3. BÖLÜM
MERKEZİ EĞİLİM ÖLÇÜLERİ
(DAĞILIM ÖLÇÜLERİ)

MEDYAN, MOD ve KARTİLLER

Verilerin incelenmesinde ortalama olmayan ama ortalamalarda kullanılan medyan (Ortanca), mod ve kartil kavramları mevcuttur. Şimdi onları inceleyelim.

Medyan (Ortanca)

3.1. Tanım: Serisinin büyüklük sırasına göre serilmiş bir rakamın orta noktasında yer alan değere medyan denir. Serisindeki sayıların sayısı tek ise tam ortadaki sayı medyandır. Serisindeki sayıların sayısı çift ise tam ortadaki bu iki sayının aritmetik ortalaması medyandır.

Örnek: {5, 7, 8, 13, 20} serisinin medyanı 8 dir. Çünkü 8 sayısı serisinin tam ortadaki sayıdır.

Örnek: {104, 182, 302, 308, 510, 640} serisi 6 tane olduğundan çift elemanlı olduğundan medyan,

$$\frac{302+308}{2}=305$$

dir.//

Sınıflandırılmış serisinde medyan $\frac{n+1}{2}$ inci terimin değeridir.

Örnek: Bir sınıftaki öğrencilerin kiloları (x_i) ve frekans (f_i) dağılımı aşağıdaki şekildedir.

Öğrenci Sayısı (f_i)	Kilosu (x_i)
5	32
6	33
7	34
5	35
2	36

Bu sınıfta 25 öğrenci olduğundan $\frac{n+1}{2}=\frac{25+1}{2}=13$. terim medyandır. 13. terim 34 olduğundan bu serisinin medyanı 34 dür.

3.2. Tanım: Sınıflandırılmış frekans dağılımında serisinin her terimi toplanarak ayrıca yazılıyorsa, bu yazıma kümülatif frekans denir. Eğer sınıflandırılmış frekans dağılımında kümülatif frekans varsa, $\frac{n}{2}$ inci terimi içeren sınıf medyan sınıfıdır.

Örnek: Yukarıdaki örnek kümülatif frekansı;

Öğrenci Sayısı (f_i)	Kilosu (x_i)	Kümülatif Frekans
5	32	5
6	33	11
7	34	18
5	35	23
2	36	25

$$\frac{n}{2}=\frac{25}{2}=12{,}5<13$$

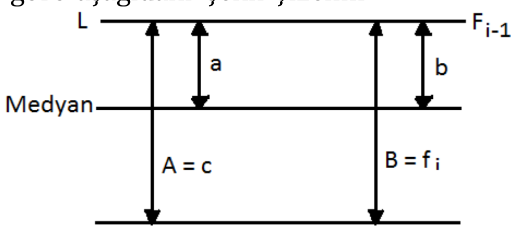
olur.

3.1. Teorem: Gruplandırılmış frekans dağılımında,
L medyan sınıfının alt sınır değeri,
c sınıf aralığı,
n toplam frekans sayısı,
f_i medyanı içeren sınıf frekansı,
F_{i-1} medyan sınıfından önceki sınıfların (kümülatif) frekansları toplamı olmak üzere gruplandırılmış frekans dağılımının medyanı;

$$\text{Medyan}=L+\frac{c}{f_i}\left(\frac{n}{2}-F_{i-1}\right)$$

şeklindedir.

İspat: Verilere göre aşağıdaki şekli çizelim.



$$\frac{A}{a}=\frac{B}{b}$$

$$\frac{c}{\text{Medyan}-L}=\frac{f_i}{\frac{n}{2}-F_{i-1}}$$

$$\text{Medyan}=L+\frac{c}{f_i}\left(\frac{n}{2}-F_{i-1}\right)$$

Örnek:

Sınıf	f_i	Kümülatif Frekans
0-2 den az	2	2
2-4 den az	4	6
4-6 den az	3	9
6-8 den az	1	10

Çözüm: Kümelatif frekanslarda $\frac{n}{2}$ inci terimi içeren sınıf medyan sınıf olacağından $\frac{n}{2}=\frac{10}{2}=5$ dir.

5. terim 2-4 aralığında olunca $L=2$ dir.

5. terim 2-4 aralığında sınıf aralığı $c=2$ dir.

5. terimi içeren sınıf frekansı $f_i=4$ dir.

5. teriminden önceki tüm sınıfların frekansları toplamı $F_{i-1}=2$ dir.

$$\text{Medyan}=L+\frac{c}{f_i}\left(\frac{n}{2}-F_{i-1}\right)=2+\frac{2}{4}\left(\frac{10}{2}-2\right)=3{,}5$$

2.3. Uyarı: 1. Medyan üzerinde cebirsel işlemler yapılmaz.

2. Farklı alt serilerin medyanları biliniyorsa, bu serileri birleştiğinde medyan değerleri değişir.

Mod (Tepe Değeri)

3.3. Tanım: Bir serisinde en çok tekrarlanan değere mod (tepe değeri) denir.

Örnek: Bir sınıftaki 7 öğrencinin serisi aşağıdaki gibidir.

Kız, Erkek, Erkek, Kız, Kız, Erkek, Kız

Bu serisinde en çok tekrar eden cinsiyet "Kız" olduğu için modu "Kız"dır.

Örnek: {15, 22, 18, 19, 15, 18, 20, 15} serisi için tepe değerini bulmak için her değerin frekansı (sıklığı) bulunur:

Değer	**Frekans**
15	3
18	2
19	1
20	1
22	1

En fazla tekrarlanan değer 15'dir (3 kez) bu nedenle bu dağılımın mod (tepe değeri) 15'dir.

3.2. Teorem: Gruplandırılmış frekans dağılımında,
L En büyük sıklığın bulunduğu sınıfın alt sınırı,
c Sınıf aralığı,
d_1 En büyük sıklığın bir önceki sıklığı,
d_2 En büyük sıklığın bir sonraki sıklığı
olmak üzere gruplandırılmış frekans dağılımının modu (tepe noktası);

$$\text{Mod} = L + c\left(\frac{d_1}{d_1 + d_2}\right)$$

şeklindedir.

3.1. Teoreme benzer yolla ispatı yapılır.

Örnek:

Veriler (x_i)	Frekanslar (f_i)
11-20	3
21-30	5
31-40	8
41-50	4
51-60	2
Toplam	22

L = 32, d_1= 8 – 5 = 3, d_2 = 8 – 4 = 4 olduğundan

$$\text{Mod}=L+c\left(\frac{d_1}{d_1+d_2}\right)=32+10\left(\frac{3}{3+4}\right)=36,28$$

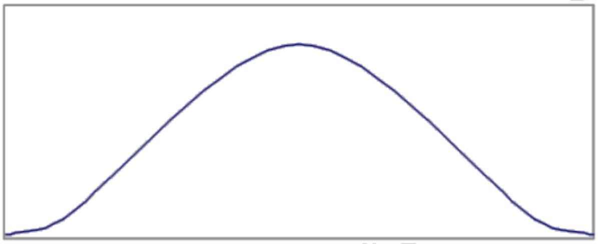
bulunur.

2.4. Uyarı: 1. Mod üzerinde cebirsel işlemler yapılmaz.

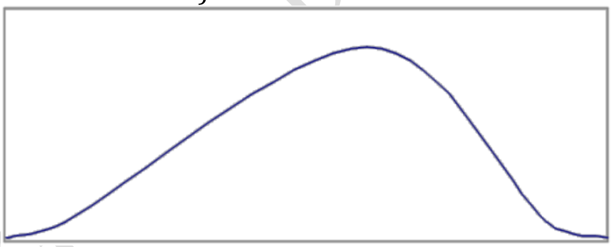
2. Bazı verilerin ortalaması ve medyanı olduğu halde modu olmayabilir.

Aritmetik Ortalama, Medyan (Ortanca), Mod (Tepe Değeri) Arasındaki İlişkiler

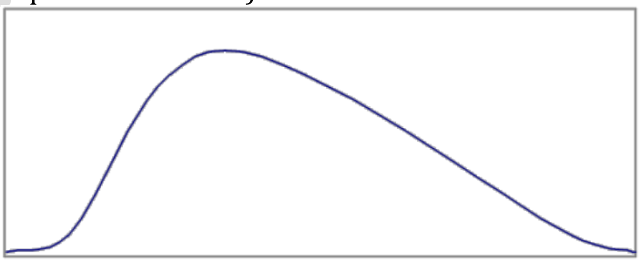
1. Aritmetik Ortalama = Medya = Mod ise; sıklık değeri simetriktir.



2. Aritmetik Ortalama < Medya < Mod ise; sıklık değeri sola doğrudur (ileri de buna sola çarpıktır denecektir).



3. Aritmetik Ortalama > Medya > Mod ise; sıklık değeri sağa doğrudur (ileri de buna sağa çarpıktır denecektir).



Örnek: Bir sınıftaki öğrencilerin boyları (x_i) ve frekans (f_i) dağılımı aşağıdaki şekildedir.

Öğrenci Sayısı (f_i)	Kilosu (x_i)
5	132
6	133
7	134
5	135
2	136

Bu sınıfta 25 öğrenci olduğundan,

Aritmetik Ortalama: $\bar{x} = \dfrac{5.132+6.133+7.134+5.135+2.136}{25} = 133{,}72$

Medyan: $\dfrac{n+1}{2} = \dfrac{25+1}{2} = 13$. terim olup Medyan = 134

Mod: En çok tekrar eden sayı 7 tane olup Mod = 134

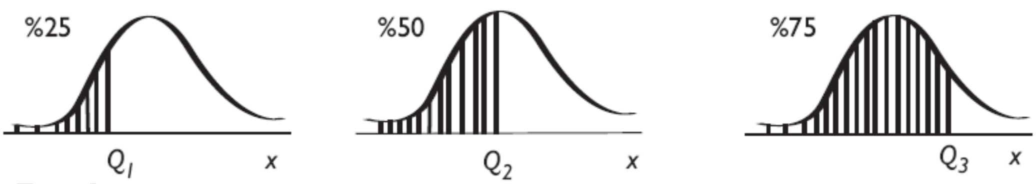
Aritmetik Ortalama < Medya = Mod olup sıklık değeri sola çarpıktır.

Kartiller (Çeyreklikler)

3.4. Tanım: Bir serisini dört eşit değere ayıran değerlere kartil (Çeyreklikler) denir. Kartil 3 değerden oluşur. Bunlar;

Q_1 için $\dfrac{n}{4} + \dfrac{1}{2}$, Q_2 için $\dfrac{n+1}{2}$, Q_3 için $\dfrac{3n}{4} + \dfrac{1}{2}$

şeklinde olur. Dikkat edersek Q_2 aynı zamanda medyan değeridir.



Kartillerin gösterimi

Örnek: Aşağıda verilen serisinde kartilleri bulalım.

{18, 25, 30, 32, 32, 44, 45, 52, 65, 77, 98}

$\dfrac{n}{4} + \dfrac{1}{2} = \dfrac{11}{4} + \dfrac{1}{2} = 3{,}25$ olup 3. terim 1. Kartildir. $Q_1 = 30$

$\frac{n+1}{2}=\frac{11+1}{2}=6$ olup 6. terim 2. Kartildir (Medyandır). $Q_2=44$

$\frac{3n}{4}+\frac{1}{2}=\frac{3.11}{4}+\frac{1}{2}=8{,}75$ olup 9. terim 3. Kartildir. $Q_3=65$

3.3. Teorem: Gruplandırılmış frekans dağılımında,
L medyan içiren sınıflar aralığının alt sınırı,
c sınıf aralığı,
n toplam frekans sayısı,
f_i medyanı içeren sınıf frekansı,
F_{i-1} medyan sınıfından önceki sınıfların (kümülatif) frekansları toplamı olmak üzere gruplandırılmış frekans dağılımının kartilleri;

$$Q_1=L+\frac{c}{f_i}\left(\frac{n}{4}-F_{i-1}\right)$$

$$Q_2=L+\frac{c}{f_i}\left(\frac{n}{2}-F_{i-1}\right)$$

$$Q_3=L+\frac{c}{f_i}\left(\frac{3n}{4}-F_{i-1}\right)$$

şeklindedir.

3.1. Teoreme benzer yolla ispatı yapılır.

Örnek:

Sınıf	f_i	Kümülatif Frekans
10-20 den az	23	23
20-30 den az	32	55
30-40 den az	35	90
40-50 den az	10	100

Çözüm: **1. Kartil** hesaplanması;

Kümelatif frekanslarda $\frac{n}{4}$ inci terimi içeren sınıf medyan sınıf olacağından $\frac{n}{4}=\frac{100}{4}=25$ dir.

25. terim 20-30 aralığında olunca $L=20$ dir.

25. terim 20-30 aralığında sınıf aralığı $c=10$ dir.

25. terimi içeren sınıf frekansı $f_i = 32$ dir.

25. teriminden önceki tüm sınıfların frekansları toplamı $F_{i-1} = 23$ dür.

$$Q_1 = L + \frac{c}{f_i}\left(\frac{n}{4} - F_{i-1}\right) = 20 + \frac{10}{32}\left(\frac{100}{4} - 23\right) = 20{,}63$$

2. Kartil (Medyan) hesaplanması;

Kümelatif frekanslarda $\frac{n}{2}$ inci terimi içeren sınıf medyan sınıf olacağından $\frac{n}{2} = \frac{100}{2} = 50$ dir.

50. terim 20-30 aralığında olunca $L = 20$ dir.

50. terim 20-30 aralığında sınıf aralığı $c = 10$ dir.

50. terimi içeren sınıf frekansı $f_i = 32$ dir.

50. teriminden önceki tüm sınıfların frekansları toplamı $F_{i-1} = 23$ dür.

$$Q_2 = L + \frac{c}{f_{med}}\left(\frac{n}{2} - f_a\right) = 20 + \frac{10}{32}\left(\frac{100}{2} - 23\right) = 28{,}44$$

3. Kartil hesaplanması;

Kümelatif frekanslarda $\frac{3n}{4}$ inci terimi içeren sınıf medyan sınıf olacağından $\frac{3n}{4} = \frac{3.100}{4} = 75$ dir.

75. terim 30-40 aralığında olunca $L = 30$ dir.

75. terim 30-40 aralığında sınıf aralığı $c = 10$ dir.

75. terimi içeren sınıf frekansı $f_i = 35$ dir.

75. teriminden önceki tüm sınıfların frekansları toplamı $F_{i-1} = 55$ dür.

$$Q_3 = L + \frac{c}{f_i}\left(\frac{3n}{4} - F_{i-1}\right) = 30 + \frac{10}{35}\left(\frac{3.100}{4} - 55\right) = 35{,}71$$

olur.

3.5. Tanım: Q_1 ve Q_3 kartiller olmak üzere,

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

eşitliğine çeyrek ayrılış (sapma) denir.

Örnek: {16, 17, 5, 4, 18, 2, 3, 6, 7} serisinin çeyrek ayrılışını bulunuz.

Çözüm: {2, 3, 4, 5, 6, 7, 16, 17, 18}

Medyan = 6

$$Q_1 = \frac{3+4}{2} = 3{,}5 \text{ ve } Q_1 = \frac{16+17}{2} = 16{,}5$$

$$Q = \frac{Q_3 - Q_1}{2} = \frac{16{,}5 - 3{,}5}{2} = 6{,}5$$

Desil (Ondabirlikler) ve Santiller (Yüzdebirlikler)

3.6. Tanım: Bir serisini on eşit değere ayıran değerlere desil (ondabirlikler) ve yüz eşit değere ayıran değerlere santil (yüzdebirlikler) denir. Desil 9, santil 99 değerden oluşur. Medyan beşinci deşile, ellinci santile eşittir.

Desil ve santil, kartilin aynı özelliklerini sağladığından tekrar anlatıma ihtiyaç duyulmamıştır.

DAĞILIM GENİŞLİĞİ (ARALIK)

Bazı serilerinin modu, medyanı veya aritmetik serileri aynı olabilir. Bu durum analizler için yeterli kalmayabilir. Bazen dağılım genişliği (aralığına) da ihtiyaç vardır.

3.7. Tanım: Bir veri kümesindeki en büyük değer ile en küçük değer arasındaki farka dağılım genişliği (aralığı) denir.

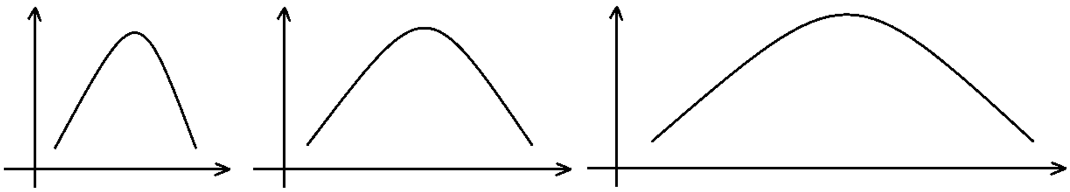
DG = Maksimum – Minimum

Örnek: X = {48, 49, 49, 50, 51, 53}
Y = {35, 41, 50, 55, 58, 61}
Z = {17, 21, 33, 49, 88, 92}

serilerinin aritmetik ortalaması 50 dir. Ama bu üç serilerinin dağılım genişlikleri,

DG (X) = 53–49 = 4
DG (Y) = 61–35 = 26
DG (Z) = 90–15 = 75

olur. Bu dağılımın grafiği aşağıdaki şekildeki gibidir.



Örnek: 2 firmanın 7 personellerinin 1 günlük kazançları aşağıda verilmiştir.

A firması: {348, 304, 328, 360, 342, 330, 325}
B firması: {346, 275, 318, 397, 352, 337, 310}

A ve B Firmaları için Dağılım genişlikleri:

A Firması için DG =360 – 307=53
B Firması için DG= 397 – 275=122

dir.

ORTALAMA SAPMA

Veriler arasında her zaman homojenlik olmaz. Veriler arasında farklı farklı değerler bulunabilir. Bu değerleri ortalaması ve standartları nedir ve nasıl bulunmalıdır. Bu durum dağılım ölçüleri içerisinde incelenir. Örneğin 15 yaşındaki bir gencin ortalama boyu ne olmalıdır. Hangi durumda uzun hangi durumda kısa denecektir. Veya bir kanda bulunan kolesterolün değerleri hangi aralıkta olursa zarar vermez. Sağlıklı insanda hangi aralıkta kolesterol bulunur. Veya bir bitkinin normal şartlar altında yetiştirilecek bir dönümden alınacak ürün miktarının standardı ne olmalıdır? Gibi sorulara cevabı dağılım ölçüleri içerisinde araştırılacaktır. Bunlardan ortalama sapma ve standart sapma olmak üzere iki çeşittir.

Ortalama sapma sınıflandırılmış (sıralı) ortalama sapma ve gruplandırılmış ortalama sapma olarak ikiye ayrılır.

3.8. Tanım: X serisinin aritmetik ortalaması $\overline{x}$ olsun;

$$O_s = \frac{1}{N}\sum_{i=1}^{N}|x_i - \bar{x}|$$

değerine ortalama sapma denir. Bu ana kitlede örneklem yapıldığında

$$O_s = \frac{1}{n-1}\sum_{i=1}^{n}|x_i - \bar{x}|$$

denklemi oluşur.

Dikkat edersek ana kitle tanımlanırken N sayısına bölünürken örneklemde n sayısına değil n–1 sayısına bölünmektedir. Çünkü n terim üzerinde incelenirse $\sum_{i=1}^{n}|x_i - \bar{x}| = 0$ olma durumu söz konusu olacağından n–1 terim sağlıklı sonuç verecektir.

Örnek: Bir ilkokul 1. sınıf öğrencileri arasında ağırlıkları {18, 19, 20, 21, 22, 24, 25, 26, 27, 28} kg olarak tespit ediliyor. Elde edilen bu ağırlıklara göre 1. sınıf öğrencilerinin ortalama sapmasını bulunuz.

Çözüm:

$$\bar{x} = \frac{22+23+21+20+24+24+23+25+25+23}{10} = 23$$

x_i	$\bar{x}$	$x_i - \bar{x}$	$\lvert x_i - \bar{x}\rvert$
18	23	−5	5
19	23	−4	4
20	23	−3	3
21	23	−2	2
22	23	−1	1
24	23	1	1
25	23	2	2
26	23	3	3
27	23	4	4
28	23	5	5
Σ = 10			Σ = 14

$$O_s = \frac{1}{n-1}\sum_{i=1}^{n}|x_i - \bar{x}| = \frac{15}{10-1} = \pm 1{,}67$$

şeklindedir. Elde edilen bu değerlere göre ortalama 23 kg olan öğrencilerin ±1,56 ortalama sapmaya uygundur.

3.1. Aksiyom: Semboller ortalama sapmadaki değerler olmak üzere;

i) Sınıflandırılmış ortalama sapma; $O_s = \dfrac{\sum_{i=1}^{N} f_i.|x_i - \bar{x}|}{\sum_{i=1}^{N} f_i}$

ii) Gruplandırılmış ortalama sapma; $O_s = \dfrac{\sum_{i=1}^{N} f_i.|m_i - \bar{x}|}{\sum_{i=1}^{N} f_i}$

şeklindedir.

Örnek: Bir hastanede 1 ayda doğan bebeklerin kiloları ve frekansları aşağıdaki şekildeki gibidir. Doğan çocukların ortalama sapmalarını bulunuz.

Bebek kiloları	Bebek Sayısı (f_k)
2,750	8
3,000	9
3,250	7
3,500	4
3,750	3
4,000	2
	$\Sigma = 35$

Çözüm:

$$\bar{x} = \frac{2{,}75.8 + 3.12 + 3{,}25.7 + 3{,}5.4 + 3{,}75.3 + 4.1}{33} = 3$$

x_i	$\bar{x}$	$x_i - \bar{x}$	F_i	$f_i.\|x_i - \bar{x}\|$
2,750	3	– 0,25	8	2
3,000	3	0	9	0
3,250	3	0,25	7	1,75
3,500	3	0,5	4	2
3,750	3	0,75	3	2,25
4,000	3	1	2	2
			$\Sigma = 33$	$\Sigma = 10$

$$O_s = \frac{\sum_{i=1}^{N} f_i.|x_i - \bar{x}|}{\sum_{i=1}^{N} f_i - 1} = \frac{10}{33-1} = \pm 0{,}313$$

Örnek: Bir hastanede bir ayda sigara sebebinden ölenlerin gruplandırılmış frekans dağılımı aşağıdaki şu şekildedir. (f_k = Hasta sayısı)

Yaş Aralığı	Frekans Ortalaması (x_k)	f_k
$35 \leq x < 45$	40	10
$45 \leq x < 55$	50	20
$55 \leq x < 65$	60	50
$65 \leq x \leq 75$	70	20
		$\Sigma = 100$

elde edilen bu verilere göre gruplandırılmış standart sapmayı bulunuz.

$$\bar{x} = \frac{\sum_{i=1}^{n} f_i . x_i}{\sum_{i=1}^{n} f_i} = \frac{10.40 + 20.50 + 50.60 + 20.70}{100} = \frac{5\,800}{100} = 58$$

x_k	$\bar{x}$	$x_k - \bar{x}$	$\lvert x_k - \bar{x} \rvert$	f_k	$\lvert x_k - \bar{x} \rvert . f_k$
40	58	−18	18	10	180
50	58	−8	8	20	160
60	58	2	2	50	100
70	58	12	12	20	240
				$\Sigma = 100$	$\Sigma = 680$

$$O_s = \frac{\sum_{i=1}^{n} \lvert x_i - \bar{x} \rvert . f_i}{\sum_{i=1}^{n} f_i - 1} = \frac{680}{100 - 1} = \pm 6{,}868$$

şeklindedir. Elde edilen bu değerlere göre ortalama 58 olan sigara içenlerin ölüm ortalaması $\pm$ 6,868 gruplandırılmış ortalama sapması çıkmıştır.

STANDART SAPMA

Şimdi standart sapma tanımını verelim. Standart sapma değeri de ortalama sapma değeri gibi, sınıflandırılmış standart sapma ve gruplandırılmış standart sapma olarak iki çeşittir.

3.9. Tanım: X serisinin aritmetik ortalaması $\bar{x}$ olsun;

$$s=\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(x_i-\overline{x}\right)^2} \text{ veya } s=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i^2-\overline{x}^2)}$$

değeri standart sapma denir. Bu ana kitlede örneklem yapıldığında

$$s=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2} \text{ veya } s=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i^2-\overline{x}^2)}$$

denklemi oluşur. Yalnız her zaman geçerli değildir. İleri de 2. formülün hangi durumlarda geçerli olup olmadığı bahsedilecektir.

Örnek: Bir şehirde farklı farklı prizlerin ölçümlerinde {220, 230, 200, 190, 240, 210, 220, 250, 210, 230} volt olarak tespit ediliyor. Elde edilen bu verilere göre şehrin standart sapmasını bulunuz.

Çözüm:

$$\overline{x}=\frac{220+230+200+190+240+210+220+250+210+230}{10}=220$$

1. formül

x_i	$\overline{x}$	$x_i-\overline{x}$	$\left(x_i-\overline{x}\right)^2$
220	220	0	0
230	220	10	100
200	220	−20	400
190	220	−30	900
240	220	20	400
210	220	−10	100
220	220	0	0
250	220	30	900
210	220	−10	100
230	220	10	100
Σ = 10			Σ = 3 000

$$s=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2}=\sqrt{\frac{3000}{10-1}}=\pm 18{,}26$$

2. formül

x_i	$\overline{x}$	x_i^2	$x_i^2-\overline{x}^2$
220	220	48 400	0
230	220	52 900	4 500
200	220	40 000	−8 400
190	220	36 100	−12 300

240	220	57 600	9 200
210	220	44 100	–4 300
220	220	48 400	0
250	220	62 500	14 100
210	220	44 100	–4 300
230	220	52 900	4 500
Σ = 10			Σ = 3 000

$$s=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2}=\sqrt{\frac{3000}{10-1}}=\pm 18{,}26$$

şeklindedir. Elde edilen bu değerlere göre ortalama 220 volt olan prizin ±18,26 sınıflandırılmış standart sapmasını verir.

3.2. Aksiyom: Semboller standart sapmadaki değerler olmak üzere;

i) Sınıflandırılmış standart sapma; $s=\sqrt{\dfrac{\sum_{i=1}^{N} f_i.(x_i-\bar{x})^2}{\sum_{k=1}^{N} f_i}}$

ii) Gruplandırılmış standart sapma; $s=\sqrt{\dfrac{\sum_{i=1}^{N} f_i.(m_i-\bar{x})^2}{\sum_{k=1}^{N} f_i}}$

şeklindedir.

Örnek: Bir yüksekokulda okuyan öğrencilerin boy ortalamaları aşağıdaki şu şekildedir. (f_i = Öğrenci sayısı)

Boy Ortalamaları	Frekans Ortalaması (x_i)	f_i
$145 \leq x < 155$	150	10
$155 \leq x < 165$	160	30
$165 \leq x < 175$	170	50
$175 \leq x \leq 185$	180	10
		Σ = 100

elde edilen bu verilere göre gruplandırılmış standart sapmayı bulunuz.

$$\bar{x}=\frac{\sum_{i=1}^{n} f_i.x_i}{\sum_{i=1}^{n} f_i}=\frac{10.150+30.160+50.170+10.180}{10+30+50+10}=\frac{16600}{100}=166$$

x_i	$\bar{x}$	$x_i-\bar{x}$	$(x_i-\bar{x})^2$	f_i	$(x_i-\bar{x})^2.f_i$
150	166	−16	256	10	2 560
160	166	−6	36	30	1 080
170	166	4	16	50	800
180	166	14	196	10	1960
				Σ = 100	Σ = 6 400

$$s=\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2 f_i}{\sum_{i=1}^{n} f_i-1}}=\sqrt{\frac{6400}{99}}=\pm 8{,}04$$

şeklindedir. Elde edilen bu değerlere göre ortalama 166 olan öğrencilerin ± 8 gruplandırılmış standart sapması çıkmıştır.

3.4. Teorem: Ortalama sapma O_s, Standart sapma σ olmak üzere;

$$O_s \le \sigma$$

$$\frac{1}{N}\sum_{i=1}^{N}|x_i-\bar{x}| \le \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^2}$$

eşitsizliği mevcuttur.

2.2. Teoreminde x_i yerine $x_i-\bar{x}$ alınırsa istenen elde edilir.

VARYANS ve KOVARYANS

3.10. Tanım: Verilerin ortalama etrafında nasıl yayıldıklarını (dağıldıklarının) ölçüsüne varyans denir.

$$V(X)=s^2=\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^2$$

s^2 denklemi varyansı verir ve V(X) ile gösterilir.

Örnek: Bir firma 8 günlük televizyon satış miktarı aşağıdaki verildiği gibidir. Bu serisinin varyansını bulunuz?

{8, 15, 14, 10, 21, 17, 5, 30}

Çözüm: Verilerin aritmetik ortalaması

$$\overline{x}=\frac{8+15+14+10+21+17+5+30}{8}=15$$

olmak üzere varyası ise;

1. formül

x_k	$\overline{x}$	$x_k-\overline{x}$	$(x_k-\overline{x})^2$
8	15	−7	49
15	15	0	0
14	15	−1	1
10	15	−5	25
21	15	6	36
17	15	2	4
5	15	−10	100
30	15	15	225
n = 8			Σ = 440

$$V(X)=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x})^2=\frac{440}{8-1}=62{,}86$$

2. formül

x_k	$\overline{x}$	x_k^2	$x_k^2-\overline{x}^2$
8	15	64	−161
15	15	225	0
14	15	196	−29
10	15	100	−125
21	15	441	216
17	15	289	64
5	15	25	−200
30	15	900	675
n = 8			Σ = 440

$$V(X)=\frac{1}{n-1}\sum_{i=1}^{n}(x_i^2-\overline{x}^2)=\frac{440}{8-1}=62{,}86$$

3.5. Teorem: Kareli ortalamanın karesi ile aritmetik ortalamanın karesi arasındaki fark varyansa eşittir.

$$V(X) = s^2 = KO^2 - \bar{x}^2$$

Bu teoremin ispatı aşikârdır.

3.3. Aksiyom: Semboller standart sapmadaki değerler olmak üzere;

i) Sınıflandırılmış serilerin varyansı;

$$s^2 = \frac{\sum_{i=1}^{N} f_i.(x_i - \bar{x})^2}{\sum_{i=1}^{N} f_i} \quad \text{veya} \quad s^2 = \frac{\sum_{i=1}^{N} f_i.(x_i^2 - \bar{x}^2)}{\sum_{i=1}^{N} f_i}$$

ii) Gruplandırılmış serilerin varyansı;

$$s^2 = \frac{\sum_{i=1}^{N} f_i.(m_i - \bar{x})^2}{\sum_{i=1}^{N} f_i} \quad \text{veya} \quad s^2 = \frac{\sum_{i=1}^{N} f_i.(m_i^2 - \bar{x}^2)}{\sum_{i=1}^{N} f_i}$$

şeklindedir.

3.11. Tanım: X ve Y iki seri olsun.

$$Cov(XY) = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})$$

denklemine kovaryans denir ve Cov(XY) ile gösterilir.

Kovaryans, seriler arasındaki ilişkinin varlığını ve yönünü (aynı veya ters) belirleyen bir ölçüdür. Kovaryans iki değişkenin birlikte değime derecesini gösterir. İki değişken aynı yönde değişiyorsa, büyük X ve Y ya da küçük ile X ile Y olaylarının meydana gelme frekansı diğer durumlardan daha büyük olur.

2.3. Sonuç: 1. Bir X serisinde bütün terimlerine aynı sayı eklenir veya çıkarırsak serisinin varyansı değişmez.

$$V(X) = V(X \pm K)$$

$$\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2 = \frac{1}{N}\sum_{i=1}^{N}[(x_i \pm K) - (\bar{x} \pm K)]^2$$

2. Bir serisinin bütün terimlerini aynı sayılsa çarptığımızda (böldüğümüzde) varyans çarptığımızda (böldüğümüzde) sayının karesiyle orantılı olarak büyür (küçülür).

$$V(cX)=c^2V(X)$$
$$\frac{1}{N}\sum_{i=1}^{N}(cx_i-c\bar{x})^2=\frac{c^2}{N}\sum_{i=1}^{N}(x_i-\bar{x})^2$$

$$V\left(\frac{X}{c}\right)=\frac{1}{c^2}V(X)$$
$$\frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i}{c}-\frac{\bar{x}}{c}\right)^2=\frac{1}{c^2N}\sum_{i=1}^{N}(x_i-\bar{x})^2$$

3. Birbiri ile ilişki iki serisinin terimlerinin karşılıklı olarak toplanması (çıkarılması) suretiyle elde edilen serinin varyansı, bu serinin varyanslarının toplamı ile kovaryansının 2 katının toplamına (farkına) eşittir.

$$V(X+Y)=V(X)+V(Y)+2Cov(XY)$$
$$V(X-Y)=V(X)+V(Y)-2Cov(XY)$$

$$\frac{1}{N}\sum_{i=1}^{N}[(x_i+y_i)-(\bar{x}+\bar{y})]^2$$
$$=\frac{1}{N}\sum_{i=1}^{N}[(x_i-\bar{x})+(y_i-\bar{y})]^2$$
$$=\frac{1}{N}\sum_{i=1}^{N}[(x_i-\bar{x})^2+(y_i-\bar{y})^2+2(x_i-\bar{x})(y_i-\bar{y})]$$
$$=\frac{1}{N}\left[\sum_{i=1}^{N}(x_i-\bar{x})^2+\sum_{i=1}^{N}(y_i-\bar{y})^2+2\sum_{i=1}^{N}(x_i-\bar{x})(y_i-\bar{y})\right]$$

Not: Birbirinden bağımsız iki serisinin terimlerinin karşılıklı olarak toplanması (çıkarılması) suretiyle elde edilen serinin varyansı, bu serinin varyansları toplamına eşittir. X ve Y birbirinden bağımsız olduğundan Cov(XY) = 0 olur.

$$V(X+Y)=V(X)+V(Y)$$
$$V(X-Y)=V(X)+V(Y)$$

Örnek: 70 erkekteki ürik asit konsantrasyon sonuçları aşağıdaki gibi bulunmuştur. Buna göre varyansını ve standart sapmasını bulunuz.

Ürik Asit (mg/dl)	(m_i)	(f_i)	$f_i\, m_i$	$f_i\, m_i^2$
1,6-2,0	1,8	6	10,8	19,44
2,1-2,5	2,3	7	16,1	37,03
2,6-3,0	2,8	20	56,0	156,80
3,1-3,5	3,3	21	69,3	228,69
3,6-4,0	3,8	12	45,6	173,29
4,1-4,5	4,3	4	17,2	73,96
		Σ = 70	Σ = 215	Σ = 689,21

$$V(X)=\frac{\sum_{i=1}^{N} f_i(m_i-\bar{x})^2}{\sum_{i=1}^{N} f_i}$$

$$=\frac{1}{\sum_{i=1}^{n} f_i-1}\left(\sum_{i=1}^{n} f_i m_i^2-\frac{\left(\sum_{i=1}^{n} f_i m_i\right)^2}{\sum_{i=1}^{n} f_i}\right)$$

$$=\frac{689{,}21-\frac{215^2}{70}}{69}$$
$$=0{,}419$$

Standart sapma $s=\sqrt{0{,}419}=0{,}647\,\text{mg}/\text{dl}$ olur.

STANDART HATA, GÜVEN ARALIĞI ve DEĞİŞİM KATSAYISI

Standart Hata

3.12. Tanım: Standart sapmanın denek sayısının kareköküne bölümüne standart hata denir ve s_x ile gösterilir. Yani

$$s_x=\frac{s}{\sqrt{n}}=\sqrt{\frac{V(X)}{n}}$$

dir.

Örnek: Bir önceki örneğin standart hatasını bulunuz.

Çözüm: $s_x = \frac{s}{\sqrt{n}} = \frac{0,647}{\sqrt{70}} = 0,077$ //

Standart hata örnekleme dağılımındaki ortalamaların standart sapmasıdır. Seçilen örneklerde ortalamalar arasındaki yaygınlığı gösterir. Standart hata örnek büyüklüğünün fonksiyonudur. Böylece n değeri arttıkça hata miktarı küçülür.

Standart hata ortalamalar ile ilgilidir, deneklerle ilgili değildir. Aritmetik ortalamada oluşan hatanın belirlemesi için yapılır. Ayrıca standart sapma gibi değişkenin dağılışı hakkında bilgi vermez. //

Güven Aralığı

3.13. Tanım: Ortalamalar O ve standart hata s_x ile göstermek üzere

$$O \pm 2\, s_x$$

denklemine güven aralığı denir.

Örnek: İki grup üzerinde yapılan bir çalışmada grupların yaş ve boy değerine bakılıyor. Bu çalışmada aşağıdaki sonuçlar elde ediliyor. Bu sonuçlara göre güven aralığını bulunuz.

1. Grup	2. Grup
Ortalama = 14,2	Ortalama = 16,7
Standart hata = 0,9	Standart Hata = 0,6

Çözüm: Güven aralığı;

1. Grup $14,2 \pm 0,9$ olup $[13,3\ ;\ 15,1]$

2. Grup $16,7 \pm 0,6$ olup $[16,1\ ;\ 17,3]$

sınırları içindedir.

Değişim (Varyasyon) Katsayısı

3.14. Tanım: Standart sapma σ, aritmetik ortalama $\overline{x}$ olmak üzere;

$$DK = \frac{s}{\overline{x}}$$

denklemine değişim (varyasyon) katsayısı denir ve DK ile gösterilir.

Örnek: X_i = {8, 15, 14, 10, 21, 17, 5, 30} serisinin varyansı yukarıda 62,86 olarak bulunmuştu. Bu serinin değişim katsayısı

$$\overline{x} = \frac{8+15+14+10+21+17+5+30}{8} = 15$$

$$DK=\frac{\sigma}{\bar{x}}=\frac{\sqrt{62,86}}{15}=0,5258$$

Değişim katsayısı % 52,58 dir. //

Bu seri 5 ile 30 arasında serpiştirilmiş ve aritmetik ortalaması 15 olmuştur. Bu sebepten bu serisi %52,58 farklılık göstermektedir.

Değişim katsayısı farklı serilerin değişkenliklerini kıyaslamada bir ölçü olabilir. Basit (frekansız), sınıflandırılmış ve gruplanmış seriler için uygun olan bu ölçü, seriler arasındaki cins ve büyüklük farklılığını ortadan kaldırır.

Değişim katsayısı küçük olan serilerin diğerlerine göre daha az değişken olduğu söylenir. Bunun anlamı ise seri terimlerinin aritmetik ortalama etrafında daha homojen olarak dağıldığıdır.

A ve B gibi iki farklı popülasyondaki değişkenliği karşılaştırmak istersek doğrudan standart sapma veya standart hatalara bakmak yanıltıcı olabilir.

Ana kitlenin ortalamaları büyüklük olarak birbirinden çok farklı ise standart ayrılış edilir, bu değişim (varyasyon) katsayısıdır, karşılaştırmada bu kullanılır. Örneğin fiillerin ağırlığı mı daha değişkendir, insanların ağırlığı mı sorusu cevabı için değişim katsayısı kullanılır.

Aynı şekilde özellikler farklı birimlerle ölçülüyorsa, örneğin üzerinde deneme yapılan farelerin kan şekerimi daha değişkendir yoksa vücut ağırlıklarımı sorusu ile karşılaşılsa bunun için değişim katsayısına bakmak gerekir.

Ortalamanın sıfıra yakın olduğu durumlarda değişim (varyasyon) katsayısı güvenirliği azalır. Bu gibi durumlarda dikkatli olmak gerekir.

Örnek: Bay X üzerinde farklı iki yöntemle ölçülen kolesterol miktarları aşağıda verilmiştir.

$X:\{177,\ 193,\ 195,\ 209,\ 226\}$ mg / ml
$Y:\{192,\ 197,\ 200,\ 202,\ 209\}$ mg / ml

Bu iki serisinin analizi;

$AO(X)=\frac{100}{5}=200$	$AO(Y)=\frac{100}{5}=200$
$V(X)=\frac{201360-200000}{4}=340$	$V(X)=\frac{200158-200000}{4}=39,5$
$s=18,44$	$s=6,28$

$DK(X)=\dfrac{18{,}44}{200}=0{,}0922$ yani % 9,22	$DK(X)=\dfrac{6{,}28}{200}=0{,}0314$ yani % 3,14

X yönteminin değişim katsayısı Y yönteminden büyük olduğundan, X yöntemi daha değişkendir Yani Y yöntemi daha sağlıklı sonuç vermektedir.

Örnek: Bir televizyon üreticisinin A ve B olmak üzere iki cins üretmiştir. Televizyonların ortalama ömürleri $\bar{x}=74750$ saat ve $\bar{y}=93750$ saattir, standart sapmaları $s_A=14000$ saat ve $s_B=15300$ saattir. Buna göre hangi televizyon daha büyük değişime sahiptir?

Çözüm: $DK_A=\dfrac{14000}{74750}=0{,}1873$ olduğundan DK % 18,73

$$DK_B=\frac{15300}{93750}=0{,}1632 \text{ olduğundan DK \% } 16{,}32$$

olur. Buna göre A türü televizyonların değişim katsayısı daha fazla olduğundan B türü değişim katsayısı tercih edilmelidir.

MONMENTLER, ÇARPIKLIK ve BASIKLIK

Momentler

3.15. Tanım: X serisinin aritmetik ortalaması $\bar{x}$ ve sapmanın derecesi r olsun;

$$m_r=\frac{1}{N}\sum_{i=1}^{N}x_i^r$$

denklemine sıfıra göre moment ve

$$\mu_r=\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^r$$

denklemine aritmetik ortalamaya göre moment denir. Bu serilerde $r=1$ için $\mu_1=0$ ve $r=2$ için $\mu_1=\sigma^2$ olup varyanstır. (Örnekleme göre benzer şekilde tanımlanır.)

Sıfıra göre momentler bilindiğinde, aritmetik ortalamaya göre momentler bulunabilir. Bu şekilde elde edilen denklemlere König Teoremi adı verilir.

$$\mu_2=m_2-m_1^2$$
$$\mu_3=m_3-3m_1m_2+2m_1^3$$
$$\mu_4=m_4-4m_1m_3+6m_1^2m_2-3m_1^4$$

Örnek: X={1, 3, 4, 6, 10, 12} serisinin 4. dereceye kadar sıfıra ve aritmetik ortalamaya göre momentini bulunuz.

Çözüm:

x_i	x_i^2	x_i^3	x_i^4	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^3$	$(x_i - \bar{x})^4$
1	1	1	1	–5	25	–125	625
3	9	27	81	–3	9	–27	81
4	16	64	256	0	0	–8	16
6	36	216	1296	4	16	0	0
10	100	1000	20736	6	36	64	256
12	144	1728	32370	0	90	216	1296
Σ = 36	Σ=306	Σ=3036	Σ=32370	Σ = 0	Σ = 90	Σ = 120	Σ = 2274

Sıfıra göre momentler:

$$m_1 = \frac{36}{6} = 6,\ m_2 = \frac{306}{6} = 51,\ m_3 = \frac{3036}{6} = 506,\ m_4 = \frac{32370}{6} = 5395$$

Aritmetik ortalamaya göre momentler:

$$\mu_1 = \frac{0}{6} = 0,\ \mu_2 = \frac{90}{6} = 15,\ \mu_3 = \frac{120}{6} = 20,\ \mu_4 = \frac{2274}{6} = 379$$

veya

$$\mu_2 = m_2 - m_1^2 = 51 - 6^2 = 15$$

$$\mu_3 = m_3 - 3m_1m_2 + 2m_1^3 = 506 - 3.6.51 + 2.6^3 = 20$$

$$\mu_4 = m_4 - 4m_1m_3 + 6m_1^2m_2 - 3m_1^4 = 5395 - 4.6.506 + 6.6^2 - 3.6^4 = 379$$

olur.

3.4. Aksiyom: Semboller 3.15. tanımdaki olmak üzere;

i) Sıfıra göre sınıflandırılmış moment; $m_r = \dfrac{\sum_{i=1}^{N} f_i x_i^r}{\sum_{i=1}^{N} f_i}$

ii) Sıfıra göre gruplandırılmış moment; $m_r = \dfrac{\sum_{i=1}^{N} f_i m_i^r}{\sum_{i=1}^{N} f_i}$

iii) Aritmetik ortalamaya göre sınıflandırılmış moment; $\mu_r = \dfrac{\sum_{i=1}^{N} f_i (x_i - \bar{x})^r}{\sum_{i=1}^{N} f_i}$

iv) Aritmetik ortalamaya göre gruplandırılmış moment; $\mu_r = \dfrac{\sum_{i=1}^{N} f_i (m_i - \bar{x})^r}{\sum_{i=1}^{N} f_i}$

şeklindedir. (Örnekleme göre benzer şekilde tanımlanır.)

Örnek: X={2, 3, 4, 6} serisinin frekansları 3, 6, 4, ve 7 dir. 3. dereceye kadar sıfıra ve aritmetik ortalamaya göre sınıflandırılmış momentini bulunuz.

Çözüm:

x_i	f_i	x_i^2	x_i^3	$f_i x_i$	$f_i x_i^2$	$f_i x_i^3$
2	3	4	8	6	12	24
3	6	9	27	18	54	162
4	4	16	64	16	64	256
6	7	36	216	42	252	1512
Σ = 15	Σ = 20			Σ = 82	Σ = 382	Σ = 1954

Sıfıra göre sınıflandırılmış momentler:

$m_1 = \dfrac{82}{20} = 4{,}1$, $m_2 = \dfrac{382}{20} = 19{,}1$, $m_3 = \dfrac{1954}{20} = 97{,}7$

x_i	f_i	$\bar{x}$	$x_i - \bar{x}$	$f_i(x_i - \bar{x})$	$f_i(x_i - \bar{x})^2$	$f_i(x_i - \bar{x})^3$
2	3	4,1	–2,1	–6,3	13,25	–27,783
3	6	4,1	–1,1	–6,6	7,26	–7,26
4	4	4,1	–0,1	–0,4	0,04	–0,004
6	7	4,1	1,9	13,3	25,27	48,013
Σ = 15	Σ = 20			Σ = 0	Σ = 45,8	Σ = 12,24

Aritmetik ortalamaya göre sınıflandırılmış momentler:

$\mu_1 = \dfrac{0}{20} = 0$, $\mu_2 = \dfrac{45{,}8}{20} = 2{,}29$, $\mu_3 = \dfrac{12{,}24}{20} = 0{,}612$

olur.

Örnek: Aşağıda verilen verilere göre 3. dereceye kadar sıfıra ve aritmetik ortalamaya göre gruplandırılmış momentini bulunuz.

Sınıflar	f_i
0- 2 den az	2
2-4 den az	1
4-6 dan az	8
6-8 den az	5

Çözüm:

m_i	f_i	m_i^2	m_i^3	$f_i m_i$	$f_i m_i^2$	$f_i x_i^3$
1	2	1	1	2	2	2
3	1	9	27	3	9	27
5	8	25	125	40	200	1000
7	5	49	343	35	245	1715
	$\Sigma = 16$			$\Sigma = 80$	$\Sigma = 456$	$\Sigma = 2744$

Sıfıra göre gruplandırılmış momentler:

$$m_1 = \frac{80}{16} = 5,\ m_2 = \frac{456}{16} = 28{,}5,\ m_3 = \frac{2744}{16} = 171{,}5$$

m_i	f_i	$\bar{x}$	$m_i - \bar{x}$	$f_i(m_i - \bar{x})$	$f_i(x_i - \bar{x})^2$	$f_i(x_i - \bar{x})^3$
1	2	5	–4	–8	32	–128
3	1	5	–2	–2	4	–8
5	8	5	0	0	0	0
7	5	5	2	10	20	40
	$\Sigma = 16$			$\Sigma = 0$	$\Sigma = 56$	$\Sigma = -96$

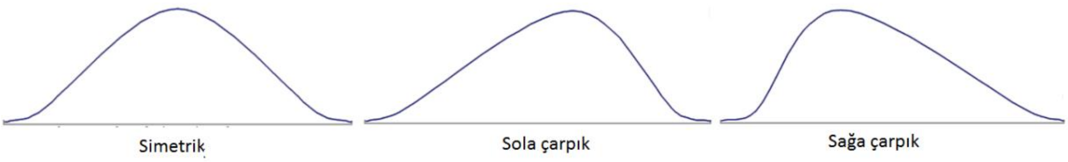
Aritmetik ortalamaya göre gruplandırılmış momentler:

$$\mu_1 = \frac{0}{16} = 0,\ \mu_2 = \frac{56}{16} = 3{,}5,\ \mu_3 = -\frac{96}{16} = -6$$

olur.

Çarpıklık

3.16. Tanım: Bir dağılımının simetrik oluşu veya simetrilikten ayrılış durumuna çarpıklık denir. Yani çarpıklık, bir dağılımın ortalaması etrafındaki asimetri derecesini belirtir. Çarpıklık aşağıdaki şekildeki gibi simetrik, sağa çarpık ve sola çarpık olmak üzere üç çeşittir.



3.17. Tanım: 3. dereceden moment μ_3, standart sapma s olmak üzere;

$$ÇK = \frac{\mu_3}{s^3}$$

denklemine çarpıklık katsayısı denir ve ÇK ile gösterilir. Bir seride;

i) ÇK = 0 ise dağılım simetriktir.

ii) ÇK < 0 ise dağılım sola doğru çarpık ya da – yöne eğilimlidir.

iii) ÇK > 0 ise dağılım sağa doğru çarpık ya da + yöne eğilimlidir.

iv) Çarpıklık katsayısı 0'a yakınsa zayıf çarpık, sıfırdan uzaksa kuvvetli çarpık adı verilir.

Örnek: X={2, 3, 4, 6} serisinin frekansları 3, 6, 4, ve 7 dir. Çarpıklık katsayısını bulunuz.

Çözüm:

x_i	f_i	$\bar{x}$	$x_i-\bar{x}$	$f_i(x_i-\bar{x})^2$	$f_i(x_i-\bar{x})^3$
2	3	4,1	–2,1	13,25	–27,783
3	6	4,1	–1,1	7,26	–7,26
4	4	4,1	–0,1	0,04	–0,004
6	7	4,1	1,9	25,27	48,013
	Σ = 20			Σ = 45,8	Σ = 12,24

Aritmetik ortalamaya göre sınıflandırılmış momentler:

$$\mu_3 = \frac{\sum_{i=1}^{N} f_i(x_i-\bar{x})^3}{\sum_{i=1}^{N} f_i} = \frac{12,24}{20} = 0,612$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N} f_i.(x_i-\bar{x})^2}{\sum_{k=1}^{N} f_i}} = \sqrt{\frac{45,8}{20}} = 1,513$$

$$ÇK = \frac{\mu_3}{s^3} = \frac{0{,}612}{1{,}513^3} = 0{,}177 > 0$$

Çarpıklık pozitif olduğundan serisi sağa çarpık (eğik), fakat çarpıklık zayıftır.

Not: Çarpıklık katsayısı

$$ÇK \cong \frac{n}{(n-1)(n-2)} \sum_{i=1}^{n} \left(\frac{x_i - \overline{x}}{s} \right)^3$$

ile yaklaşık değeri bulunur.

Örnek: X = {–1,64, –0,35, –0,10, 0,00, 2,83, 1,57} serisinin yaklaşık olarak çarpıklık katsayısını bulunuz ve sonucu yorumlayınız.

Çözüm: $\overline{x} = \dfrac{-1{,}64 - 0{,}35 - 0{,}10 + 0{,}00 + 2{,}83 + 1{,}60}{6} = 0{,}39$

x_i	$\overline{x}$	$x_i - \overline{x}$	$(x_i - \overline{x})^2$	$(x_i - \overline{x})^3$
–1,64	0,39	–2,03	4,121	–8,365
–0,35	0,39	–0,74	0,548	–0,405
–0,10	0,39	–0,49	0,24	–0,118
0,00	0,39	–0,39	0,152	0,059
2,83	0,39	2,44	5,954	14,527
1,60	0,39	1,21	1,464	1,772
			Σ = 12,479	Σ = 8

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - \overline{x})^2} = \sqrt{\frac{12{,}479}{6}} = 1{,}442$$

$$ÇK \cong \frac{n}{(n-1)(n-2)} \sum_{i=1}^{n} \left(\frac{x_i - \overline{x}}{\sigma} \right)^3$$

$$= \frac{6}{(6-1)(6-2)} \cdot \frac{8}{1{,}442^3}$$

$$= 0{,}838 > 0$$

Serisi hafif sağa çarpıktır.

3.18. Tanım: $ÇK_{P1} = \dfrac{\overline{x} - Mod}{s}$ ve $ÇK_{P2} = \dfrac{3(\overline{x} - Medyan)}{s}$

denklemlerine 1. ve 2. Pearson Çarpıklık katsayısı (ölçüsü) denir.

1. Pearson Çarpıklık katsayısı çarpıklık katsayısı gibi simetrik serilerde sıfır, sola çarpık serilerde negatif ve sağa çarpık serilerde pozitiftir.

Aşırı çarpık serilerde (asimetrik-simetrik olmayan) çarpıklık ölçüsü değerleri 1'e yakınsa serinin çarpık olduğu (simetrik olmadığı) söylenebilir.

3.19. Tanım: Q_1, Q_2,Q_3 ler karteller olmak üzere;

$$ÇK_B = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_2) + (Q_2 - Q_1)}$$

denklemine Bowley çarpıklık katsayısı (ölçüsü) denir. Çarpıklık katsayısı 0'a yakınsa seri simetrik bir seridir. +1'e yakın sağa çarpık, –1'e yakın seri sola çarpık denilebilir.

Örnek: Aşağıdaki veriler için Pearson ve Bowley çarpıklık katsayısını bulup, sonucu yorumlayınız?

X = {3, 4, 5, 2, 3, 4, 5, 6, 4, 7}

Çözüm:

AO = 4,3   s = 1,49   Mod = 4   Medyan = 4 Q_1 = 3 Q_3 = 5,25

$$ÇK_{P1} = \frac{\bar{x} - Mod}{s} = \frac{4{,}3 - 4}{1049} = 0{,}2$$

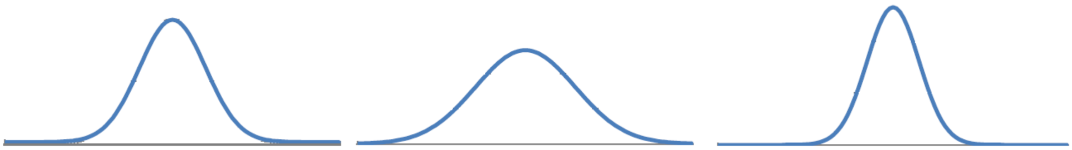
$$ÇK_{P2} = \frac{3(\bar{x} - Medyan)}{s} = \frac{3(4{,}3 - 4)}{1{,}49} = 0{,}6$$

$$ÇK_B = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_2) + (Q_2 - Q_1)} = \frac{(5{,}25 - 4) - (4 - 3)}{(5{,}25 - 4) + (4 - 3)} = 0{,}11$$

Değer pozitif olduğundan, seri hafif sağa çarpık seridir.

Basıklık

3.20. Tanım: Bir dağılımın x eksenindeki veriler büyüklük küçüklüğüne göre normal basık, geniş basık ve dar basık (sivri) olmak üzere üç çeşittir.



Normal Basık   Geniş Basık   Dar Basık (Sivri)

Bir serinin normal olabilmesi için hem simetrik hem de normal bir yüksekliğe sahip olması gerekir.

3.21. Tanım: 4. dereceden moment μ_4, standart sapma s, varyans V(X) olmak üzere;

$$BK = \frac{\mu_4}{s^4} = \frac{\mu_4}{V(X)^2}$$

denklemine basıklık katsayısı denir ve BK ile gösterilir. Bir seride;

i) BK = 3 ise dağılım normal basıklıktadır.

ii) BK < 3 ise dağılım geniş basıklıktadır.

iii) BK > 3 ise dağılım dar basıklıktadır (sivridir).

iv) Basıklık katsayısı 3'a yakınsa zayıf basık, üçten uzaksa kuvvetli basık adı verilir.

Örnek: 4. dereceden momenti $\mu_4 = 7{,}86$, standart sapması $s = 1{,}6$ olan serinin basıklık katsayısını bulunuz ve yorumlayınız.

Çözüm: $BK = \frac{\mu_4}{s^4} = \frac{7{,}86}{1{,}6^4} = 1{,}199 < 3$

olduğundan kuvvetli geniş basıktır.

Örnek: 4. dereceden momenti $\mu_4 = 38$, varyansı $V(X) = 3{,}5$ olan serinin basıklık katsayısını bulunuz ve yorumlayınız.

Çözüm: $BK = \frac{\mu_4}{V(X)^2} = \frac{38}{3{,}5} = 3{,}1 > 3$

olduğundan zayıf dar basıktır (sivridir).

Not: Basıklık katsayısı

$$BK \cong \left[\frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{s}\right)^4\right] - \frac{3(n-1)^2}{(n-2)(n-3)}$$

ile yaklaşık değeri bulunur. Burada 0'a eşitse normal basıklık, sıfırdan büyükse geniş basıklık, sıfırdan küçükse dar basıklık (sivrilik) olduğunu gösterir.

Örnek: X = {3, 4, 5, 2, 3, 4, 5, 6, 4, 7} olan serinin yaklaşık basıklık katsayını bulunuz.

Çözüm: n = 10, AO = 4,3, s = 1,49

$$BK \cong \left[\frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{s}\right)^4\right] - \frac{3(n-1)^2}{(n-2)(n-3)}$$

$$=\left[\frac{10(11+1)}{(10-1)(10-2)(10-3)}\left\{\left(\frac{3-4,3}{1,49}\right)^4+\left(\frac{4-4,3}{1,49}\right)^4+...+\left(\frac{7-4,3}{1,49}\right)^4\right\}\right]-\frac{3(10-1)^2}{(10-2)(10-3)}$$
$$=-0,15$$

Seri hafif dar basıktır (sivridir).

STANDART PUANLAMALAR

Bu veri grubundaki bir puanın, standart sapmaya göre ortalamadan ne kadar uzakta olduğunu tespit edebilmek ya da farklı veri gruplarına ait verileri birbirleriyle karşılaştırabilmek için bu puanların standart sapması yani aynı standart puana dönüştürülmesi gerekmektedir. z ve T puanları bu işlemi gerektiren puanlama türleridir.

Örneğin, farklı derslere ait puanları karşılaştırmak ya da öğrencinin grup içindeki başarısını yorumlamak için bu puanlar kullanılır.

3.22. Tanım: Ham puan HP, aritmetik ortalama AO ve Standart sapma s olmak üzere;

$$z=\frac{HP-AO}{s}$$

denklemine z puanı denir. z puanı herhangi bir ham puanın sınıf ortalamasının kaç standart sapma uzakta olduğunu gösterir.

Örnek: Bir sınavın aritmetik ortalaması 55, standart sapması 2,5 dir. Bu sınavdan 50 ve 65 olan iki öğrencinin z puanı

$$z=\frac{50-55}{2,5}=-2 \text{ ve } z=\frac{65-55}{2,5}=4$$

olur. Bu durum 50 alan öğrenci negatif yönde 2 standart sapma, 65 alan öğrenci pozitif yönde 4 standart sapma uzaklıkta olur.

Örnek: Yabancı Dil Sınavına giren bir öğrencinin 3 sınavdan elde ettiği netleri, Sınava giren öğrencilerin netlerin ortalaması, Sınavın standart sapması aşağıdaki gibidir.

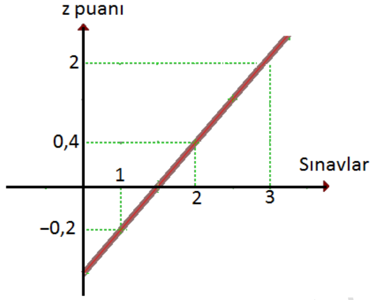
	Net	Sınava giren netlerin ortalaması	Sınavın standart sapması
1. YDS	45	35	5
2. YDS	50	35	8
3. YDS	55	37	9

Bu öğrenci hangi sınavda daha başarılıdır.

$$z_1=\frac{45-35}{5}=2,\; z_2=\frac{50-35}{7}=2{,}14,\; z_3=\frac{55-37}{9}=2$$

2. sınavda daha başarılı olmuştur.

Örnek: Matematik dersine ait üç sınavdan sırasıyla 60, 55 ve 70 notlarını alan bir öğrencinin z puanı aşağıdaki şekilde gibidir.



2. sınavın standart sapması 15 olduğuna göre, 2. Sınavdan 70 alan başka bir öğrencinin z puanını bulunuz.

Çözüm: 2. Sınavdan 55 alan öğrencinin z puanı 0,4 tür.

$$0{,}4=\frac{55-\bar{x}}{15} \text{ ise } \bar{x}=49$$

2. Sınavdan 70 alan başka bir öğrencinin z puanı,

$$z=\frac{70-49}{15}=1{,}4$$

dür. //

z puanı bazı durumlarda negatif çıkmaktadır. Karşılaştırmanın daha rahat anlaşılabilmesi için z puanının T puanına dönüştürülmesi daha uygundur. Şimdi T puanını tanımlayalım.

3.23. Tanım: T = 10.z+50 denklemine T puanı denir.

Örnek: z puanı −2 olan bir puanın T puanını bulunuz.

Çözüm: T = 10.(−2)+50 = 40 dir.

Örnek: z puanı 3 den küçük ise T puanının en büyük doğal sayı değeri nedir?

Çözüm: T < 10.3+50 = 80 ise T nin en büyük değeri 79 dur.

Örnek: Bir öğrencinin T puanı ile z puanının toplamı 72 ise z puanı nedir?

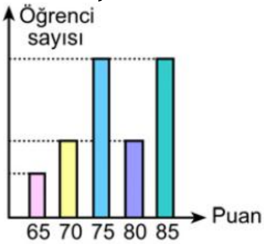
Çözüm: T + z = 72

10.z + 50 + z = 72

z = 2

ÜNİVERSİTE GİRİŞ SINAV SORULARI

1. Tüm değerlerin eşit sayıda tekrar etmediği bir veri grubundaki en çok tekrar eden her bir değer, bu veri grubunun tepe değeri (mod) olmaktadır. 48 öğrencinin bulunduğu bir sınıftaki öğrencilerin tamamı matematik sınavına girmiş ve bu öğrencilerin tamamının bu sınavdan aldıkları puanlara göre sayıca dağılımı aşağıdaki sütun grafiğinde verilmiştir.



Bu sınavdan alınan puanların oluşturduğu veri grubunun tepe değerleri bulunmuş ve puanları bu değerler olan toplam öğrenci sayısının 32 olduğu görülmüştür. Ayrıca, bu sınıfta bu sınavdan 70'ten yüksek puan alan öğrenci sayısı 38 olarak hesaplanmıştır. Buna göre, bu sınıfta bu sınavdan 65 puan alan öğrenci sayısı kaçtır?

A) 2 B) 3 C) 4 D) 5 E) 6

Çözüm: 75 ve 85 alan öğrenci sayısı 32 tanedir (Mod olduğundan)

70 den yüksek 38 öğrenci varsa 80 alan 6 tanedir.

70 ve 80 eşit sayıda alan öğrenci olduğundan 70 alan tanedir.

65 alan öğrenci sayısı $48-38-6=4$ tanedir.

(2019 TYT) Cevap: C