Analitik Olmayan Ortalamalar

Mod (Tepe değeri)

- Verilerin simetrik bir dağılış göstermediği durumlarda iyi bir ölçü olarak düşünülebilir.
- Bir veri grubunda en çok tekrarlanan, yani frekansı en yüksek olan değerdir.
- Tepe değer verilerin en çok hangi değer etrafında toplandığı hakkında bilgi verir.
- Eğer değer dizisinde her değer aynı frekansa sahipse o puan dizisinin modu yoktur.
- Bir seride birden fazla mod olabilir. Bu durumda değişken çok modlu olarak nitelendirilir.
- Eğer serinin histogramı çizilirse, en yüksek sütunun değeri serinin modudur. Bu sebepten mod’a tepe değeri de denir.
- Hesaplanması ve anlaşılması kolaydır
- Dağılımdaki aşırı değerlerden etkilenmez.

Mod (Tepe değeri)

- **Örnek:** Aşağıda satılan bilgisayar sayıları ile ilgili basit ve frekans serileri verilmiştir. Bu serilerin modlarını bulunuz.

Bilgisayar sayısı
30
40
50
50
50
60
60
70

Bilgisayar sayısı	Ay sayısı
30	2
40	4
50	7
60	9
70	4

Mod = 50

Mod= 60

Mod (Tepe değeri)

- **Gruplanmış seride modun hesaplanması**
- Gruplanmış seride modu belirleyebilmek için önce modun içinde bulunduğu sınıf belirlenir. Mod sınıfı frekansı en fazla olan sınıftır. *Gruplanmış seride modun hesaplanabilmesi için serinin sınıf aralığının eşit olması gerekir*. Çünkü sınıf içindeki frekansların dağılımı sınıf aralıklarının büyüklüğüne göre değişir. Eğer sınıf aralıkları eşit değilse eşit hale getirmek gerekir. Eşit hale getirilemiyorsa modun bu şekilde hesaplanması uygun olmaz. Bu sınıf içindeki modun değeri aşağıdaki formülle bulunur.

$$Mod = l_1 + \frac{\Delta_1}{\Delta_1 + \Delta_2} \cdot s$$

Yukarıdaki formülde;

l_1: mod sınıfının alt sınırı

Δ_1: mod sınıfı frekansından bir önceki sınıf frekansının farkı,

Δ_2: mod sınıfı frekansından bir sonraki sınıf frekansının farkı,

s: seride sabit olan sınıf aralığı

Mod (Tepe değeri)

- **Örnek:** Bir ilköğretim okulunda öğrencilerin günlük olarak aldıkları harçlıkların dağılımı aşağıda verilmiştir. Öğrencilerin aldıkları günlük harçlık miktarının ortalamasını mod ile belirleyiniz.

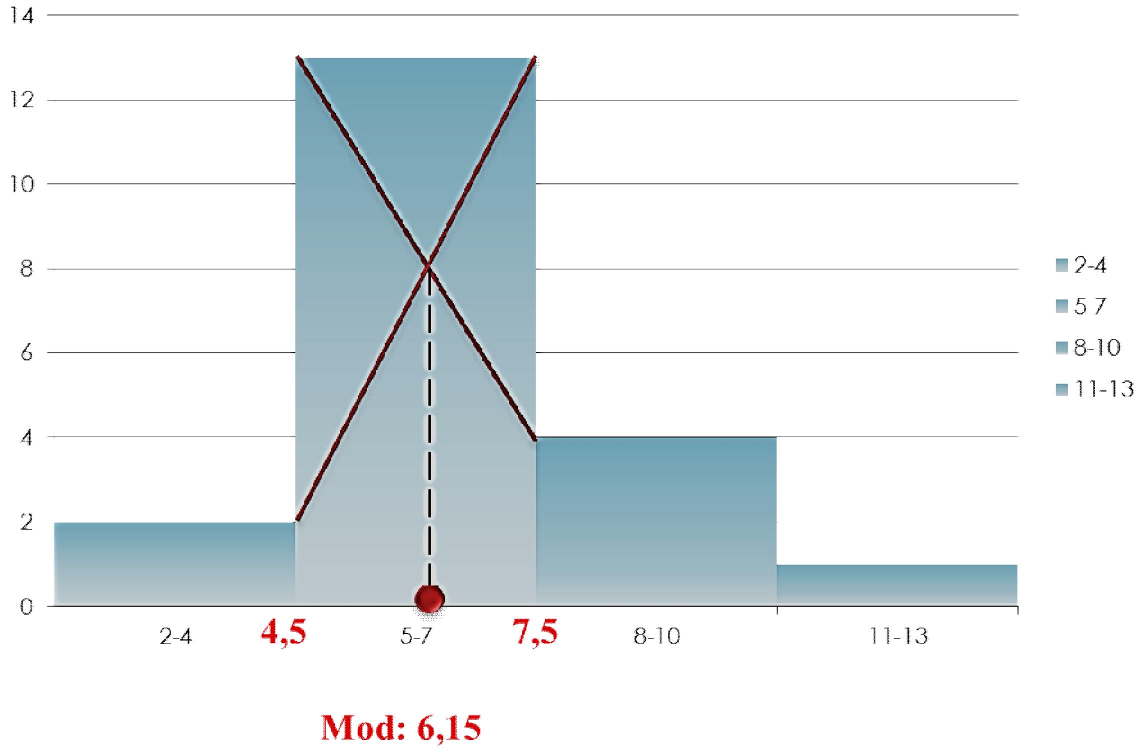
Harçlık (TL/gün)	*Öğrenci sayısı*	
0 – 0,5	30	$l_1 = 1$
0,5 – 1	50	$\Delta_1 = 100 - 50 = 50$
1 – 1,5	100	$\rightarrow$ Mod sınıfı
1,5 – 2	70	$\Delta_2 = 100 - 70 = 30$
2 – 2,5	20	$s = 0{,}5$

$$Mod = l_1 + \frac{\Delta_1}{\Delta_1 + \Delta_2} \cdot s \Rightarrow Mod = 1 + \frac{50}{50+30} \cdot 0{,}5 \Rightarrow Mod = 1{,}3125$$

$$Mod = 1{,}3125 \quad TL/\,gün$$

Histogram ile Mod Tespiti

- Modun grafikle gösterilebilmesi için serinin histogramı çizilir.
- Histogramda en yüksek sütun mod sınıfına karşılık gelir.
- Burada modun yerini tayin etmek için en yüksek sütunun üst köşegenleri ile komşu sütunların bitişik üst köşeleri çapraz olarak birleştirilir.
- İşaretlenen bu noktalar çaprazlama olarak birleştirildiklerinde doğruların kesişme noktasının yatay eksene çizilen doğrunun ekseni kestiği nokta serinin modunu verir.



Modun özellikleri

- Ortalamalar arasında en temsili olanıdır.
- Pratik hayatta çok kullanılan ortalamalardandır.
- Özellikle niteliksel serilerin ortalaması mod ile ifade edilir. Göz rengi, medeni hal, marka, cinsiyet v.s gibi değişkenler niteliksel değişkenler olup sayısal olarak ifade edilemezler
- Mod serideki aşırı değerlere karşı hassas değildir. Çünkü normal serilerde mod genellikle serinin orta bölgesinde yer alır, uç değerlerden etkilenmez.
- Yukarıdaki avantajlarının yanında analitik olmaması sebebi ile matematik işlemlere elverişli olmaması dezavantajıdır.

Medyan (Ortanca)

- **Tanım:** Serideki değerler küçükten büyüğe sıralandığında tam ortaya düşen ve seriyi iki eşit parçaya bölen değere medyan adı verilir.
- **Basit ve tasnif edilmiş seride medyanın bulunuşu:**
- Bunun için serideki değerler küçükten büyüğe sıralanır.
- Daha sonra medyana karşılık gelen değerin sıra değeri $\frac{N+1}{2}$ işlemi ile belirlenir.
- Eğer bu işlemin sonucu tam sayı ise bu sıradaki eleman medyan olarak belirlenmiş olur.
- Eğer bu işlemin sonucu kesirli çıkarsa medyan iki değerin tam ortasına düşeceğinden bu iki değerin ortalaması alınarak medyan bulunur.

Medyan (Ortanca)

- **Örnek:** Bir atölyede çalışan işçilerin belli bir günde ürettikleri kusurlu parçalarının dağılımı verilmiştir. Bu verilerden hareketle işçi başına ortalama kusurlu parça sayısını medyanla belirleyiniz.
- **Çözüm:**

Medyanın serideki sırası

$$\frac{N+1}{2} = \frac{36+1}{2} = 18{,}5$$

Kusurlu parça sayısı	İşçi sayısı		İşçi sayısı (-den az)
10	2	10 ve daha az	2
12	3	12 ve daha az	5
15	4	15 ve daha az	9
16	6	16 ve daha az	15
18	10	18 ve daha az	25
20	5	20 ve daha az	30
21	4	21 ve daha az	34
25	2	25 ve daha az	36
N	36		

$$Medyan = \mathbf{18\ parça}$$

Gruplanmış seride medyanın hesaplanması

- Gruplanmış seriler sürekli karakterde seriler olduğu için medyanın sıra değeri $N/2$ şeklinde bulunur. Bu değer toplam frekansın yarısına eşit olup serinin orta noktasını gösterir. Bu noktayı tespit etmek gruplanmış serilerde basit bir sayma işlemi ile mümkün olmaz. Bu işlemle medyanın içinde bulunduğu sınıf tespit edilir. Belirlenmiş olan medyan sınıfından hareketle aşağıdaki formül yardımı ile medyan değeri hesaplanır.

$$Medyan = l_1 + \frac{\frac{N}{2} - \sum_{i=1}^{m-1} N_i}{N_m} \cdot s_m$$

- l_1 : Medyan sınıfının alt sınırı
- s_m : Medyan sınıfının sınıf aralığı

N_m : Medyan sınıfının frekansı

$N/2$: Medyanın sıra değeri

- $\sum_{i=1}^{m-1} N_i$: Medyandan sınıfından önceki frekanslar toplamı

Gruplanmış seride medyanın hesaplanması

- **Örnek:** Bir işletmede işçilere ödenen saat ücretlerinin dağılımı aşağıda verilmiştir. Bu verilere göre medyan saat ücretini hesaplayınız.

Saat ücreti (Bin TL)	İşçi sayısı		İşçi sayısı	
500 – 600	10	600 den az	10	l_1=700
600 – 700	50	700 den az	60	$N/2$=150/2=75
700 – 800←Medyan sınıfı	40←	800 den az←	100←	$\sum N_i$= 60
800 – 1000	30		130	N_m=40
1000 – 1500	20		150	s_m= 800-700 = 100
Toplam	150			

$\frac{N}{2} = \frac{150}{2} = 75.$ sıradaki değer medyandır. Bu değer 700-800 sınıfına düşmektedir. Bu sınıf içindeki medyan değeri:

$$Medyan = l_1 + \frac{\frac{N}{2} - \sum_{i=1}^{m-1} N_i}{N_m} \cdot s_m \Rightarrow Medyan = 700 + \frac{75-60}{40} \cdot 100 \Rightarrow Medyan = 737{,}5\, TL / saat$$

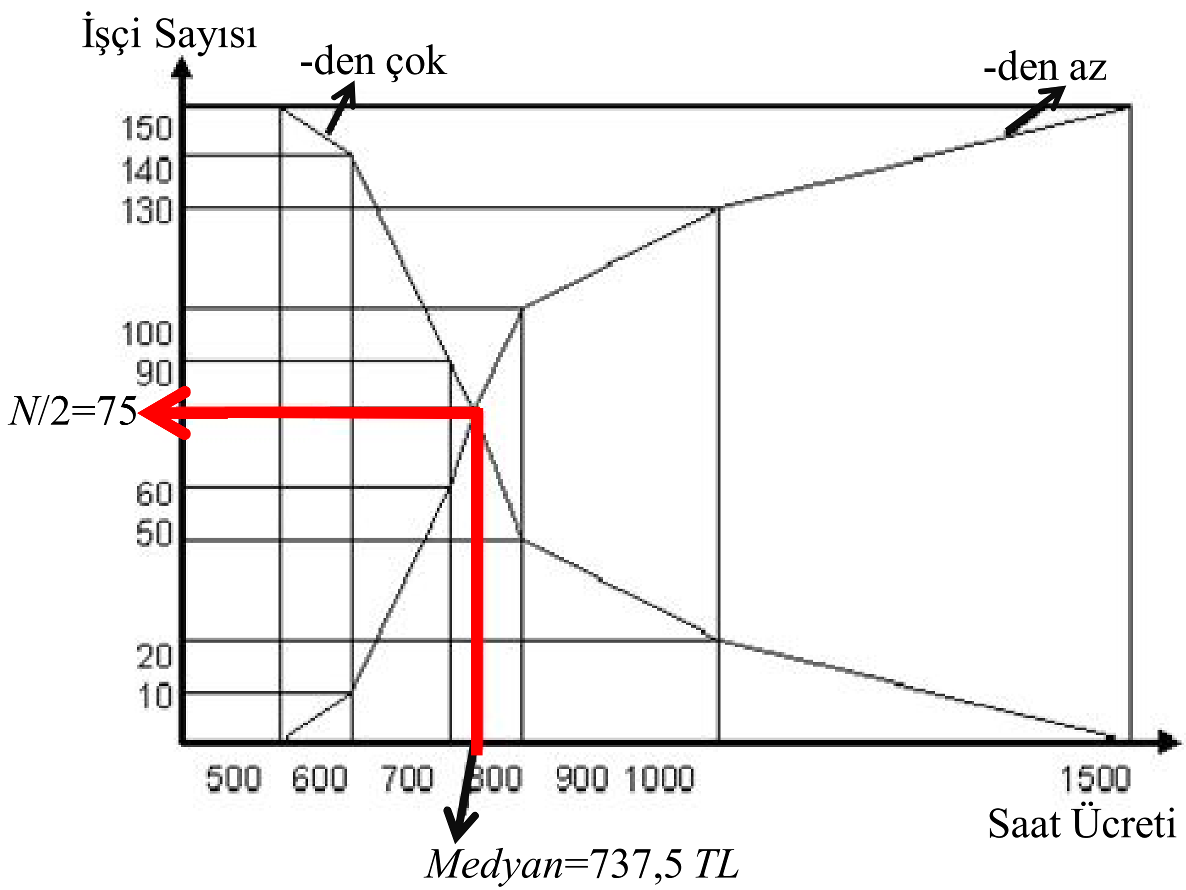
Medyanın grafikle belirlenmesi

- Medyanın grafik üzerinde gösterilebilmesi için kümülatif (-den az) ve ters kümalatif (-den çok) frekans serilerin oluşturulması gerekir.
- Bu serilerin grafiği birlikte çizildiğinde iki eğrinin birbirini kestiği noktadan yatay eksene çizilen doğrunun ekseni kestiği nokta medyan olarak tespit edilir.
- Esasen bu işlemi sadece eğrilerden birini çizmekle de yapabiliriz. Eğrilerden biri çizildiğinde Y ekseninde $N/2$ değerine karşılık gelen noktadan X eksenine paralel çizildiğinde, bu doğrunun kümülatif eğriye temas ettiği noktadan X eksenine çizilen doğrunun ekseni kestiği noktada medyanı gösterecektir.

Medyanın grafikle belirlenmesi

- **Örnek:** Bir önceki örneğin grafikle gösterimi

Kümülatif frekans dağılımı		Ters kümülatif frekans dağılımı	
Saat ücreti (Bin TL)	f_i	Saat ücreti (Bin *TL*)	f_i
500 den az	0	500 den çok	150
600 den az	10	600 den çok	140
700 den az	60	700 den çok	90
800 den az	100	800 den çok	50
1000 den az	130	1000 den çok	20
1500 den az	150	1500 den çok	0



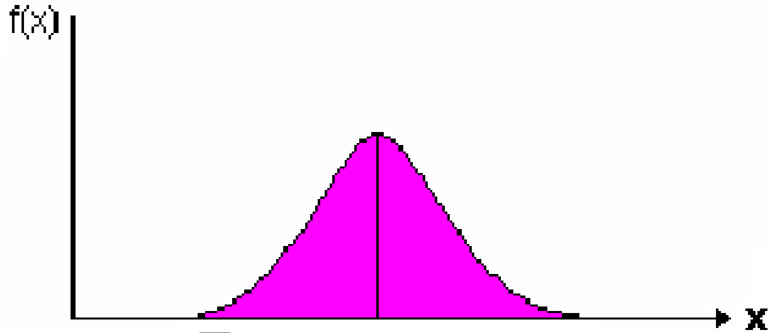
Medyanın özellikleri

- Pratik bir ortalamadır. Sadece basit bir sıralama işlemi gerektirir.
- Özellikle açık sınıflı seriler için medyan daha bir önem kazanır. Bu tür serilerde diğer ortalamalar ya hesaplanamaz, ya da açık sınıflar için sınıf sınırları farazi olarak seçilerek hesaplanabilir.
- Mod ise sınıf aralıklarının eşit olmasını gerektirdiğinden hesaplanamaz. Medyan ise böyle serilerin ortalamasında problemsiz olarak hesaplanabilir.
- Serideki aşırı değerlere karşı hassas değildir. Çünkü medyan serinin ortasına rastladığından, uçlarda oluşan aşırı değerler medyanı etkilemez.
- Medyanın zayıf tarafı serideki bütün değerleri dikkate almaması sebebi ile matematik işlemlere uygun olmamasıdır.

Mod, Medyan ve Aritmetik Ortalama Arasındaki İlişkiler

- Simetrik seride her üç ortalama birbirine eşittir.

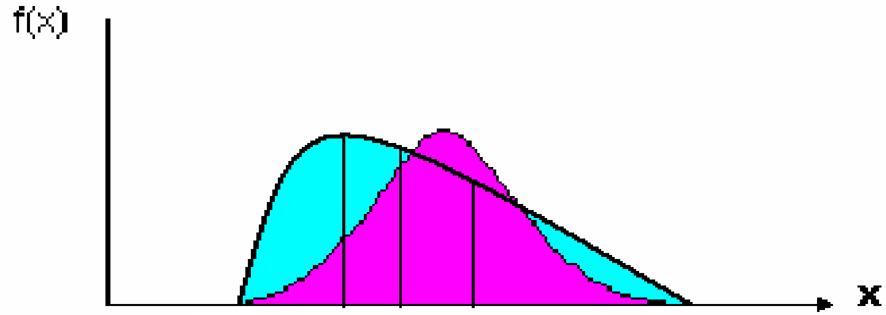
$$\overline{X} = Mod = Medyan$$



$\overline{X} = Medyan = Mod$

- Sağa çarpık serilerde

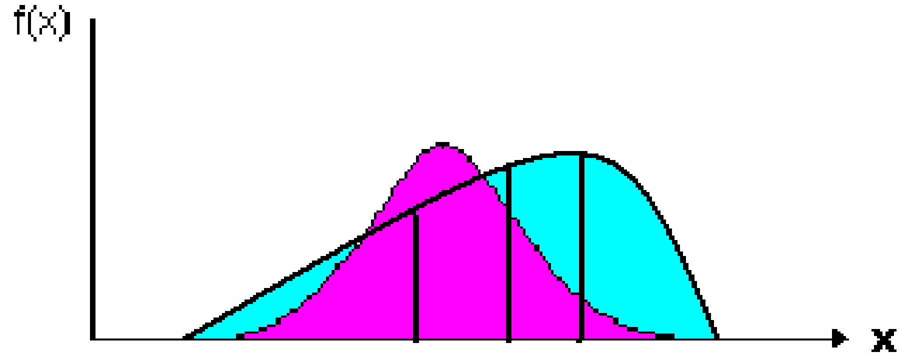
$$\overline{X} > Medyan > Mod$$



$Mod < Medyan < \overline{X}$

- Sola çarpık seride

$$\overline{X} < Medyan < Mod$$



$\overline{X} < Medyan < Mod$

Merkezi Dağılma (Sapma) Ölçüleri

- **Mutlak Sapma Ölçüleri**
  - Değişim aralığı
  - Ortalama mutlak sapma
  - Standart sapma ve Varyans
- **Nispi sapma ölçüleri**
  - Değişim Katsayısı

Sapma ölçüleri

- Bir veri setini meydana getiren elemanlar ortalama değer etrafında belirli bir dağılış gösterirler. Gözlem değerleri arasındaki farklılıktan ileri gelen bu durum istatistik olarak serinin önemli karakteristiklerinden biridir.
- Ortalamalar serinin merkezi noktasını belirlemeye yarayan ölçülerdir. Dağılma ölçüleri ise gözlem değerlerinin bu merkezi noktadan uzaklaşma durumunu ortaya koyan ölçülerdir. Aynı ortalamaya sahip seriler farklı dağılış gösterebilirler. Bu yüzden bir seriyi sadece ortalama değere göre tanımlamak yanlış olur. Bunun yanı sıra dağılışının da bilinmesi gerekir.
- Bir seride ortalamanın temsil kabiliyeti ile dağılma ölçüleri arasında ters bir ilişki vardır. Dağılışı az olan serilerin ortalamaları daha temsili oldukları halde, dağılışı fazla olanların ortalamaları seriyi daha az temsil eder. Bu sayede veri setindeki dağılışın tespiti ortalamanın temsil kabiliyeti hakkında da bilgi verecektir.

Sapma ölçüleri

X_i ve Y_i aşağıdaki gibi iki seri verilmiş olsun:

Xi	Yi
0	2
1	2
2	3
3	4
9	4

Serilerin aritmetik ortalamaları

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{15}{5} = 3$$

$$\bar{Y} = \frac{\sum_{i=1}^{N} Y_i}{N} = \frac{15}{5} = 3$$

- Her iki serinin ortalamaları aynı olmasına rağmen ortalama değer etrafında farklı dağılış göstermektedirler.

Mutlak Sapma ölçüleri

- Mutlak dağılma ölçüleri ilgili değişkenin kendi ölçüldüğü birim cinsinden (kg, cm, TL vs) sonuç verir. Bu sebeple mutlak dağılma ölçüleri olarak adlandırılırlar.

- **Değişim Aralığı**

Gözlem değerlerinin en büyük ve en küçük değeri arasındaki fark olup, verilerin ne kadarlık bir aralıkta değiştiğini gösterir.

$$R = X_{max} - X_{min}$$

X_i : 12,15,20,30,50,52,58,70,90

olan bir serinin değişim aralığı R=90-12 =78 olur. Yani gözlem değerleri 78 birimlik bir aralıkta değişme göstermektedir.

Mutlak Sapma ölçüleri

- Bu dağılım ölçüsü oldukça basit ve anlaşılır olmasına karşılık sadece iki uç değere bağlı olması sebebiyle *serideki aşırı değerlerin etkisi altında kalması* zayıf yönünü oluşturur. Sadece iki uç değeri dikkate alması diğer gözlem değerlerinin dağılımının hiç dikkate alınmamasına sebep olmaktadır.
- **Örnek:** Bir işletmede belli bir parçayı üreten işçinin bu parçayı üretim süresi gözlemlenmiş ve aşağıdaki veriler elde edilmiştir. Bu verilere göre parça üretim süresinin değişim aralığını bulunuz.

Üretim Süresi	Parça Sayısı	$\sum f_i$
30-35	10	10
35-37	30	40
37-40	40	80
40-42	35	115
42-50	20	135
50-60	5	140

- **Çözüm:** $R = X_{max} - X_{min}$

$R = 60\text{-}30 = 30$ dakika

Ortalama (mutlak) Sapma

- Bilindiği gibi sapmalar serisinin $(\sum(X_i - \bar{X}) = 0)$ (aritmetik ortalamadan sapmalar) toplamı sıfıra eşittir.

- Bu durumda sapmalar serisinin ortalaması da sıfır olacağından bir sapma ölçüsü elde etmek mümkün değildir. Serinin toplamını sıfır olmaktan kurtarabilmek için mutlak sapmalar dikkate alınabilir. Çünkü mutlak sapmalar serisinin toplamı sıfırdan büyük olacaktır.

$$(\sum \left| X_i - \bar{X} \right| \geq 0)$$

- Böylece mutlak sapmalar serisinin ortalaması alınarak yeni bir sapma ölçüsü elde edilebilir.

- Bu sapma ölçüsü diğer iki sapma ölçülerinin aksine serinin bütün değerlerini dikkate almaktadır. Bu sebeple daha kullanışlı ve daha temsili bir sapma ölçüsü elde edilmiş olmaktadır.

Ortalama (mutlak) Sapma

- **Basit Seride**

$$O.S = \frac{\sum \left| X_i - \bar{X} \right|}{N}$$

- **Frekans Serisinde**

$$O.S = \frac{\sum fi \left| X_i - \bar{X} \right|}{\sum fi}$$

- **Gruplandırılmış Seride**

$$O.S = \frac{\sum fi \left| m_i - \bar{X} \right|}{\sum fi}$$

Örnek: Bir ağrı kesicinin insanlar üzerinden ne kadar süre ile etkili olduğunu belirlemek için yapılan araştırmada, ağrı kesicinin etkinlik süresinin aşağıdaki gibi dağıldığı gözlenmiştir. Bu verilere göre etkinlik sürenin ortalama sapmasını bulunuz.

Etkin.Sür e (saat)	Hast. Say	m_i	$f_i m_i$	$m_i - \bar{X}$	$fi \left\| m_i - \bar{X} \right\|$
2 - 5	10	3,5	35	-5,8	58
5 - 8	30	6,5	195	-2,8	84
8 - 12	50	10	500	0,7	35
12 - 20	16	16	256	6,7	107,2
Toplam	106		986		284,2

- Aritmetik ortalama

$$\bar{X} = \frac{986}{106} \cong 9{,}3 \; saat$$

- Ortalama sapma

$$O.S = \frac{\sum f_i \left| m_i - \bar{X} \right|}{\sum f_i} = \frac{284{,}2}{106}$$

$$OS = 2{,}68 \text{ saat}$$

Standart Sapma

- Serideki gözlem değerlerinin aritmetik ortalamadan sapmaları (sapmalar serisi) toplamı sıfırdır. Bu durumu mutlak değer almak suretiyle önlenir.

- Ancak bu yol aritmetik işlemler için elverişli olmamaktadır. Mutlak işlemler yerine kare alma yolu ile sapmalar serisi toplamı sıfır olmaktan kurtarılabilir. Böylece yeni bir sapma ölçüsü elde edilmektedir.

- Standart sapma, sapmalar serisinin (aritmetik ortalamadan sapmalar) kareli ortalamasıdır. Yani gözlem değerlerinin aritmetik ortalamadan sapmalarının kareli ortalamasına standart sapma denir.

- Standart sapmanın karesine **varyans** adı verilir. Kütle ve örnek standart sapması için aşağıdaki formüller kullanılır.

Standart Sapma Formülleri

- **Basit Seride**

$$\sigma = \sqrt{\frac{\sum (X_i - \bar{X})^2}{N}}$$

- **Frekans Serisinde**

$$\sigma = \sqrt{\frac{\sum f_i (X_i - \bar{X})^2}{\sum f_i}}$$

- **Gruplandırılmış Seride**

$$\sigma = \sqrt{\frac{\sum f_i (m_i - \bar{X})^2}{\sum f_i}}$$

- **Örnek:** Bir beyaz eşya servis merkezine gelen günlük servis isteklerinin dağılımı ile ilgili aşağıdaki veriler elde edilmiştir. Bu verilere göre servis merkezine gelen günlük servis isteklerinin aritmetik ortalamasını ve standart sapmasını bulunuz.

Servis isteği	$X_i - \overline{X}$	$(X_i - \overline{X})^2$
3	-4	16
4	-3	9
5	-2	4
7	0	0
10	3	9
13	6	36
$\sum X_i = 42$		$\sum 74$

- **Aritmetik ortalama:** $\overline{X} = \dfrac{\sum X}{n} = \dfrac{42}{6} \Rightarrow \overline{X} = 7$

- **Standart sapma:** $\sigma = \sqrt{\dfrac{\sum (X_i - \bar{X})^2}{N}} = \sqrt{\dfrac{74}{6}} = 3{,}51$

- **Örnek:** Doğru, yanlış şeklinde cevap şıkları olan 10 soruya öğrencilerin verdikleri doğru cevap sayılarının dağılımı aşağıda verilmiştir. Bu serinin standart sapmasını ve varyansını bulunuz.

Doğru Cevap. Sayısı	Öğr. Sayısı (f_i)	$f_i X_i$	$X_i - \bar{X}$	$f_i \ (X_i - \bar{X})^2$
2	2	4	-4,69	43,99
3	4	12	-3,69	54,46
4	5	20	-2,69	36,18
5	10	50	-1,69	28,56
6	20	120	-0,69	9,52
7	30	210	0,31	2,88
8	20	160	1,31	34,32
9	10	90	2,31	53,36
10	3	30	3,31	32,87
Toplam	**104**	**696**		**296,15**

- Çözüm: Çözüm için önce aritmetik ortalamanın hesaplanması gerekir.

$$\bar{X} = \frac{696}{104} \cong 6{,}69$$

- Aritmetik ortalamadan farklar serisi oluşturularak standart sapma elde edilir.

$$\sigma = \sqrt{\frac{\sum f_i (X_i - \bar{X})^2}{\sum f_i}} = \sqrt{\frac{296{,}15}{104}} \Rightarrow \sigma \cong 1{,}687 \quad \text{varyans}: \ \sigma^2 = 2{,}847$$

- **Örnek:** Bir liseden mezun olan ve ÖSS sınavına giren öğrencilerin puanlarının dağılımı aşağıda verilmiştir. Buna göre öğrenci puanlarının standart sapmasını bulunuz.

ÖSS Puanları	Öğr.Sayısı	m_i	$f_i m_i$	$m_i - \bar{X}$	$f_i(m_i - \bar{X})^2$
90-110	10	100	1000	-37,9	14364,1
110-130	30	120	3600	-17,9	9612,3
130-150	50	140	7000	2,1	220,5
150-170	25	160	4000	22,1	12210,25
170-210	5	190	950	52,1	13572,05
Toplam	120		16550		49979,2

$$\text{Ortalama}: \quad \bar{X} = \frac{16550}{120} \cong 137{,}9 \ \ puan$$

$$\text{Standart sapma} \quad \sigma = \sqrt{\frac{\sum f_i (m_i - \bar{X})^2}{\sum f_i}} = \sqrt{\frac{49979{,}2}{120}} \Rightarrow \sigma \cong 20{,}4 \ \ puan$$

Standart sapmanın kısa yoldan hesaplanması

- $\sigma^2 = \dfrac{\sum (X_i - \bar{X})^2}{N}$ ifadesi açılarak yazılırsa;
- $\sigma^2 = \dfrac{\sum (X_i^2 - 2\bar{X}.X_i + \bar{X}^2)}{N}$ şeklinde yazılabilir.
- Bu ifade ayırarak yazılırsa $\sigma^2 = \dfrac{\sum X_i^2}{N} - 2.\bar{X}.\dfrac{\sum X_i}{N} + \dfrac{N\bar{X}^2}{N}$ olur.
- $\dfrac{\sum X_i^2}{N} = K^2$, $\dfrac{\sum X_i}{N} = \bar{X}$ olduğuna göre;
- $\sigma^2 = K^2 - 2\bar{X}.\bar{X} + \bar{X}^2 = K^2 - 2\bar{X}^2 + \bar{X}^2$ olur. Buna göre;
- Standart sapma kısa yolda $\boxed{\sigma = \sqrt{K^2 - \bar{X}^2}}$ şeklinde yazılır.
- Şu halde varyans kareli ortalamanın karesinden aritmetik ortalamanın karesinin farkına eşit olup, bunun kare kökü standart sapmaya eşit olur.

Örnek: Yukarıdaki ÖSS örneği için standart sapmayı kısa yoldan hesaplayınız.

- Aritmetik ortalama

$$\overline{X} = \frac{22000}{150} = 146,67 \ \ puan$$

- Kareli ortalama

$$K^2 = \frac{3328000}{150} = 22186,67$$

- Standart sapma

$$\sigma = \sqrt{K^2 - \overline{X}^2}$$

$$\sigma = \sqrt{22186,67 - 146,67^2}$$

$$\sigma = \sqrt{674,58} \Rightarrow \sigma = 25,97 \text{ puan}$$

ÖSS Puanları	Öğr.Sayısı	m_i	$f_i m_i$	$f_i m_i^2$
90-110	10	100	1000	100000
110-130	30	120	3600	432000
130-150	50	140	7000	980000
150-170	30	160	4800	768000
170-190	20	180	3600	648000
190-210	10	200	2000	400000
Toplam	150		22000	3328000

SHEPPARD DÜZELTMESİ

Tanım: Gruplanmış serilerde bir sınıfa düşen gözlem değerleri sınıfın orta değerine eşitmiş gibi kabul edilerek işlemler yapılması bir hataya sebep olmaktadır. Bu hatayı düzeltmek amacı ile Sheppard şu düzeltme formülünü geliştirmiştir.

$$\sigma^{\prime} = \sqrt{\sigma^{2} - \frac{s^{2}}{12}}$$

Buna göre $\sigma^{\prime} \leq \sigma$ olacaktır.

$\sigma^{\prime}$: Düzeltilmiş standart sapma
σ^{2}: Düzeltilmemiş varyans
s : Sınıf aralığı

Örnek: Önceki slayttaki örnek için elde edilen standart sapmaya Sheppard düzeltmesini uygulayınız .

Standart sapmanın özellikleri

- Matematik işlemler için uygun bir dağılma ölçüsüdür. Bu sebeple en yaygın kullanılan ölçüdür.
- Standart sapmada aritmetik ortalama gibi istatistik analiz için temel ölçülerden birisidir.
- Genel olarak standart sapma ortalama sapmadan daha büyüktür. $(OS < \sigma)$
- N_1 ve N_2 gözlemden oluşan iki serinin ortalamaları aynı ve sırayla varyansları σ_1^2 ve σ_2^2 olsun. Bu iki serinin birleştirilmiş ortak varyansı .

$$\sigma^2 = \frac{N_1.\sigma_1^2 + N_2.\sigma_2^2}{N_1 + N_2}$$ şeklinde olur.

Nispi Sapma Ölçüleri

- Mutlak dağılma ölçüleri gözlem değerlerinin ifade edildiği birimler cinsinden sonuç vermekte idi. Mutlak dağılma ölçülerinin bu özelliği farklı birimlerle ifade edilen serilerin değişkenliklerini karşılaştırma imkanı vermemektedir. Diğer taraftan aynı birimle ifade edilen serilerde bile gözlem değerleri arttıkça mutlak dağılma ölçüleri de buna paralel olarak artmaktadır. Bu durumda aynı cins serilerin dağılımlarını da mutlak sapma ölçüleri ile karşılaştırmak çoğu zaman mümkün olamamaktadır.

- Nispi dağılma ölçüleri serideki gözlem değerlerinin ölçüldüğü birim farklılıklarını ortadan kaldırmakta ve değişkenliği yüzde(%) cinsinden ifade etmektedir. Böylece nisbi dağılma ölçüleri farklı birimlerle ifade edilen ve farklı büyüklüklerdeki serileri aynı cins ve büyüklükte ifade etme imkanı tanımaktadır. Nispi sapma ölçüleri bu özellikleri dolayısıyla farklı birimlerle ölçülmüş, farklı büyüklük ve özelliklerdeki verilerin sapmalarının karşılaştırılmasına imkan sağlamaktadır

Değişim Katsayısı

- Standart sapmanın ortalamanın bir yüzdesi olarak ifade edilmesine değişim katsayısı adı verilir. Bu tanıma göre standart sapmanın büyüklüğü aritmetik ortalamaya göre ifade edilmektedir.

$$D.K = \frac{\sigma}{\overline{\overline{X}}}.100$$

- Bu ölçü farklı cins ve büyüklüklerdeki serileri aynı cins ve büyüklükte (yüzde cinsinden) ifade etme imkanı sağlamaktadır. Ancak bu ölçünün bir dezavantajı bir üst sınırının olmamasıdır. Yani değişim katsayısı %100 ü geçen değerler de alabilmesi bu ölçünün zayıf tarafıdır. Eğer bu ölçünün üst sınırı %100 olsaydı verinin değişkenliğini daha iyi yorumlamak mümkün olurdu. Özellikle ortalaması sıfıra yakın seriler için kullanımı pek uygun değildir.

- **Örnek:**

x	y
10	43
11	48
13	58
14	63
17	73
65	**285**

$\sigma_x = ?$ $\sigma_y = ?$

- **Örnek:** Konutlarda tüketilen aylık elektrik ve su miktarları için aşağıdaki veriler elde edilmiştir. Değişim katsayılarını bularak hangi grupta değişikliğin daha fazla olduğunu araştırın.

Elekt. Tük.(kw/h)	Konut Say.(f_i)	m_i	$f_i m_i$	$f_i m_i^2$
50-100	10	75	750	56250
100-150	20	125	2500	312500
150-200	30	175	5250	918750
200-300	15	250	3750	937500
300-500	5	400	200	800000

- Elektrik Tüketimi İçin:

$$\bar{X} = \frac{14250}{80} = 178{,}125 \qquad K^2 = \frac{3025000}{80} = 37812{,}5$$

$$\sigma = \sqrt{K^2 - \bar{X}^2} = \sqrt{37812{,}5 - 178{,}125^2} = \sqrt{6083{,}98} \cong 78$$

- Değişim katsayısı: $D.K = \frac{\sigma}{\bar{X}} \cdot 100 \Rightarrow D.K = \frac{78}{178{,}125} \cdot 100$

$$DK = \%44{,}8$$

Su Tük.(ton/h)	Konut Say.	m_i	$f_i m_i$	$f_i m_i^2$
5-15	10	10	100	1000
15-25	30	20	600	12000
25-35	40	30	1200	36000
35-45	20	40	800	32000
45-65	10	55	550	30250

- Su Tüketimi İçin değişim katsayısı;

$$\bar{X} = \frac{3250}{110} = 29{,}55 \qquad K^2 = \frac{111250}{110} = 1011{,}4$$

$$\sigma = \sqrt{K^2 - \bar{X}^2} = \sqrt{1011{,}4 - 29{,}55^2} = \sqrt{138{,}16} \cong 11{,}75$$

$$D.K = \frac{\sigma}{\bar{X}}.100 \Rightarrow D.K = \frac{11{,}75}{29{,}55}.100 \qquad \boxed{DK = \%39{,}76}$$

- Bu verilere göre elektrik tüketiminin değişkenliği (*DK*=44,8) su tüketiminin değişkenliğine göre (*DK*=39.7) daha fazladır.

Notlar	Öğr. Sayısı
0 - 20	2
20 - 40	5
40 - 60	10
60 - 80	3
80 - 100	2

Yukarıda verilen seri için şu ana kadar öğrendiğiniz kuralları uygulayınız.