 elemanlı kombinasyonlar :	a, b, c
2 elemanlı kombinasyonlar :	ab, ac, bc
3 elemanlı kombinasyonlar :	abc

  Dikkat edilirse kombinasyonda sıranın önemi yoktur. Örneğin ab ve ba aynı kabul edilir.
- n farklı elemanın tüm r li kombinasyonları ${}_nC_r$, $C(n,r)$, $C_{n,r}$ ya da $\binom{n}{r}$ ile gösterilir ve

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

  bağıntısı ile bulunur.

- **Örnek.** 10 nesne arasından
  - 4 nesne kaç farklı şekilde seçilebilir?
  - önce 4 nesne sonra da kalanlar içinden 3 nesne kaç farklı şekilde seçilebilir?
  - seçilecek nesnelerden biri belirli ise 4 nesne kaç farklı şekilde seçilebilir?
- **Çözüm.**
  - $\binom{10}{4} = 210$
  - Saymanın temel ilkesi gereği $\binom{10}{4}\binom{6}{3} = 4200$
  - Seçilecek nesnelerden biri belli olduğuna göre kalan 9 nesne içinden 3 nesneyi $\binom{9}{3} = 84$ farklı şekilde seçeriz.

- **Örnek.** 5 doktor 6 hemşire arasından 3 kişilik bir sağlık ekibi oluşturulacaktır.
  - Kaç farklı şekilde oluşturulabilir?
  - Oluşacak ekipte en az bir doktor bulunmak şartıyla ekip kaç farklı şekilde oluşturulabilir?
  - Ali adlı doktor ile Ayşe adlı hemşire seçilecek ekipte aynı anda bulunmayacaksa ekip kaç farklı şekilde oluşturulabilir?
- **Çözüm.**
  - $\binom{11}{3} = 165$
  - $\binom{5}{1}\binom{6}{2} + \binom{5}{2}\binom{6}{1} + \binom{5}{3}\binom{6}{0} = 145$ veya tüm durumdan hiç doktor olmayan durumu çıkarırız $\binom{11}{3} - \binom{6}{3} = 145$
  - Tüm durumdan ikisinin de aynı anda olduğu durumu çıkarırız. $\binom{11}{3} - \binom{9}{1} = 156$

6.7 Olasılık

- Olasılığın bir çok tanımı yapılmıştır. Örneğin "olasılık gelecekte gerçekleşebilecek bir olay hakkındaki ümidimizin bir ölçüsüdür" tanımı bunlardan biridir. Burada matematiksel olarak iki temel tanım vereceğiz.
- **Klasik Tanım.** Bir olayın bütün olanaklı sonuçları eşit olabilirliğe sahip olsun. Bu olayın sonucu olan bütün olanaklı sonuçları "istenen" ve "istenmeyen" şeklinde iki gruba ayıralım. İlk gruptakilerin sayısı a, ikinci gruptakilerin sayısı b ise istenen sonucun ortaya çıkması olasılığı $p = \frac{a}{a+b}$, istenmeyen sonuçların ortaya çıkması olasılığı ise $q = \frac{b}{a+b}$ dir. Bazı durumlarda p istenmeyen durumuna karşılık istenmeyen durum q yerine p' gösterimi kullanılır. Buna göre

$$\begin{aligned} p + p' &= \frac{a}{a+b} + \frac{b}{a+b} = 1 \\ p &= 1 - p' \end{aligned}$$

  olacaktır.
- Kısaca bir olayın p olasılığı

$$p = \frac{\text{İstenen durumların sayısı}}{\text{Tüm durumların sayısı}}$$

- **Göreli (Oransal) Frekans Tanımı.** Klasik tanımdaki "eşit olabilirlik" ifadesi belirsizlik anlamı taşıdığından belirsiz bir anlam taşımaktadır. Bundan dolayı daha tutarlı bir tanım yapılmalıdır. Bir deneyde olanaklı sonuçların sayısı, deney sonsuz kere tekrarlandığında bir limit değerine yaklaşır. Söz konusu olan limite "bu olayın ortaya çıkma olasılığı" denir. Bu limiti

$$p = \lim_{n\to\infty} \frac{a}{n}$$

şeklinde veya $\frac{a}{n}$ in beklenen değeri anlamında

$$p = E\left(\frac{a}{n}\right)$$

şeklinde gösteririz.
- Daima $0 \leq p \leq 1$ eşitsizliği geçerlidir. Olma olasılığı $p = 1$ olan olan olay "kesin olay", $p = 0$ olan olan olay "imkansız olay" şeklinde tanımlanır.
- **Örnek.** Aşağıdaki olayların her biri için p olasılığını bulunuz.
  - Hilesiz bir zar atıldığında tek sayı gelmesi,
  - Hilesiz bir para iki kez atıldığında en az bir tura gelmesi,
  - Hilesiz bir çift zar atıldığında toplamın 7 gelmesi,
  - Hileli bir para 100 kez atıldığında 56 kez tura gelmektedir. Bu paranın yazı gelmesi olasılığı kaçtır?

- **Çözüm.**
  - Tüm durum 6 tane ve istenen durumlar $\{1, 3, 5\}$ olduğundan $p = \frac{3}{6} = \frac{1}{2}$.
  - Tüm durum $2 \cdot 2 = 4$ tane ve istenen durumlar TY,YT,TT olacağından $p = \frac{3}{4}$.
  - Tüm durum $6 \cdot 6 = 36$ ve istenen durumlar $(1,6), (6,1), (2,5), (5,2), (3,4), (4,3)$ olduğundan Tüm durum $p = \frac{6}{36} = \frac{1}{6}$.
  - Para 100 kez atıldığında 56 tura geldiyse 44 defa yazı gelmiştir. Demek ki yazı gelme olasılığı $p = \frac{44}{100} = 0.44$ tür. Bu örnekte paranın hilesiz olduğu söylenseydi yazı gelme olasılığı ilk durumlardan bağımsız olarak $p = \frac{50}{100} = 0.5$ olacaktı.
- **Örnek.** Deney hilesiz bir para ile bir zarın atılması olsun. Paranın tura ve zarın asal gelme olasılığını bulunuz.
- **Çözüm.** Saymanın temel ilkesi gereği $p = \frac{1}{2} \cdot \frac{3}{6} = \frac{1}{4}$.
- **Örnek.** İçinde aynı büyüklükte 4 kırmızı, 6 mavi bilye bulunan bir kutudan yerine geri koymadan arka arkaya 2 bilye çekiliyor. Buna göre,
  - Birincinin kırmızı, ikincinin mavi olma olasılığını,
  - İkisinin de mavi olma olasılığını,
  - İkisinin de aynı renk olma olasılığını bulunuz.

- **Çözüm.**
  - $p = \frac{4}{10} \cdot \frac{6}{9} = \frac{4}{15}$
  - $p = \frac{6}{10} \cdot \frac{5}{9} = \frac{1}{3}$
  - $p = \frac{4}{10} \cdot \frac{3}{9} + \frac{6}{10} \cdot \frac{5}{9} = \frac{7}{15}$
- **Not.** Bir kutudan bilye çekme olayında çekilen bilyeler geri konmayacaksa bilyeleri aynı anda çekmekle arka arkaya çekmek arasında fark yoktur.
- **Örnek.** Bir önceki örnekte bilyeler kutuya geri atılıyorsa aynı olasılıkları bulunuz.
  - $p = \frac{4}{10} \cdot \frac{6}{10} = \frac{6}{25}$
  - $p = \frac{6}{10} \cdot \frac{6}{10} = \frac{9}{25}$
  - $p = \frac{4}{10} \cdot \frac{4}{10} + \frac{6}{10} \cdot \frac{6}{10} = \frac{13}{25}$
- **Örnek.** İçinde aynı büyüklükte 6 kırmızı, 8 mavi bilye bulunan bir kutudan aynı anda 5 bilye çekiliyor. Çekilen bilyelerden ikisinin kırmızı üçünün mavi olma olasılığını bulunuz.
- **Çözüm.** $p = \frac{\binom{6}{2}\binom{8}{3}}{\binom{14}{5}} = \frac{60}{143}$

6.8 Bağımlı ve Bağımsız Olaylar

- **Tanım.** İki olaydan birinin olması diğerini etkilemiyorsa bu olaylara "bağımsız olaylar", etkiliyorsa "bağımlı olaylar" denir. Örneğin, bir ailedeki ikinci çocuğun cinsiyeti ilk çocuğun cinsiyetinden veya aynı anda olan farklı iki maçın sonuçları birbirinden bağımsızdır. Buna karşılık içinde mavi ve kırmızı bilyelerin bulunduğu bir torbadan geri konulmaksızın bilye çekildiğinde ikincinin mavi olma olayı ilkinin rengi ile bağımlıdır.
- E_1 ve E_2 iki olay olsun. Bunların olma olasılıklarını $P(E_1)$ ve $P(E_2)$ ile bu iki olayın aynı anda olma olasılığını da $P(E_1 E_2)$ ile gösterelim. Bu iki olay bağımsız ise

$$P(E_1 E_2) = P(E_1) \cdot P(E_2)$$

dir. Bu eşitlik ikiden fazla bağımsız olay için de yazılabilir.
- **Örnek.** Birbirinden bağımsız olarak ateş eden iki askerden birincisinin bir hedefi vurma olasılığı $\frac{5}{12}$, ikincisinin aynı hedefi vurma olasılığı $\frac{1}{6}$ dır. Askerler hedefe 1 kez ateş ettiklerinde hedefin yalnızca birinci asker tarafından vurulma olasılığını bulunuz.
- **Çözüm.** Hedefi sadece ilk asker vuracağına göre ikinci asker vuramamalıdır. İkinci askerin vuramama olasılığı $1 - \frac{1}{6} = \frac{5}{6}$ olduğundan $p = \frac{5}{12} \cdot \frac{5}{6} = \frac{25}{72}$ bulunur.

- **Örnek.** A ve B adlı kişiler 12 kez satranç oynuyor, bunlardan 6'sını A, 4'ünü B kazanıyor ve 2'si beraberlikle sonuçlanıyor. 3 oyundan oluşan bir maç yapmaya karar veriyorlar. Maç sonuçları birbirinden bağımsız ise
  - A nın üç oyunu da kazanması,
  - 2 oyunun beraberlikle sonuçlanması,
  - A ve B nin sırayla kazanması,
  - B nin en az bir oyunu kazanması olasılıklarını bulunuz.
- **Çözüm.** A nın herhangi bir oyunu kazanma olasılığı $\frac{6}{12} = \frac{1}{2}$, B nin herhangi bir oyunu kazanma olasılığı $\frac{4}{12} = \frac{1}{3}$ ve beraberlik olasılığı $\frac{2}{12} = \frac{1}{6}$ olduğundan
  - $p = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$
  - Berabere sonuçlanmayan oyun birinci ikinci veya üçüncü olabileceğinden ve bir oyunun berabere sonuçlanmama olasılığı $1 - \frac{1}{6} = \frac{5}{6}$ olduğundan

$$p = \frac{5}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} + \frac{1}{6} \cdot \frac{5}{6} \cdot \frac{1}{6} + \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{5}{6} = \frac{5}{72}$$

  - Kazanma sırası ABA veya BAB olacağından $p = \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{2} \cdot \frac{1}{3} = \frac{5}{36}$
  - Tüm durumun olasılığından B nin hiç kazanamama olasılığını çıkarırız. $p = 1 - \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} = \frac{19}{27}$

6.9 Koşullu Olasılık

- **Tanım.** E_1 ve E_2 iki bağımlı olay olsun. Bunlardan E_1 olayının gerşçekleştiği bilindiğinde E_2 olayının meydana gelme olasılığı $P(E_2|E_1)$ ile gösterilir ve E_2 olayının E_1 olayına bağlı koşullu olasılığı denir. $P(E_1) > 0$ olmak üzere

$$P(E_2|E_1) = \frac{P(E_1 E_2)}{P(E_1)}$$

ile bulunur.
- **Örnek.** Bir öğrencinin Hukuk dersinde başarılı olma olasılığı $P(H) = \frac{1}{4}$, Hukuk ve Matematik derslerinden ikisinde de başarılı olma olasılığı $P(HM) = \frac{3}{20}$ dir. Bu öğrencinin Hukuk dersinde başarılı olma koşuluyla Matematik dersinde de başarılı olma olasılığı kaçtır?
- **Çözüm.** Bize $P(M|H)$ soruluyor.

$$P(M|H) = \frac{P(HM)}{P(H)} = \frac{\frac{3}{20}}{\frac{1}{4}} = \frac{3}{5}$$

bulunur.

6.10 Ayrık Olaylar

- **Tanım.** Aynı anda meydana gelmesi mümkün olmayan olaylara "ayrık olaylar" denir. Örneğin iki takım arasında oynanan bir maçtaki kazanmak, kaybetmek ve berabere kalmak olayları aynı anda olamayacağından ayrık olaylardır. E_1 ve E_2 iki ayrık olay ise $P(E_1 E_2) = 0$ dır.
- E_1 ve E_2 iki olay olmak üzere E_1 veya E_2 nin olma olasılığı $P(E_1 + E_2)$ ile gösterilir

$$P(E_1 + E_2) = P(E_1) + P(E_2) - P(E_1 E_2)$$

ile bulunur. Eğer E_1 ve E_2 ayrık olaylar ise

$$P(E_1 + E_2) = P(E_1) + P(E_2)$$

olacaktır. Bu sonuç ikiden fazla ayrık olay için de söylenebilir.
- **Örnek.** Bir öğrencinin İstatistik ve Yönetim derslerinin sınavlarında başarılı olma olasılıkları $P(\dot{I}) = \frac{4}{5}$, $P(Y) = \frac{3}{4}$ olarak veriliyor. Bu öğrencinin bu derslerden başarılı olması birbirinden bağımsız ise İstatistik veya Yönetim derslerinin sınavlarından başarılı olma olasılığını bulunuz.

- **Çözüm.** Başarı durumları bağımsız olduğundan $P(\dot{I}Y) = P(\dot{I})P(Y)$ dir. Buna göre

$$\begin{aligned} P(\dot{I} + Y) &= P(\dot{I}) + P(Y) - P(\dot{I}Y) = P(\dot{I}) + P(Y) - P(\dot{I})P(Y) \\ &= \frac{4}{5} + \frac{3}{4} - \frac{4}{5} \cdot \frac{3}{4} = \frac{19}{20} \end{aligned}$$

- **Örnek.** Hilesiz bir para ile bir çift zarın atılması deneyinde paranın tura veya zarların üst yüzüne aynı sayı gelme olasılığını bulunuz.
- **Çözüm.** Tura gelme olasılığını $P(T)$ ile zarların aynı gelme olasılığını $P(A)$ ile gösterelim. $P(T) = \frac{1}{2}$ ve $P(A) = \frac{6}{36} = \frac{1}{6}$ dır. Diğer taraftan bu iki olay birbirinden bağımsızdır. Bu durumda istenen olasılık

$$\begin{aligned} P(T + A) &= P(T) + P(A) - P(TA) = P(T) + P(A) - P(T)P(A) \\ &= \frac{1}{2} + \frac{1}{6} - \frac{1}{2} \cdot \frac{1}{6} = \frac{7}{12} \end{aligned}$$

olarak bulunur.

7. Binom,Normal ve Poisson Dağılımları

7.1 Olasılık Dağılımları

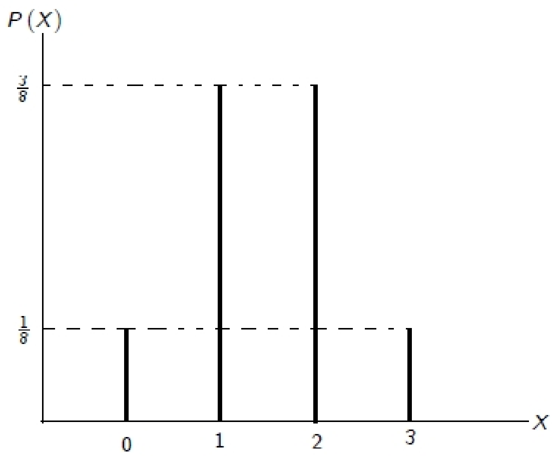
- **Tanım.** Rassal (tesadüfi) bir olayın (ya da deneyin) sonuçlarını, sayısal değerlerle ifade eden değişkene, rassal değişken denir. Rassal değişkenler, aldıkları değerlere göre **kesikli** ya da **sürekli** olarak adlandırılırlar. Değer kümesi sayılabilir olan rassal değişkenler **kesikli**, sayılamayan olan rassal değişkenler ise **sürekli** olarak isimlendirilir. Bir kavşakta meydana gelen trafik kazalarının sayısı kesikli rassal değişken, yarışmacıların 100 metreyi koşma süreleri sürekli rassal değişkendir.
- **Tanım.** X kesikli rassal değişkenin olmak üzere bu değişkenin aldığı değerler ile bu değerlere karşılık gelen $P(X=x)$ olasılıklarına **olasılık dağılımı** denir. Bir fonksiyonunun olasılık dağılımı olarak tanımlanabilmesi için aşağıdaki iki şartı sağlaması gerekir:
  1. X kesikli rassal değişkeninin, herhangi bir x olma olasılığı $[0,1]$ aralığında olmalı yani $0 \leq P(X=x) \leq 1$ olmalı,
  2. X kesikli rassal değişkeninin x in tüm olası değerlerine eşit olma olasılıklarının toplamı 1 e eşit olmalıdır yani $\sum P(X=x) = 1$ olmalıdır.

- **Örnek**. Bir çift hilesiz zar atıldığında üst yüze gelen sayıların toplamı X olsun. Buna göre, X in olasılık dağılımı nasıl olur?
- **Çözüm.** X en az 2, fazla 12 olacaktır. Bu durumda X in olasılık dağılımı

X	2	3	4	5	6	7	8	9	10	11	12
$P(X)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

  şeklinde olacaktır.
- **Örnek**. Kız ve erkek çocuklara ilişkin olasılıkların eşit olduğunu varsayarsak X, 3 çocuklu bir ailedeki erkek çocuk sayısını göstersin. Buna göre, X in olasılık dağılımı nasıl olur? Bu dağılımın grafiğini çiziniz.
- **Çözüm.** X en az 0, fazla 3 olacaktır. Bu durumda X in olasılık dağılımı ve bu dağılımın grafiği ise aşağıdaki gibi olacaktır.

X	0	1	2	3
$P(X)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

- 

- **Örnek.** X rassal değişkeninin aldığı değerler ve bu değerlere karşılık gelen olasılıklar aşağıdaki tabloda verilmiştir. $P(X = x)$ fonksiyonunun olasılık dağılımı olup olmadığını belirleyiniz.

X	0	1	2	3	4
$P(X)$	$\frac{3}{20}$	$\frac{2}{10}$	$\frac{3}{10}$	$\frac{2}{10}$	$\frac{3}{20}$

- **Çözüm.** $\sum P(X) = \frac{3}{20} + \frac{2}{10} + \frac{3}{10} + \frac{2}{10} + \frac{3}{20} = 1$ olduğundan $P(X)$ fonksiyonu olasılık dağılımıdır.
- **Örnek.** X rassal değişkeninin olasılık dağılımı aşağıdaki tabloda verilmiştir. c değerini bulunuz.

X	1	2	3	4
$P(X)$	0.12	0.16	0.54	c

- **Çözüm.** Bir olasılık dağılımında $\sum P(X) = 1$ olduğundan $0.12 + 0.16 + 0.54 + c = 1$ denkleminden $c = 0.18$ bulunur.

7.2 Matematiksel Beklenti (Beklenen Değer)

- **Tanım.** X kesikli rassal değişkeni $X_1, X_2, , ..., X_n$ değerlerini $p_1, p_2, , ..., p_n$ olasılıkları ile alabilsin. Bu durumda X kesikli rassal değişkeninin **matematiksel beklentisi** (veya beklenen değeri)

$$E(X) = p_1 \cdot X_1 + p_2 \cdot X_2 + \cdots + p_n \cdot X_n = \sum_{j=1}^{n} (p_j \cdot X_j)$$

şeklinde tanımlanmaktadır. Bu değer aynı zamanda kesikli rassal değişkenin ortamalasıdır ve bazen μ (mü) ile gösterilir. Yani $\mu = E(X)$ anlamına gelir.
- **Örnek.** Hilesiz bir zar atılıyor. Anlaşmaya göre A, babasından her atışta kaç gelirse o kadar dolar alacaktır. Matematiksel beklentiyi bulunuz.
- **Çözüm.** X üst yüze gelen sayı olmak üzere olasılık dağılımı

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

şeklinde olacaktır. Buna göre beklenti

$$E(X) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = 3.5$$

dolar olacaktır. Demek ki A, atış başına 3.5 dolar kazanmayı beklemektedir. Bu gerçek bir değer olmayıp matematiksel bir ümit anlamına gelir.

- **Örnek.** Bir adam, bir piyango bileti satın alırsa 5000$ lık bir ikramiyeyi 0.001 olasılıkla ya da 2000$ lık bir ikramiyeyi 0.003 olasılıkla kazanabilmektedir. Piyangonun adil olması için bilete ödenen uygun fiyat ne olmalıdır?
- **Çözüm.** Matematiksel beklenti

$$E = 5000 \cdot 0.001 + 2000 \cdot 0.003 = 11$$

olduğundan biletin fiyatı 11$ olmalıdır.
- **Örnek.**Belirli bir iş girişiminde bulunan bir bayan 0.6 olasılıkla 300$ kazanacak ya da 0.4 olasılıkla 100$ kaybedecektir. Beklentisini belirleyiniz.
- **Çözüm.** $E = 300 \cdot 0.6 - 100 \cdot 0.4 = 140$ olduğundan bayan 140$ kazanmayı beklemektedir.

- X kesikli rassal değişkeninin standart sapması

$$\sigma = \sqrt{\sum_{j=1}^{n} \left(X_j^2 \cdot p_j \right) - \mu^2}$$

ile bulunur. Kesikli bir rassal değişkenin standart sapması, olasılık dağılımının yayılmasının bir ölçüsüdür. Standart sapma değerinin büyük olması, X değerlerinin ortalama etrafında geniş bir aralıkta değerler aldığını gösterirken, küçük standart sapma değeri bu aralığın dar olduğunu, gözlenen X değerlerinin ortalamaya çok yakın değerler aldığını ifade eder.
- **Örnek.** X arızalı parça sayısını göstermek üzere, 400 parçalık bir üretimdeki arızalı parçalara ait olasılık dağılımı aşağıda verilmiştir. X değerinin standart sapmasını bulunuz.

X	0	1	2	3	4	5
$P(X)$	0.02	0.2	0.3	0.3	0.1	0.08

- **Çözüm.** Önce ortalamayı bulmalıyız. Ortalama

$$\begin{aligned} \mu &= E(X) = \sum_{j=0}^{5} (p_j \cdot X_j) \\ &= 0 \cdot 0.02 + 1 \cdot 0.2 + 2 \cdot 0.3 + 3 \cdot 0.3 + 4 \cdot 0.1 + 5 \cdot 0.08 = 2.5 \end{aligned}$$

- Standart sapma ise

$$\begin{aligned}\sigma &= \sqrt{\sum_{j=1}^{5}\left(X_j^2\cdot p_j\right)-\mu^2}=\sqrt{\sum_{j=1}^{5}\left(X_j^2\cdot p_j\right)-2.5^2}\\ &= \sqrt{0^2\cdot 0.02+1^2\cdot 0.2+2^2\cdot 0.3+3^2\cdot 0.3+4^2\cdot 0.1+5^2\cdot 0.08-2.5^2}\\ &= 1.2\end{aligned}$$

olacaktır. Demek ki 1.2 standart sapma ile üretilen 400 üründen ortalama 2.5 tanesi arızalıdır.
- **Chebyshev teoremi.** $k>1$ olmak üzere eğri altında kalan alanın en az $1-\frac{1}{k^2}$ kadarı $[\mu-k\cdot\sigma,\mu+k\cdot\sigma]$ aralığında yer alır.
- Son örnekte $k=2$ için en az $1-\frac{1}{2^2}=0.75$ i yani en az yüzde 75 kadarı $[2.5-2\cdot 1.2, 2.5+2\cdot 1.2]=[0.1,4.9]$ aralığındadır. Başka bir ifade ile üretilen 400 üründen en az yüzde 75 inin arızalı parça sayısı $[0.1,4.9]$ aralığındadır.

7.3 Binom Dağılımı

- **Tanım**. Eğer bir deney aşağıdaki dört koşulu sağlıyorsa bu deneye **binom deneyi** denmektedir:
  1. n tane özdeş deneme vardır. Yani verilen deney n kez özdeş (aynı) koşulda tekrarlanmaktadır.
  2. Her denemenin sadece ve sadece iki sonucu vardır. Bu sonuçlara genellikle istenen ya da istenmeyen denmektedir.
  3. p istenen olasılığı, q ise istenmeye olasılığı olmak üzere $p+q=1$ dir. p ve q olasılıkları her deneme için aynıdır.
  4. Bir denemenin sonucu öteki denemenin sonucunu etkilememektedir. Yani denemeler bağımsızdır.

- Bir binom deneyinde

$$\begin{aligned} N &: \text{toplam deneme sayısı}\\ p &: \text{istenen sonuç elde edilme olasılığı}\\ q &= 1-p : \text{istenmeyen sonuç elde edilme olasılığı}\\ X &: \text{istenen sonuç sayısı}\end{aligned}$$

olmak üzere, N denemeden X tanesinin başarılı olma olasılığı

$$P(X)=\binom{N}{X}p^X q^{N-X}$$

ile bulunur. Bu eşitliğe binom dağılımı denir. Burada $\binom{N}{X}$, N nin X li kombinasyonudur.

- **Örnek**. Hilesiz bir madeni para 6 defa atıldığında
  - tam 2 tura gelme,
  - en az 5 tura gelme,
  - en fazla 2 tura gelme olasılığını bulunuz.
- **Çözüm.** $N=6$, $X=2$, $p=q=\frac{1}{2}$ olduğundan binom dağılımı formülünden
  - $X=2$ için

$$P(X=2)=\binom{6}{2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^4=\frac{15}{64}$$

  - $X=5,6$ için

$$\begin{aligned}P(5\le X\le 6) &= \binom{6}{5}\left(\frac{1}{2}\right)^5\left(\frac{1}{2}\right)^1+\binom{6}{6}\left(\frac{1}{2}\right)^6\left(\frac{1}{2}\right)^0\\ &= \frac{7}{64}\end{aligned}$$

  - $X=0,1,2$ için

$$\begin{aligned}P(0\le X\le 2) &= \binom{6}{0}\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^6+\binom{6}{1}\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^5+\binom{6}{2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)\\ &= \frac{11}{32}\end{aligned}$$

- Binom dağılımının

Ortalaması	$\mu = Np$
Varyansı	$\sigma^2 = Npq$
Standart sapması	$\sigma = \sqrt{Npq}$
Moment çarpıklık katsayısı	$\alpha_3 = \dfrac{q-p}{\sqrt{Npq}}$
Moment Basıklık katsayısı	$\alpha_4 = 3 + \dfrac{1-6pq}{Npq}$

  ile bulunur.
- Bir para 64 defa atıldığında bulunan turaların
  - Ortalaması $\mu = Np = 64 \cdot \frac{1}{2} = 32$
  - Varyansı $\sigma^2 = Npq = 64 \cdot \frac{1}{2} \cdot \frac{1}{2} = 16$
  - Standart sapması $\sigma = \sqrt{Npq} = \sqrt{16} = 4$
  - Moment çarpıklık katsayısı $\alpha_3 = \dfrac{q-p}{\sqrt{Npq}} = \dfrac{\frac{1}{2}-\frac{1}{2}}{\sqrt{Npq}} = 0$
  - Moment Basıklık katsayısı $\alpha_4 = 3 + \dfrac{1-6pq}{Npq} = 3 + \dfrac{1-6\cdot\frac{1}{2}\cdot\frac{1}{2}}{16} = 2.97$ olur.

- **Örnek.** Yapılan bir araştırmayla bir kasabadaki erişkinlerin % 58'inin psikolojik sorunu olduğu bulunmuştur. Bu kasabadan rassal 25 erişkin seçilmiştir. X , bu örneklemdeki psikolojik sorunu olan kişi sayısını göstermek üzere, X 'in olasılık dağılımının ortalama ve standart sapmasını bulunuz.
- **Çözüm.** Ortalama $\mu = Np = 25 \cdot \frac{58}{100} = 14.5$, standart sapma $\sigma = \sqrt{25 \cdot \frac{58}{100} \cdot \frac{42}{100}} = 2.47$ olur. Demek ki bu kasabada seçilen 25 kişiden 2.47 standart sapma ile 14.5 tanesi psikolojik sorunludur.
- **Örnek.** Bir kargo şirketinin paketlerden %2 sini belirlenen sürede yerine ulaştıramadığı bilinmektedir. Bir müşteri 10 tane paketi bu kargo firmasına vererek göndermek istediğinde bu paketlerden
  - sadece 2 tanesinin zamanında yerine ulaşmama olasılığını
  - en fazla 1 tanesinin zamanında yerine ulaşmama olasılığını binom dağılımına göre bulunuz.
- **Çözüm.** $N = 10, p = \frac{2}{100}, q = \frac{98}{100}$ olduğundan
  -
$$P(X=2) = \binom{10}{2}\left(\frac{2}{100}\right)^2\left(\frac{98}{100}\right)^8 = 0.02$$
  -
$$\begin{aligned} P(0)+P(1) &= \binom{10}{0}\left(\frac{2}{100}\right)^0\left(\frac{98}{100}\right)^{10} + \binom{10}{1}\left(\frac{2}{100}\right)^1\left(\frac{98}{100}\right)^9 \\ &= 0.98 \end{aligned}$$

7.4 Poisson Dağılımı

- Poisson olasılık dağılımı, binom dağılımı gibi X in kesikli bir rassal değişken olması durumunda yaygın olarak kullanılan dağılımlardan biridir. Verilen bir aralıkta bir olayın tekrar etme sayısı λ (lambda) olmak üzere bu aralıkta aynı olayın X kez olma olasılığı
$$P(X) = \frac{e^{-\lambda} \cdot \lambda^X}{X!}$$
  ile bulunur. Burada $e = 2.7182...$ sayısıdır ve $X = 0, 1, 2, 3, ...$ gibi kesikli değerler alabilir.
- Uygulamada Binom dağılımını kullanmak bazen güç olabilir. Böyle durumlara alternatif olarak Poisson dağılımı kullanılır.
- **Örnek.** Yapılan bir araştırmada $18-24$ yaş grubundaki tüketicilerin ayda ortalama 6.9 kez alışverişe çıktıkları bulunmuştur. Poisson olasılık dağılımına uyduğu düşünülen rassal değişken için, bu yaş grubunun ayda tam 5 kez alışverişe çıkma olasılığını bulunuz.
- **Çözüm.** $\lambda = 6.9$ ve $X = 5$ için poisson olasılık dağılımını bulmak istiyoruz. Poisson olasılık dağılımı formülünden
$$P(X=5) = \frac{e^{-6.9} \cdot 6.9^5}{5!} = 0.13$$

- **Örnek.** Bir fabrikada üretilen gereçlerin % 10 u kusurlu çıkmaktadır. Rastgele seçilen 10 gereçlik bir örnekten tam 2 sinin kusurlu olma olasılığını
  - Binom olasılık dağımını kullanarak,
  - Poisson olasılık dağılımını kullanarak bulunuz.
- **Çözüm.**
  - $N = 10, X = 2, p = \frac{10}{100}, q = \frac{90}{100}$ olduğundan Binom olasılık dağılımından
$$P(X=2) = \binom{10}{2}\left(\frac{10}{100}\right)^2\left(\frac{90}{100}\right)^8 = 0.19$$
  - $\lambda = Np = 10 \cdot \frac{10}{100} = 1$ ve $X = 2$ için Poisson olasılık dağılımından
$$P(X=2) = \frac{e^{-1} \cdot 1^2}{2!} = 0.18$$
  Dikkat edilirse her iki değer de birbirine çok yakındır. Binom dağılımından elde edilen sonuç gerçek olasılık değeri, poisson olasılık dağılımından elde edilen sonuç yaklaşık olasılık değeridir.

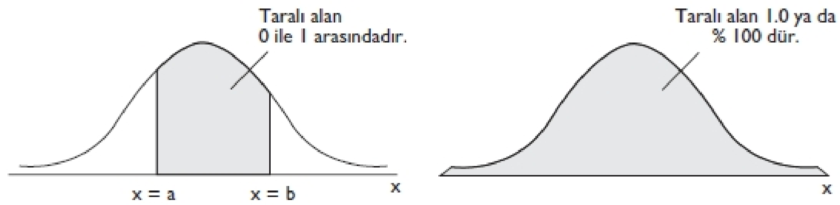
- Poisson dağılımın

Ortalaması	$\lambda = \mu = Np$
Varyansı	$\sigma^2 = \lambda$
Standart sapması	$\sigma = \sqrt{\lambda}$
Moment çarpıklık katsayısı	$\alpha_3 = \dfrac{1}{\sqrt{\lambda}}$
Moment Basıklık katsayısı	$\alpha_4 = 3 + \dfrac{1}{\lambda}$

olacaktır. Bu eşitlikler, Binom dağılımında N büyük ve bir olayın olma olasılığı p sıfıra yakınken yani $p \approx 0$ ve $q \approx 1$ iken elde edilir. Genelde $p \leq 0.1$ ve $\lambda = \mu = Np \leq 5$ olduğunda (bu durumda $N \geq 50$ olur) Poisson olasılık dağılımı, Binom olasılık dağılımına oldukça yakın değerler verir.

7.5 Normal Dağılım

- X rassal değişkeninin kesikli olmadığı yani sürekli olduğu durumlarda kullanılan önemli bir dağılım türüdür.
- Normal dağılım aşağıdaki özelliklere sahiptir:
  - X in olasılığı $[0, 1]$ aralığındadır.
  - X in alabileceği tüm değerlerin olasılıkları toplamı 1 dir
  - Eğri ortalamaya göre simetriktir.
  - Eğrinin iki ucu (kuyruğu) sonsuza gitmektedir.
- Normal dağılımın olasılık dağılım eğrisi



şeklindedir. X in a ve b arasında olma olasılığı yukarıdaki şekilde taralı alana eşittir ve bu olasılık $P(a < X < b)$ ile gösterilir.

- Normal dağılım fonksiyonu μ normal dağılımın ortalaması, σ standart sapması olmak üzere

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

şeklindedir. Bu durumda $P(a < X < b)$ olasılığı ile Normal dağılım eğrisi altında $X = a$ ile $X = b$ arasıdaki alana eşit olacaktır. Bu alan ise

$$P(a < X < b) = \frac{1}{\sigma\sqrt{2\pi}} \int_a^b e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dX$$

integrali ile bulunur.
- Eğer bir normal dağılımda $\mu = 0$ ve $\sigma = 1$ olursa bu normal dağılıma **standart normal dağılım** denir. Herhangi bir X değişkeni

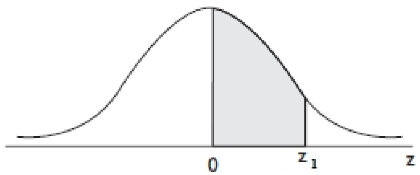
$$z = \frac{X - \mu}{\sigma}$$

yardımıyla z değerlerine dönüştürülebilir. Bu durumda $P(a < X < b)$ olasılığı $P(z_1 < z < z_2)$ olasılığına dönüşür. Buna normal dağılımın standart hale getirilmesi denir. Bu dönüşüm sonrasında normal dağılım fonksiyonu ise

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

şekline dönüşür. Bu dönüşüm alanları değiştirmez.

- Standart normal eğri altındaki 0 dan z_1 e kadar olan alan ya da başka bir ifadeyle yani $P(0 < z < z_1)$ olasılığı şekildeki taralı alanı verir.

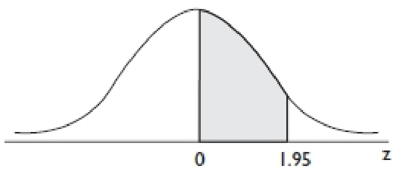


- Bu alan ise (veya olasılık)

$$P(0 < z < z_1) = \frac{1}{\sqrt{2\pi}} \int_0^{z_1} e^{-\frac{z^2}{2}} dz$$

integraline eşittir. Bu integrali hesaplamak mümkün olmadığından istenen alan (veya olasılık) tablo yardımıyla bulunur.

- **Örnek.** Standart normal eğri altında $z = 0$ ile $z = 1.95$ arasındaki alanı bulunuz.
- **Çözüm.** İstenen alan aşağıdaki taralı alandır:



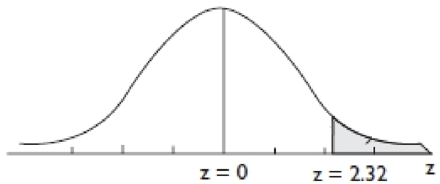
Bu taralı alanı veren integral

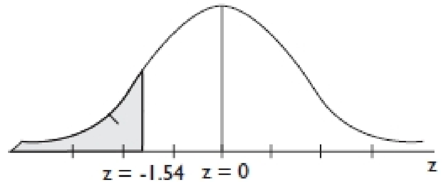
$$P\left(0 < z < 1.95\right) = \frac{1}{\sqrt{2\pi}} \int_0^{1.95} e^{-\frac{z^2}{2}} dz$$

belirli integralidir. Bu integrali hesaplamak kolay olmadığından (gerçekte tam çözümü yoktur) bu integralin yaklaşık çözümü standart normal dağılım tablosundan 0.4744 olarak bulunur.

- **Örnek.** Standart normal eğri altında $z = -2.17$ ile $z = 0$ arasındaki alanı bulunuz
- **Örnek.** Simetriden dolayı bu alan yerine $z = 0$ ile $z = 2.17$ arasındaki alanı da bulabiliriz. Tablo yardımıyla bu alan 0.4850 olarak bulunacaktır.

- **Örnek.** Standart normal eğri altında olan ve
  - $z = 2.32$ nin sağdaki alanı,
  - $z = -1.54$ ün solundaki alanı bulunuz
- **Çözüm.** İstenen alanlar aşağıdaki şekilde gösterilmiştir:



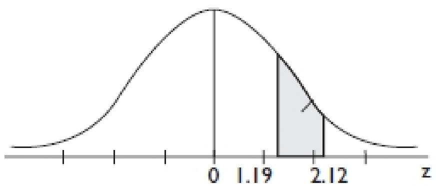


  - Standart normal eğri simetrik olduğundan ortalamanın sağındaki tüm alan 0.5 tir.Önce $z = 0$ ile $z = 2.32$ arasındaki alanı tablo yardımıyla buluruz. Bulduğumuz bu alanı 0.5 ten çıkararark istenen alanı bulmuş oluruz. O halde istenen alan $P\left(z \geq 2.32\right) = 0.5 - 0.4898 = 0.0102$ olur.
  - Tabloda negatif z değerleri olmadığından simetriden dolayı istenen alan yerine $z = 1.54$ ün sağıdaki alanı bulabiliriz. Bir önceki düşünceyle istenen alanı

$$P\left(z \leq -1.54\right) = P\left(z \geq 1.54\right) = 0.5 - 0.4382 = 0.0618$$

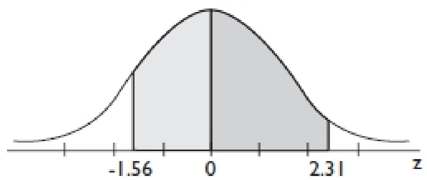
olarak elde ederiz.

- **Örnek.** Standart normal eğri altında aşağıdaki olasılıkları (yani alanları) bulunuz.
  - $P\left(1.19 \leq z \leq 2.12\right)$
  - $P\left(-1.56 \leq z \leq 2.31\right)$
  - $P\left(-0.75 \leq z\right)$
- **Çözüm.**
  - İstenen alan aşağıdaki taralı alandır.



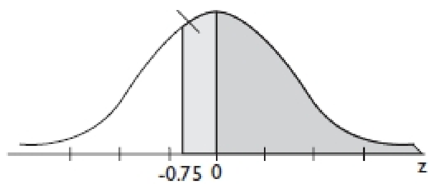
Tablodan 0 ile 2.12 arasındaki alanı ve 0 ile 1.19 arasındaki alanı buluruz ve bu alanların farkını alırız. Bu durumda istenen olasılığı $P\left(1.19 \leq z \leq 2.12\right) = 0.4830 - 0.3830$ olarak bulmuş oluruz.

- Benzer olarak
  - İstenen alan aşağıdaki grafikte gösterilmiştir:



Yine simetri yardımıyla
$P\left(-1.56 \leq z \leq 2.31\right) = P\left(0 \leq z \leq 2.31\right) + P\left(0 \leq z \leq 1.56\right)$ yazılabilir. Bu durumda istenen alan $0.4406 + 0.4896 = 0.9302$ bulunur.

  - İstenen alan aşağıdaki alandır:

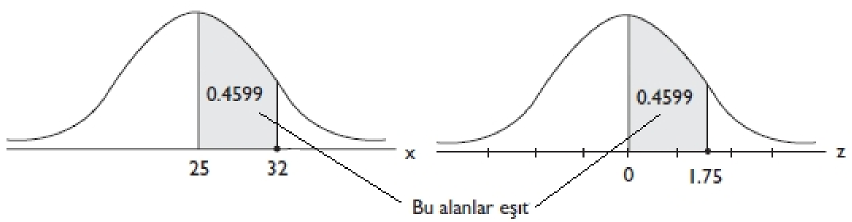


Tablo yardımıyla istenen alan $0.5 + 0.2734 = 0.7734$ bulunacaktır.

- **Not.** Standart normal eğri tablosunda en son değer $z = 3.99$ değeridir. Bu değerden daha büyük değerleri için tabloya bakılarak alan bulunamaz. Bu sorunun üstesinden gelmek için $P(0 \leq z \leq 3.99) = 0.5$ olarak kabul edilir. Yani $z = 3.99$ un sağındaki alan 0 kabul edilir. Örneğin $P(0 \leq z \leq 5.12) = 0.5$ tir.
- **Örnek.** Matematik final sınavında ortalama 72, standart sapma 15 olmuştur.
  - 60 alan öğrencinin standart puanını (yani z değerini),
  - 93 alan öğrencinin standart puanını (yani z değerini),
  - 72 alan öğrencinin standart puanını (yani z değerini) bulunuz.
- **Çözüm.** $\mu = 72, \sigma = 15$ ve $z = \frac{X-\mu}{\sigma}$ olduğundan
  - 60 alan öğrencinin standart puanı
  $$z = \frac{60-72}{15} = -0.8$$
  - 93 alan öğrencinin standart puanı
  $$z = \frac{93-72}{15} = 1.4$$
  - 72 alan öğrencinin standart puanı
  $$z = \frac{72-72}{15} = 0$$
  olur.

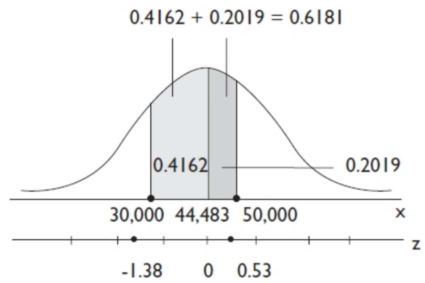
- X değerleri belli iken z değerleri $z = \frac{X-\mu}{\sigma}$ ile bulunuyordu. Bu formülde X yalnız bırakılırsa
$$X = \mu + z \cdot \sigma$$
elde edilir.
- **Örnek.** Matematik final sınavında ortalama 72, standart sapma 15 olmuştur. Buna göre
  - $z = -1$
  - $z = 1.6$ standart puanlarına karşılık gelen X değerlerini bulunuz.
- **Çözüm.** $X = \mu + z \cdot \sigma$ olduğundan
  - $z = -1$ e karşılık gelen X değeri
  $$X = \mu + z \cdot \sigma = 72 + (-1) \cdot 15 = 57$$
  - $z = 1.6$ ya karşılık gelen X değeri
  $$X = \mu + z \cdot \sigma = 72 + 1.6 \cdot 15 = 96$$
  olacaktır.

- **Örnek.** X sürekli rassal değişkeni 25 ortalama ve 4 standart sapma ile normal dağılmaktadır. $X = 25$ ile $X = 32$ arasındaki alanı bulunuz.
- **Çözüm.** Verilen normal dağılımda $\mu = 25$ ve $\sigma = 4$ olarak verilmiş. X değerlerini z değerlerine dönüştürüp tablo yardımıyla istenen alanları bulacağız.
  - $X = 25$ ise $z = \frac{X-\mu}{\sigma} = \frac{25-25}{4} = 0$ ve $X = 32$ ise $z = \frac{X-\mu}{\sigma} = \frac{32-25}{4} = 1.75$ olur. $P(25 < X < 32) = P(0 < z < 1.75)$ olduğundan aşağıda verilen alanlardan üstteki yerine alttakini bulacağız:



Tablodan $P(0 < z < 1.75) = 0.4599$ olarak bulunacaktır.

- **Örnek.** Bir ülkede aile başına düşen gelir 44483 dolar ve 10500 dolar standart sapma ile normal dağılım göstermektedir. Bu ülkeden rassal seçilen bir ailenin $30000 - 50000$ dolar arasında gelire sahip olma olasılığını bulunuz.
- **Çözüm.** Bizden $P(30000 \leq X \leq 50000)$ olasılığı isteniyor. Standart puanlar
$$z = \frac{X-\mu}{\sigma} = \frac{30000-44483}{10500} = -1.38$$
$$z = \frac{X-\mu}{\sigma} = \frac{50000-44483}{10500} = 0.53$$
olduğundan istenen alan aşağıdaki taralı alandır.



Tablo yardımıyla istenen olasılık aşağıdaki gibi bulunur:
$$P(30000 \leq X \leq 50000) = P(-1.38 \leq z \leq 0.53) = 0.6181$$

- **Örnek.** Biyolojiden yapılan test sınavında 10 soru sorulmuş ve her doğru soruya 1 puan yanlış soruya 0 puan verilmiştir. Sınavın ortalaması 6.7, standart sapması da 1.2 puan olmuştur. Buna göre,
  - en alttaki %10 sınıfına giren en yüksek puanı,
  - en üstteki %10 sınıfına giren en düşük puanı bulunuz.
- **Çözüm.**
  - Soldaki alan 0.10 olduğuna göre bu alana en yakın z değeri -1.28 dir.

$$\begin{aligned} -1.28 &= \frac{X-6.7}{1.2} \\ X &= 5.2 \end{aligned}$$

ve buna en yakın tamsayı olarak 5 olur.

  - Sağdaki alan 0.10 olduğuna göre bu alana en yakın z değeri 1.28 dir. Buna göre

$$\begin{aligned} 1.28 &= \frac{X-6.7}{1.2} \\ X &= 8.2 \end{aligned}$$

ve buna en yakın tamsayı olarak 8 olur.

- Normal dağılımın

Ortalaması	μ
Varyansı	σ^2
Standart sapması	σ
Moment çarpıklık katsayısı	$\alpha_3 = 0$
Moment Basıklık katsayısı	$\alpha_4 = 3$

şeklindedir.

7.6 Normal Dağılım ile Binom Dağılımı Arasındaki İlişki

- N büyükse ve p veya q sıfıra çok yakın değilse Binom dağılımdan elde edilen $\mu = N\cdot p$ ortalaması $\sigma = \sqrt{N\cdot p\cdot q}$ standart sapması

$$z = \frac{X-\mu}{\sigma} = \frac{x - N\cdot p}{\sqrt{N\cdot p\cdot q}}$$

ile standart hale getirilir ve Normal dağılım ile sonuca gidilir. N büyüdükçe Normal dağılımdan elde edilen olasılık değeri Binom dağılımından elde edilen olasılık değerine yaklaşır öyle ki limit durumunda iki olasılık birbirine eşit olur. Uygulamada $Np > 5$ veya $Nq > 5$ için iki olasılık dağılımı birbirine oldukça yakın değerler alır.
- Bilindiği gibi Binom dağılımı kesikli, Normal dağılım sürekli bir dağılımdır. Kesikli bir değişken $X = 0, 1, 2, 3, ..$ gibi değerler alabildiğinden böyle bir değişkeni sürekli hale getirmek için herhangi bir X değeri yerine $[X-0.5, X+0.5]$ aralığı alınır. Örneğin $X = 5$ ise bunun yerine X in $[4.5, 5.5]$ aralığında olduğu kabul edilir.

- **Örnek.** Hatasız bir madeni para 10 kez atıldığında en az 3 en fazla 6 kez tura gelme olasılığını,
  - Binom dağılımına göre,
  - Normal dağılıma göre hesaplayınız.
- **Çözüm.**
  - $N = 10, p = q = \frac{1}{2}$ olduğundan

$$\begin{aligned} P(3 \le X \le 6) &= \frac{15}{128} + \frac{105}{512} + \frac{63}{256} + \frac{105}{512} \\ &= \frac{99}{128} = 0.7734 \end{aligned}$$

olur. Ayrıca ortalama $\mu = N\cdot p = 10\cdot\frac{1}{2} = 5$, standart sapma $\sigma = \sqrt{N\cdot p\cdot q} = \sqrt{10\cdot\frac{1}{2}\cdot\frac{1}{2}} = 1.58$ olur.

  - Verileri sürekli hale getirirken binom dağılımındaki 3 yerine normal dağılımda $[2.5-3.5]$ aralığı, benzer olarak binom dağılımındaki 6 yerine normal dağılımda $[5.5-6.5]$ aralığı alınır. Bundan dolayı en az 3 en fazla 6 kez tura gelme olasılığı yerine normal dağılımda X in 2.5 ile 6.5 aralığında olma olasılığını bulacağız. Buna göre

$$\begin{aligned} 2.5 \text{ in standart değeri} &= \frac{2.5-5}{1.58} = -1.58 \\ 6.5 \text{ in standart değeri} &= \frac{6.5-5}{1.58} = 0.95 \end{aligned}$$

olur. Tablodan $P(-1.58 \le z \le 0.95) = 0.7718$ olarak elde edilir.

- **Örnek.** $1-4$ yaş grubunda emniyet kemeri kullananların %75 olduğu varsayılıyor. $1-4$ yaş grubundakileri içeren rastgele 100 otomobil durdurulduğunda 70 veya daha az sayıda emniyet kemeri takan ile karşılaşma olasılığını Normal dağılıma göre bulunuz.
- **Çözüm.**
  - $N = 100, p = \frac{75}{100}, q = \frac{25}{100}$ olduğundan $\mu = N \cdot p = 100 \cdot \frac{75}{100} = 75$ ve $\sigma = \sqrt{N \cdot p \cdot q} = \sqrt{100 \cdot \frac{75}{100} \cdot \frac{25}{100}} = 4.33$ olur. $X = 70$ yerine değişkeni sürekli hale getirmek için $X = 70.5$ alır ve standart puanı hesaplarsak

$$z = \frac{X-\mu}{\sigma} = \frac{70.5-75}{4.33} = -1.04$$

olur. Demek ki $P\left(z \leq -1.04\right)$ olasılığını veya simetriden dolayı $P\left(z \geq 1.04\right)$ olasılığını bulacağız. Tablo yardımıyla istenen olasılık 0.1492 olarak elde edilir.

8. Temel Örnekleme Teorisi

8.1 Örnekleme Teorisi

- **Tanım.** Planlanan bir istatistiksel araştırma için tanımlanan sonlu evrenin bütün birimleri üzerinden araştırmaya konu olan değişkenler itibarıyla veri derleniyorsa yapılan işleme **tamsayım** denir. Tamsayım sonucu elde edilen veri kümesinin çözümlenmesiyle elde edilen bilgiler (**parametre değerleri**) veri derleme ve çözümleme hatası işlenmemiş ise kesin ve doğru bilgilerdir. Tamsayım genellikle sonlu ve küçük hacimli evrenlere uygulanır. Bununla birlikte büyük hacimli sonlu evrenlerin bütün birimlerine ulaşabilmek olanaklıysa karşılaşılan özel problemin çözümü için mümkün bütün verilerin elde edilmesine gereksinim varsa tamsayım yapılmalıdır.
- **Tanım**. Tanımlanan evrenden onu ilgilenilen değişkenler bakımından temsil eden sınırlı sayıda birimin belirli yöntemler kullanılarak seçilmesi işlemine örnekleme, seçilen birimlerin oluşturduğu topluluğa **örneklem** denir. Seçilen birimlerin alındığı evrensel kümeye de **yığın** denir. Örnekleme teorisi gözlenen iki örnek arasındaki farkın şansa bağlı mı yoksa gerçekten anlamlı fark olup olmadığının belirlenmesinde kullanılır. Genel olarak örnekleme teorisi yığından çekilen örnekler yardımıyla o yığın hakkında yapılan çıkarımların bir çalışmasıdır. Olasılık teorisi yardımıyla bu çıkarımların doğruluğu veya yanlışlığı kanıtlanır.

8.2 Örnekleme Dağılımları

- **Tanım.** Yığından alınan örnek bilgilerinin hesabıyla bulunan niteliklere (örneğin örnek ortalaması, varyans) **örnek istatistikleri** veya kısaca **istatistik** denir.
- **Tanım.** Örnekleme teorisi sonucu elde edilen çıkarımların geçerli olabilmesi için seçilen örnekler yığını en iyi temsil edecek şekilde seçilmelidir. Eğer yığındaki her birimin seçilecek örneklemde çıkma şansı eşit ise bu örneklemeye **tesadüfi örnekleme** denir.
- **Tanım.** Belirlenen bir yığından (yerine koyarak veya koymayarak) seçilen bütün N çaplı mümkün örnekleri düşünelim. Her bir örnek için farklı bir istatistik (ortalama, standart sapma vb.) hesaplanabilir. Bu yolla bir istatistik dağılımı bulunur. Bu bulunan dağılım **örnekleme dağılımı** olarak isimlendirilir.

8.3 Ortalamaların Örnekleme Dağılımları

- N_p ile sonlu bir yığındaki eleman sayısını, N ile de buradan yerine konmadan seçeceğimiz bütün örneklerin sayısını (çapını) gösterelim. Yığın ortalamasını μ, yığın standart sapmasını σ, ortalamaların örnekleme dağılımının ortalamasını $\mu_{\overline{X}}$, ortalamaların örnekleme dağılımının standart sapmasını $\sigma_{\overline{X}}$ ile gösterelim. Yığın **sonlu sayıda** eleman içeriyorsa

$$\begin{aligned} \mu_{\overline{X}} &= \mu \\ \sigma_{\overline{X}} &= \frac{\sigma}{\sqrt{N}}\sqrt{\frac{N_p - N}{N_p - 1}} \end{aligned}$$

olur.
- Yığın **sonsuz sayıda** elemandan oluşuyorsa veya örnekler yerine konularak seçiliyorsa N_p eleman arasından N tanesi $(N_p)^N$ farklı şekilde seçilebilir. Eğer alınan elemanlar geri konmuyorsa bu durumda N_p eleman arasından N tanesi $\binom{N_p}{N}$ farklı şekilde seçilebilir.

- Yığın sonsuz sayıda elemandan oluşuyorsa veya örnekler yerine koyularak seçiliyorsa bu durumda yukarıdaki sonuçlar

$$\mu_{\overline{X}} = \mu$$
$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{N}}$$

sonuçlarına indirgenir.

- **Örnek.** Yığın $2, 3, 6, 8, 11$ sayılarından oluşmaktadır. Bu yığından yerine koyarak çekilen 2 şer çaplı bütün örnekleri düşünelim.
  - Yığın ortalamasını (μ)
  - Yığın standart sapmasını (σ)
  - ortalamaların örnek dağılımının ortalamasını $(\mu_{\overline{X}})$
  - ortalamaların örnek dağılımının standart sapmasını $(\sigma_{\overline{X}})$ bulunuz.

- Çözüm.
  - Yığın ortalaması
$$\mu = \frac{2+3+6+8+11}{5} = 6$$
  - Yığın standart sapması
$$\sigma = \sqrt{\frac{(2-6)^2 + (3-6)^2 + (6-6)^2 + (8-6)^2 + (11-6)^2}{5}} = 3.29$$
  - 5 eleman arasından yerine konarak 2 eleman (yani iki çaplı bir örneklem) $5 \cdot 5 = 25$ farklı sayıda seçilebilir. Bu seçimler

(2, 2)	(2, 3)	(2, 6)	(2, 8)	(2, 11)
(3, 2)	(3, 3)	(3, 6)	(3, 8)	(3, 11)
(6, 2)	(6, 3)	(6, 6)	(6, 8)	(6, 11)
(8, 2)	(8, 3)	(8, 6)	(8, 8)	(8, 11)
(11, 2)	(11, 3)	(11, 6)	(11, 8)	(11, 11)

şeklindedir.

- Bu ikili örneklere ait ortalamalar sırasıyla

2.0	2.5	4.0	5.0	6.5
2.5	3.0	4.5	5.5	7.0
4.0	4.5	6.0	7.0	8.5
5.0	5.5	7.0	8.0	9.5
6.5	7.0	8.5	9.5	11.0

şeklindedir. Bu değerlerin ortalaması yani ortalamaların örnekleme dağılımının ortalaması

$$\mu_{\overline{X}} = \frac{\text{Yukarıdaki bütün ortalamaların toplamı}}{25} = 6$$

olarak bulunur. Dikkat edilirse $\mu = \mu_{\overline{X}}$ bulduk.

  - Yukarıdaki ortalamaların standart sapması klasik standart sapma formülü ile elde edilebilir. Bulunacak değer
$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{N}} = \frac{3.29}{\sqrt{2}} = 2.33$$
olarak bulunacaktır.

- **Örnek.** Bir önceki örneği yerine koymadan seçim yapıldığı durum için çözünüz.
- **Çözüm.** Yerine koymadan 2 eleman seçildiğine göre tüm seçimler $\binom{5}{2} = 10$ farklı şekildedir. Bunlar $(2, 3)$, $(2, 6)$, $(2, 8)$, $(2, 11)$, $(3, 6)$, $(3, 8)$, $(3, 11)$, $(6, 8)$, $(6, 11)$, $(8, 11)$ dir.
  - Sonuç bir önceki sorudaki ile aynıdır.
  - Sonuç bir önceki sorudaki ile aynıdır.
  - yukarıdaki ikli örneklerin ortalamaları sırasıyla $2.5, 4, 5, 6.5, 4.5, 5.5, 7, 7, 8.5, 9.5$ olduğundan
$$\mu_{\overline{X}} = \frac{2.5+4+5+6.5+4.5+5.5+7+7+8.5+9.5}{10} = 6$$
  - Ortalamaların örnekleme dağılımının standart sapması
$$\sigma_{\overline{X}} = \sqrt{\frac{(2.5-6)^2 + (4-6)^2 + \cdots + (9.5-6)^2}{10}} = \sqrt{4.05} = 2.01$$
olur. Bunun yerine bu standart sapmayı
$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{N}}\sqrt{\frac{N_p - N}{N_p - 1}} = \frac{3.29}{\sqrt{2}}\sqrt{\frac{5-2}{5-1}} = 2.01$$
ile de bulabilirdik.

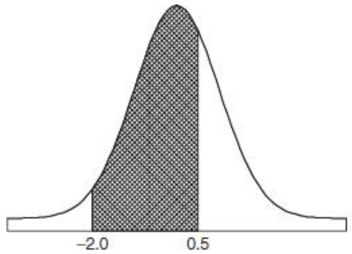
- N nin çok büyük değerleri ($N \geq 30$) için ortalamaların örnekleme dağılımı, yığına bakılmaksızın ortalaması $\mu_{\overline{X}}$, standart sapması $\sigma_{\overline{X}}$ olan yaklaşık olarak normal dağılıma sahip olur. Yığın normal dağılıma sahip ise bu durumda örnekleme dağılımı N nin küçük ($N < 30$) olması durumunda bile normal dağılıma sahip olacaktır.
- **Örnek.** Bir üniversitedeki 3000 bayan öğrencinin boylarının ortalaması 68 ve standart sapması 3 inç ile normal dağılmaktadır. Her biri 25 öğrenciden oluşan 80 örnek bulunsun.
  - Yerine koyarak,
  - Yerine koymadan beklenen ortalamayı ve ortalamaların örnekleme dağılımlarının sonuçlarının standart sapmasını bulunuz.
- **Çözüm.** Seçilecek örneklerin sayısı yerine konulmayacaksa 3000^{25}, yerine konulacaksa $\binom{3000}{25}$ tir. Bu değerler oldukça büyük olduklarından gerçek örnekleme dağılımını bulamayız Bunun yerine sadece deneysel örnekleme dağılımını bulabiliriz. İki örnekleme birimi de birbirine oldukça yakındır.
  - $$\mu_{\overline{X}} = \mu = 68 \text{ ve } \sigma_{\overline{X}} = \frac{\sigma}{\sqrt{N}} = \frac{3}{\sqrt{25}} = 0.6$$
  - $$\mu_{\overline{X}} = \mu = 68 \quad \text{ve} \quad \sigma_{\overline{X}} = \frac{\sigma}{\sqrt{N}}\sqrt{\frac{N_p - N}{N_p - 1}} = \frac{3}{\sqrt{25}}\sqrt{\frac{3000-25}{3000-1}} = 0.6$$

  Görüldüğü gibi her iki değer de birbirine yakındır.

- **Örnek.** Bir önceki örnekteki verileri kullanarak boy uzunluğu 66.8 ile 68.3 arasında olan öğrenci sayısını bulunuz.
- **Çözüm.** Standart puanlar

$$z_1 = \frac{\overline{X} - \mu_{\overline{X}}}{\sigma_{\overline{X}}} = \frac{66.8 - 68}{0.6} = -2 \quad \text{ve} \quad z_2 = \frac{\overline{X} - \mu_{\overline{X}}}{\sigma_{\overline{X}}} = \frac{68.3 - 68}{0.6} = 0.5$$

olduğundan bir öğrencinin boy uzunluğunun 66.8 ile 68.3 arasında olma olasılığı aşağıdaki alana eşit olacaktır.



Standart normal eğri tablosundan istenen alan 0.6687 olacaktır. Örneklerin beklenen değeri yani boy uzunluğu istenen aralıkta olana öğrenci sayısı

$$\mu = Np = 80 \cdot 0.6687 = 53.496 \approx 53$$

olarak bulunur.

8.4 Oranların Örnekleme Dağılımları

- Yığın sonsuz ve bir olayın meydana gelme olasılığı (istenen durum olarak veya başarılı olarak isimlendirilir) p, meydana gelmeme olasığı $q = 1 - p$ olsun. Bu yığından seçilen bütün N çaplı mümkün örnekleri ve her bir örnekteki istenen durumların oranının P olduğunu kabul edelim. Bu durumda ortalaması $\mu_{\overline{X}}$, standart sapması $\sigma_{\overline{X}}$ olan oranların örnekleme dağılımı

$$\mu_P = p \text{ ve } \sigma_P = \sqrt{\frac{pq}{N}} = \sqrt{\frac{p(1-p)}{N}}$$

ile bulunur. Bu eşitlikler daha önce verilen

$$\mu_{\overline{X}} = \mu \text{ ve } \sigma_{\overline{X}} = \frac{\sigma}{\sqrt{N}}$$

eşitliklerinde $\mu = p$ yazılarak bulunur. Yığın çok büyük olduğunda ($N \geq 30$) örnekleme dağılım yaklaşık olarak normal dağılır. Eğer yığın sonsuz sayıda elemandan oluşuyorsa bu durumda bu eşitlikler aşağıdaki gibi olacaktır.

$$\mu_P = p \text{ ve } \sigma_P = \sqrt{pq} = \sqrt{p(1-p)}$$

- Kesikli değişkenlere ait oranların örnekleme dağılımı bulunurken standart puanlara $\frac{1}{2N}$ düzeltmesi uygulanır ve bu değerler

$$z_1 = \frac{P_1 - \mu_P - \frac{1}{2N}}{\sigma_P} \text{ ve } z_2 = \frac{P_2 - \mu_P + \frac{1}{2N}}{\sigma_P}$$

şeklinde elde edilir. Böylece kesikli değişken sürekli hale getirilir ve normal dağılım uygulanır.
- **Örnek.** Hilesiz bir para 120 defa atıldığında %40 ile %60 arasında tura gelme olasılığını bulunuz.,
- **Çözüm.** İki farklı yöntem ile yapılabilir. İlki daha önce kullandığımız yöntemdir.
  - **Birinci yöntem.** 120 atışın %40'ı 48, %60'ı 72 dir. Turaların sayısı kesikli bir değişken olduğu için bu değişkeni turaların sayısını $[47.5, 72.5]$ aralığında alarak sürekli hale getiririz. Bu durumda turalar için beklenen değer yani ortalama $\mu = Np = 120 \cdot \frac{1}{2} = 60$ ve standart sapma $\sigma = \sqrt{Npq} = \sqrt{120 \cdot \frac{1}{2} \cdot \frac{1}{2}} = 5.48$ olur. Bu durumda standart puanlar

$$z_1 = \frac{\overline{X} - \mu_{\overline{X}}}{\sigma_{\overline{X}}} = \frac{47.5 - 60}{5.48} = -2.28 \quad \text{ve} \quad z_2 = \frac{\overline{X} - \mu_{\overline{X}}}{\sigma_{\overline{X}}} = \frac{72.5 - 60}{5.48} = 2.28$$

olacaktır. İstenen olasılık alan tablosundan 0.9774 olarak bulunacaktır.

- İkinci yöntem ise oranların örnekleme dağılımı kullanılarak yapılacaktır.
  - **İkinci yöntem.** Ortalama ve standart sapma değerleri

$$\mu_P = p = \frac{1}{2} \quad \text{ve } \sigma_P = \sqrt{\frac{pq}{N}} = \sqrt{\frac{\frac{1}{2}\cdot\frac{1}{2}}{120}} = 0.0456$$

Diğer taraftan $P_1 = \frac{40}{100} = 0.4$ ve $P_2 = \frac{60}{100} = 0.6$ olduğundan standart puanlar

$$\begin{aligned} z_1 &= \frac{P_1 - \mu_P - \frac{1}{2N}}{\sigma_P} = \frac{0.4 - 0.00417 - 0.5}{0.0456} = -2.28 \\ z_2 &= \frac{P_2 - \mu_P + \frac{1}{2N}}{\sigma_P} = \frac{0.6 + 0.00417 - 0.5}{0.0456} = 2.28 \end{aligned}$$

dir. İstenen alan yine bir önceki gibi 0.9774 olarak bulunacaktır.

- **Örnek.** Belirli bir makine tarafından üretilen aletlerin %2 si bozuktur. Bu makine tarafından üretilen 400 aletten
  - %3 veya daha fazlasının
  - %2 veya daha azının kusurlu bulunma olasılığını bulunuz.
- **Çözüm.**

$$\mu_P = p = 0.02 \text{ ve } \sigma_P = \sqrt{\frac{pq}{N}} = \sqrt{\frac{0.02\cdot 0.98}{400}} = 0.007$$

  - $P_1 = \frac{3}{100} = 0.03$

$$z_1 = \frac{P_1 - \mu_P - \frac{1}{2N}}{\sigma_P} = \frac{0.03 - \frac{1}{800} - 0.02}{0.007} = 1.25$$

olduğundan istenen olasılık $z = 1.25$ in sağındaki alan olan 0.1056 değerine eşit olacaktır.

  - $P_2 = \frac{2}{100} = 0.02$

$$z_1 = \frac{P_1 - \mu_P + \frac{1}{2N}}{\sigma_P} = \frac{0.02 + \frac{1}{800} - 0.02}{0.007} = 0.18$$

olduğundan istenen olasılık $z = 0.18$ in solundaki alan olan 0.5714 değerine eşit olacaktır.

- **Örnek.** Seçim sonuçları belirli bir adayın tüm oyların %46 sını aldığını göstersin. Oy verenlerin tamamının arasından tesadüfi olarak seçilen 200 kişiden en az 101 tanesinin oylarını bu adaya verme olasılığını bulunuz.
- **Çözüm.**

$$\mu_P = p = 0.46 \text{ ve } \sigma_P = \sqrt{\frac{pq}{N}} = \sqrt{\frac{0.46\cdot 0.54}{200}} = 0.0352$$

$P_1 = \frac{101}{200} = 0.505$ olduğundan

$$z_1 = \frac{P_1 - \mu_P - \frac{1}{2N}}{\sigma_P} = \frac{0.505 - \frac{1}{400} - 0.46}{0.0352} = 1.21$$

olur. Standart normal eğri altında bulunan ve $z = 1.21$ değerinin sağında bulunan alan istenen olasılığa eşittir ve bu değer $0.5 - 0.3869 = 0.1131$ dir.

8.5 Farkların ve Toplamların Örnekleme Dağılımı

- İki farklı sonlu yığın verilsin. Birinci yığından N_1 çaplı bir örnek için S_1 istatistiği hesaplansın. S_1 istatistiği için ortalama μ_{S_1}, standart sapma σ_{S_1} olsun. Benzer şekilde ikinci yığından N_2 çaplı bir örnek için S_2 istatistiği hesaplansın. S_2 istatistiği için ortalama μ_{S_2}, standart sapma σ_{S_2} olsun. Bu durumda $S_1 - S_2$ farkı için örnekler birbirinden bağımsız olmak koşulu ile ortalama ve standart sapma

$$\begin{aligned} \mu_{S_1-S_2} &= \mu_{S_1} - \mu_{S_2} \\ \sigma_{S_1-S_2} &= \sqrt{(\sigma_{S_1})^2 + (\sigma_{S_2})^2} \end{aligned}$$

şeklinde olacaktır.

- Eğer yığın sonsuz elemanlı ise veya seçilen örnekler geri konuyorsa

$$\begin{aligned} \mu_{S_1-S_2} &= \mu_{S_1} - \mu_{S_2} \\ \sigma_{S_1-S_2} &= \sqrt{\frac{(\sigma_{S_1})^2}{N_1} + \frac{(\sigma_{S_2})^2}{N_2}} \end{aligned}$$

olacaktır.

- Eğer P_1 ve P_2 oranları için farkların ortalaması ve standart sapması bulunmak isteniyorsa ve yığın sonsuz sayıda elemandan oluşuyorsa

$$\begin{aligned} \mu_{P_1-P_2} &= \mu_{P_1} - \mu_{P_2} \\ \sigma_{P_1-P_2} &= \sqrt{\frac{p_1 q_1}{N_1} + \frac{p_2 q_2}{N_2}} \end{aligned}$$

eşitliği kullanılır.
- N_1 ve N_2 değerleri 30 dan büyük eşit olduğunda oranların farklarının örnekleme dağılımı normal dağılıma yakınsar.
- Toplamlara ait ortalama ve standart sapmalar ise örnekler birbirinden bağımsız olduklarında ve yığın sonsuz sayıda elemandan oluştuğunda

$$\begin{aligned} \mu_{S_1+S_2} &= \mu_{S_1} + \mu_{S_2} \\ \sigma_{S_1+S_2} &= \sqrt{(\sigma_{S_1})^2 + (\sigma_{S_2})^2} \end{aligned}$$

bağıntıları kullanılarak bulunur.
- Benzer bağıntılar yığın sonlu sayıda elemandan oluştuğunda da bulunabilir.

- **Örnek.** S_1 yığını $3, 7, 8$ ve S_2 yığını ise $2, 4$ olsun. Buna göre
  - μ_{S_1} ve μ_{S_2}
  - $\mu_{S_1-S_2}$
  - σ_{S_1} ve σ_{S_2}
  - $\sigma_{S_1-S_2}$ yi hesaplayınız.,
- **Çözüm.**
  - $\mu_{S_1} = \dfrac{3+7+8}{3} = 6$ ve $\mu_{S_2} = \dfrac{2+4}{2} = 3$
  - S_1 yığını ile S_2 yığınlarının elemanları farkı

$$\begin{array}{cccccccc} 3-2 & 7-2 & 8-2 & \text{veya} & 1 & 5 & 6 \\ 3-4 & 7-4 & 8-4 & & -1 & 3 & 4 \end{array}$$

olduğundan $\mu_{S_1-S_2} = \dfrac{1+5+6+(-1)+3+4}{6} = 3$ tür. Bunu yerine daha kısa yoldan $\mu_{S_1-S_2} = \mu_{S_1} - \mu_{S_2} = 6 - 3 = 3$ şeklinde de yapılabilir.
  - $\sigma_{S_1} = \sqrt{\dfrac{(3-6)^2 + (7-6)^2 + (8-6)^2}{3}} = \sqrt{\dfrac{14}{3}}$ ve

$$\sigma_{S_1} = \sqrt{\frac{(2-3)^2 + (4-3)^2}{2}} = 1$$

  - $\sigma_{S_1-S_2}$ klasik yoldan yapılabilir. Bunun yerine

$$\sigma_{S_1-S_2} = \sqrt{(\sigma_{S_1})^2 + (\sigma_{S_2})^2} = \sqrt{\left(\sqrt{\frac{14}{3}}\right)^2 + 1^2} = \sqrt{\frac{17}{3}} \text{ bulunur.}$$

- **Örnek.** A fabrikasının ürettiği elektrik ampülleri 200 saatlik standart sapma ile 1400 saatlik ortalamam ömre sahiptir. B firmasının ürettiği ampüller 100 saatlik standart sapma ile 1200 saatlik ortalama ömre sahiptir. Her bir firmadan 125 tesadüfi örnek test edildiğinde A firmasının ürettiği ampüllerin ortalama ömrünün B firmasından en az 160 saat fazla olma olasılığını bulunuz. (Üretilen ampül sayısının sonsuz sayıda olduğunu kabul ediniz)
- **Çözüm.** $\mu_A = 1400, \mu_B = 1200, \sigma_A = 200$ ve $\sigma_B = 100$ olarak verilmiş. Her bir firmadan 125 tesadüfi örnek test edildiğine göre $N_1 = N_2 = 125$ tir. Buna göre,

$$\begin{aligned} \mu_{S_1-S_2} &= \mu_{S_1} - \mu_{S_2} = 1400 - 1200 \\ \sigma_{S_1-S_2} &= \sqrt{\frac{(\sigma_{S_1})^2}{N_1} + \frac{(\sigma_{S_2})^2}{N_2}} = \sqrt{\frac{200^2}{125} + \frac{100^2}{125}} = 20 \end{aligned}$$

Bu durumda 160 saat fark için standart puan

$$z = \frac{160 - 200}{20} = -2$$

olacağından standart normal eğri altında $z = -2$ nin sağındaki alan 0.9772 istenen olasılıktır.

- **Örnek.** A ve B yazı tura oyunu oynamaktadır. Her biri hilesiz bir parayı 50 defa atmaktadır. A, B den 5 veya daha fazla sayıda sayıda tura atarsa oyunu kazanacak aksi halde oyunu B kazanacaktır. Oyunu A nın kazanma olasılığı nedir?
- **Çözüm.** PA ile A nın tura oranını PB ile B nin tura oranını gösterelim. Para hilesiz olduğuna göre $\mu_{PA} = \mu_{PB} = p = \frac{1}{2}$ dir. Bu durumda

$$\begin{aligned} \mu_{PA} - \mu_{PB} &= \frac{1}{2} - \frac{1}{2} = 0 \\ \sigma_{PA-PB} &= \sqrt{\frac{p_1 q_1}{N_A} + \frac{p_2 q_2}{N_B}} = \sqrt{\frac{\frac{1}{2}\cdot\frac{1}{2}}{50} + \frac{\frac{1}{2}\cdot\frac{1}{2}}{50}} = 0.1 \end{aligned}$$

olur. A, B den 5 veya daha fazla sayıda sayıda tura atarsa oyunu kazanacağına göre 5 kesikli değişkeni sürekli hale getirilirse tura gelme sayısı 4.5 veya daha fazla sayıda tura anlamına gelir. Bu durumda tura gelme oranı $P = \dfrac{4.5}{50} = 0.09$ ve bu orana ait standart puan

$$z = \frac{P - \mu_P - \frac{1}{2N}}{\sigma_P} = \frac{0.09 - \frac{1}{100} - 0}{0.1} = 0.8$$

olur. Standart normal eğri altında $z = 0.8$ in sağındaki alan yani istenen olasılık $0.5 - 0.2881 = 0.2119$ olacaktır.