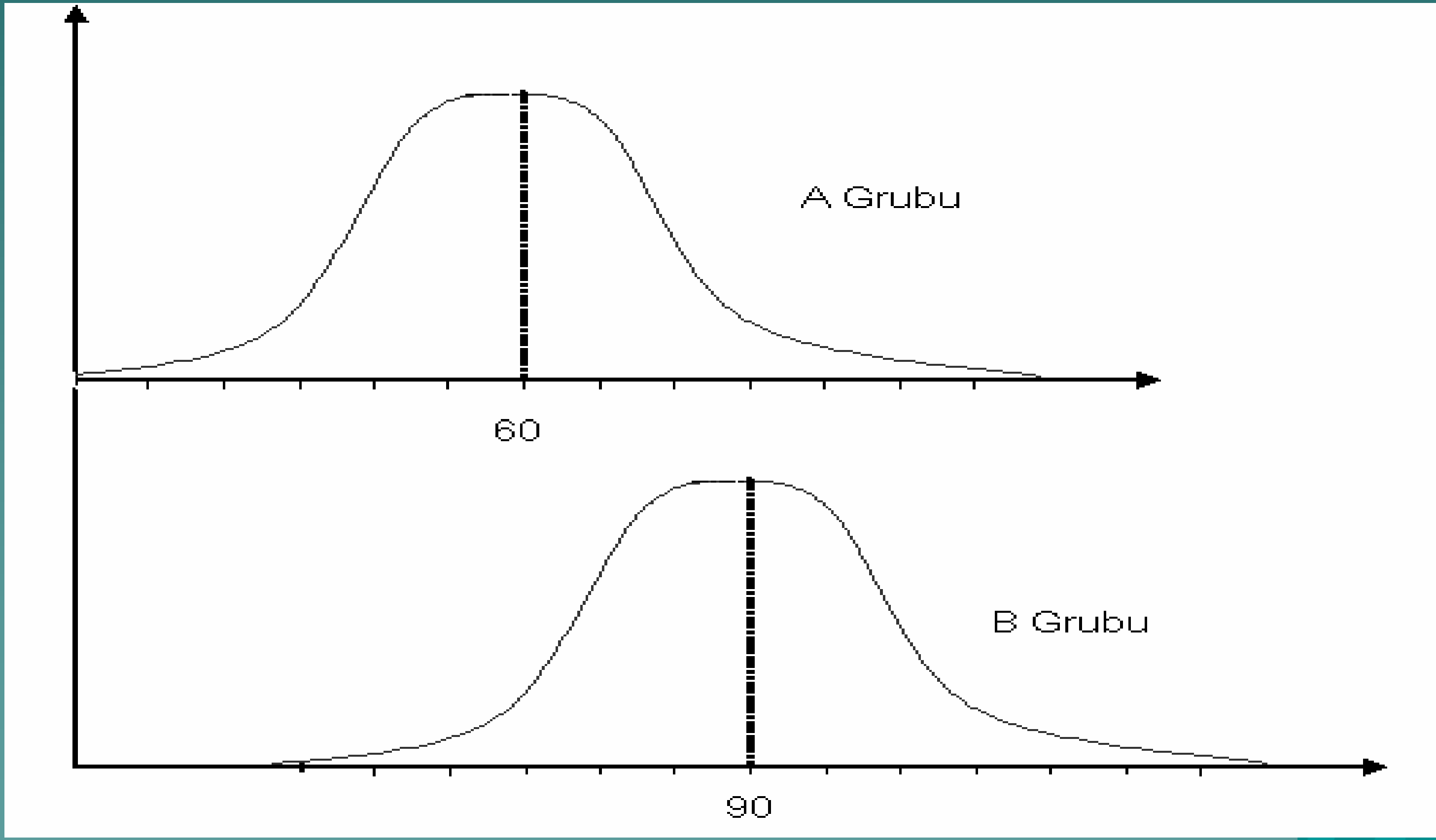
BÖLÜM 3

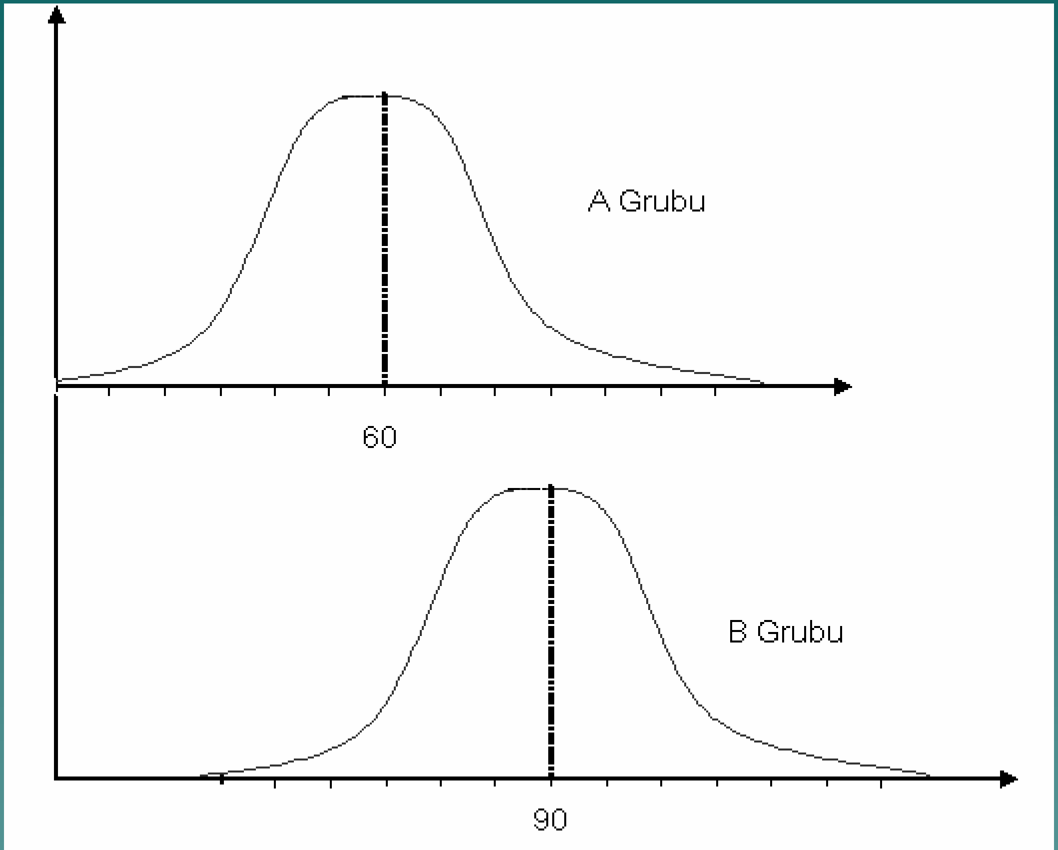
YER ÖLÇÜLERİ

Doç.Dr. Suat ŞAHİNLER

İkinci bölümde verilerin frekans tablolarının hazırlanması ve grafiklerin çizilmesindeki esas amaç; gözlemlerin doğal olarak ait oldukları populasyon dağılışını belirlemek ve dağılışın genel özelliklerini ortaya koymaktı.

Şimdi ise A ve B gibi iki gruba ait frekans poligonlarının aşağıdaki şekilde olduğunu varsayalım .





Şekilden inceleneceği üzere her iki gruba ait frekans poligonlarının benzer şekle sahiptir. Buna karşın, B grubunun frekans poligonu A'nın sağında bulunmaktadır. Ayrıca A grubunun ortalaması 60, B'nin ise 90'dır. Benzer dağılıma sahip her iki grup arasındaki farklılığı, ortalamalar yani 60 ile 90'ı karşılaştırmak suretiyle ortaya koymak mümkündür. Diğer bir ifadeyle, poligon A ile poligon B'yi mukayese etmek yerine poligonları temsil eden ortalamaları mukayese edersek, dolayısıyla poligonları mukayese etmiş oluruz. Çünkü her iki grubun gösterdikleri dağılışın şekli birbirinin benzeridir.

Yukarıdaki açıklamalardan anlaşılacağı üzere gözlemlerin en çok toplandığı kısım ele alınan ölçü bakımından dağılışın merkezini oluşturur, bu da ortalamadır.Benzer şekilde poligon (dağılışın) sağında ve solunda uzanan kuyruklar da gözlemler arasındaki varyasyonu (farklılığı) ifade eder.

Bir dağılışta gözlemlerin en çok toplandığı noktayı kantitatif bir değer olarak belirleyen ölçülere Yer Ölçüleri denir. Bu bakımdan yer ölçüleri gözlemlerin çoğunluğunu temsil ederler. Yer ölçüleri aynı zamanda dağılışın merkezini veya X ekseni üzerindeki yerini, bir diğer ifade ile dağılışın merkezinin orijine olan uzaklığını belirlerler.

En çok bilinen ve kullanılan yer ölçüleri

- Aritmetik ortalama,
- Tartılı ortalama,
- Medyan (orta değer),
- Mod (tepe değeri),
- Geometrik ortalama,
- Harmonik ortalama,
- Kareli ortalama ve
- Kantiller (bölenler)

olarak sıralanabilir. Bu bölümde bunlar arasından ilk dördü incelenecektir.

Aritmetik Ortalama

Aritmetik ortalama istatistikte çok kullanılması nedeniyle çok iyi bilinmesi gereken bir ortalama veya bir yer ölçüsüdür. Bu ortalama hem ham verilerden hem de sınıflandırılmış verilerden kolayca hesaplanabilir.

Aritmetik ortalama, herhangi bir örneği meydana getiren gözlem değerlerinin toplamının, toplam gözlem sayısına bölünmesiyle elde edilen değer olarak tanımlanabilir.

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

Burada;
$\bar{x}$: aritmetik ortalamayı
x_i : i nci gözlem değerini,
n : gözlem sayısını
göstermektedir.

Örnek : Beş lise öğrencisinin matematik dersinden aldıkları notlar sırasıyla 7,6,5,8,4 puan olduğuna göre öğrencilerin notlarının aritmetik ortalamasını bulunuz.

Çözüm: Aritmetik ortalama formülüne göre;

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{x_1 + x_2 + x_3 + x_4 + x_5}{5} = \frac{7+6+5+8+4}{5} = \frac{30}{5} = 6$$

olarak hesaplanır.

Örnekte de görüldüğü üzere aritmetik ortalamanın hesaplama tekniği çok basittir. Basit bir matematiksel formülle ifade edilmesi ve kolayca uygulanabilmesi yanında teorik bakımdan gösterdiği çeşitli özellikleri nedeniyle bu ortalamanın geniş bir uygulama alanı vardır.

Sınıflandırılmış Verilerde Aritmetik Ortalamanın Hesaplanması

Verilerin sınıflandırılması amacıyla düzenlenen frekans tabloları bazı bilgilerin kaybına sebep olmaktadır. Bu nedenle aritmetik ortalamanın gerçek değerinin bulunması oldukça güçtür. Bununla beraber, frekans tablolarında sınıf değerlerini veya sınıf orta değerlerini gözlenen değer gibi dikkate alarak aritmetik ortalama hesaplanabilir. Bu amaçla

$$\bar{x} = \frac{\sum_{j=1}^{k} f_j x_j}{\sum_{j=1}^{k} f_j} = \frac{f_1 x_1 + f_2 x_2 + \ldots + f_n x_n}{f_1 + f_2 + \ldots + f_n}$$

Burada;
$\bar{x}$: aritmetik ortalamayı
x_j : j nci sınıfın sınıf değerini
f_j : j nci sınıfın frekansını
k : sınıf sayısını
göstermektedir.

Örnek: Kırk koyunun ağırlıkları ile ilgili frekans tablosu için aritmetik ortalamayı hesaplayınız.

Sınıf Limitleri (Koy. Ağr.(kg))		f (Koy.Say)
39	42	1
43	46	4
47	50	8
51	54	11
55	58	7
59	62	5
63	66	2
67	70	2
Toplam		40

Önce sınıf değerleri bulunur ve formül gereği frekanslarla çarpılır ve toplanırsa…

Sınıf Limitleri (Koy. Ağr.(kg))		f_j (Koy.Say)	Sınıf Değeri (x_j)	$f_j x_j$
39	42	1	40.5	40.5
43	46	4	44.5	178.0
47	50	8	48.5	388.0
51	54	11	52.5	577.5
55	58	7	56.5	395.5
59	62	5	60.5	302.5
63	66	2	64.5	129
67	70	2	68.5	137
Σ		40	-	2148

$$\overline{x} = \frac{\sum_{j=1}^{k} f_j x_j}{\sum_{j=1}^{k} f_j} = \frac{1*40.5 + 4*44.5 + ... + 2*68.5}{1+4+...+2} = \frac{2148}{40} = 53.7\,kg$$

olarak hesaplanır.

Örnek 3. 3: Bir koyun sürüsünde 2 yaşlı anaçlardan doğan dişi kuzuların doğum ağırlıklarına(kg) ait aşağıdaki frekans tablosu için göre aritmetik ortalamayı bulunuz(Düz güneş ve ark.,1983).

Doğum Ağırlığı (kg)	Kuzu sayısı (f)	Sınıf Değeri (x)	fx
2.5-2.7	3	2.6	7.8
2.8-3.0	6	2.9	17.4
3.1-3.3	6	3.2	19.2
3.4-3.6	8	3.5	28.0
3.7-3.9	7	3.8	26.6
4.0-4.2	6	4.1	24.6
4.3-4.5	2	4.4	8.8
4.6-4.8	2	4.4	9.4
Toplam	40	-	141.8

$$\overline{x} = \frac{\sum_{j=1}^{k} f_j x_j}{\sum_{j=1}^{k} f_j} = \frac{141.8}{40} = 3.545kg$$

Aritmetik ortalama örnekteki aşırı değerlerden çok etkilenir . Aritmetik ortalamanın bu zayıf yönüne karşın aşağıda verilmiş olan bazı matematik özellikleri aritmetik ortalamayı istatistik analiz açısından önemli kılar .

a)Gözlemlerin aritmetik ortalamadan farklarının (sapmalarının) cebirsel toplamı sıfıra eşittir .Yani,

$$\sum_{i=1}^{n}(x_i-\overline{x})=\sum_{i=1}^{n}x_i-\sum_{i=1}^{n}\overline{x}=n\overline{x}-n\overline{x}=0$$

$$\overline{x}=\frac{\sum_{i=1}^{n}x_i}{n}\Rightarrow\sum_{i=1}^{n}x_i=n\overline{x}$$

Yukarıdaki beş öğrencinin notları ile ilgili veri için bu durumu gösterelim.

$$\sum_{i=1}^{n}(x_i-\overline{x})=(7-6)+(6-6)+(5-6)+(8-6)+(4-6)$$

$$=1+0-1+2-2=0$$

b) Gözlemlerin, aritmetik ortalamadan sapmalarını kareleri toplam minimumdur. Diğer bir deyişle, gözlemlerin ortalamadan farklı bir değerden sapmalarının kareleri toplamı, ortalamadan sapmalarının kareleri toplamından daha büyüktür.

$$\sum_{i=1}^{n}(x_i-\overline{x})^2<\sum_{i=1}^{n}(x_i-A)^2, A\neq\overline{x}$$

Yukarıdaki beş öğrencinin notları ile ilgili veri için bu durumu gösterelim.

$$\sum_{i=1}^{n}(x_i-\overline{x})^2=(7-6)^2+(6-6)^2+(5-6)^2+(8-6)^2+(4-6)^2$$

$$=1^2+0^2-1^2+2^2-2^2=10$$

$A=4\neq 6=\overline{x}$ ise

$$\sum_{i=1}^{n}(x_i-A)^2=\sum_{i=1}^{5}(x_i-4)^2=(7-4)^2+(6-4)^2+(5-4)^2+(8-4)^2+(4-4)^2$$

$$=3^2+2^2+1^2+4^2-0^2=30$$

Böylece

$$\sum_{i=1}^{n}(x_i-\overline{x})^2<\sum_{i=1}^{n}(x_i-A)^2$$

10<30

Olur.

Tartılı Ortalama:

Gözlemlerin temsil ettikleri değer bakımından farklılık gösterdikleri durumlarda kullanılan bir yer ölçüsüdür.

$$\bar{x}_t = \frac{\sum_{j=1}^{k} t_j x_j}{\sum_{j=1}^{k} t_j} = \frac{t_1 x_1 + t_2 x_2 + \ldots + t_n x_n}{t_1 + t_2 + \ldots + t_n}$$

Tartılı ortalama, aynı örnek içinde farklılık gösteren gözlemlerin ortalamasını bulmak için bunların temsil ettiği birim sayılarına göre tartılması gerekir

Örnek: Bir öğrencinin değişik derslerine ait aşağıda verilen notlarının tartılı ortalamasını hesaplayınız.

Dersler	Haftalık Kredi saati	Aldığı Not
Türk Dili	2	100
Gıda kimyası	3	85
İstatistik	4	65
Toplam	9	250

Dersler	Haftalık Kredi saati (t)	Aldığı Not (x)	tx
Türk Dili	2	100	200
Gıda Kimyası	3	85	255
İstatistik	4	65	260
Toplam	9	250	715

$$\bar{x}_t = \frac{\sum_{j=1}^{k} t_j x_j}{\sum_{j=1}^{k} t_j} = \frac{715}{9} = 79.44$$

Dersler	Haftalık Kredi saati (t)	Aldığı Not (x)	tx
Türk Dili	2	100	200
Gıda Kimyası	3	85	255
İstatistik	4	65	260
Toplam	9	250	715

Bu tabloya ait aritmetik ortalama hesaplanırsa

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{250}{3} = 83.33$$

Derslerdeki başarı önceki durumdan farklı olsa idi aritmetik ortalama değişmeyecekti ancak tartılı ortalama ise

Dersler	Haftalık Kredi saati (t)	Aldığı Not (x)	tx
Türk Dili	2	65	130
Gıda Kimyası	3	85	255
İstatistik	4	100	400
Toplam	9	250	785

$$\bar{x}_t = \frac{\sum_{j=1}^{k} t_j x_j}{\sum_{j=1}^{k} t_j} = \frac{785}{9} = 87.22$$

olarak hesaplanacaktır.

Medyan (Orta Değer):

Medyan, küçükten büyüğe doğru sıralanmış gözlemlerde ortaya düşen değer olarak tanımlanabilir. Başka bir deyişle medyan, örnekteki gözlemleri iki eşit kısma ayıran değer olup gözlemlerin %50'si bu değerden küçük, %50'si ise bu değerlerden büyüktür.

a) Eğer n tek ise,

$$med=x_{((n+1)/2)}$$

Örnek:

X={6,18,12,62,15,10,17} için medyanı bulunuz.

Çözüm:

Öncelikle bu veriler büyüklük sırasına göre dizilir,

X={6,10,12,15,17,18,62}, n=7 olduğuna göre

$$med=x_{((n+1)/2)}=x_{((7+1)/2)}=x_{(4)}=15$$

Eğer n çift ise ,

b) Eğer n çift ise ,

$$med = \frac{x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}}{2}$$

Örnek:
X={6,18,12,62,15,10,17,40} için medyanı bulunuz.

Çözüm:
Öncelikle bu veriler büyüklük sırasına göre dizilir,
X={6,10,12,15,17,18,40,62}, n=8 olduğuna göre

$$med = \frac{x_{(\frac{8}{2})} + x_{(\frac{8}{2}+1)}}{2} = \frac{x_{(4)} + x_{(5)}}{2} = \frac{15+17}{2} = 16$$

Sınıflandırılmış Verilerde Medyan

$$med = b_s + \frac{(\frac{n}{2} - F_{m-1})}{f_m} * c$$

Burada;
b_s:medyanın içinde bulunduğu sınıfın alt sınıf sınırını,
n: toplam gözlem sayısını,
F_{m-1}: medyanın içinde bulunduğu sınıfdan bir önceki sınıfın “den daha az” yığmalı frekansını,
fm:medyanın içinde bulunduğu sınıfın frekansını,
c:sınıf aralığını

göstermektedir.

Örnek: 40 koyunun ağırlığı ile ilgili frekans tablosuna ait medyanı hesaplayınız

Sınıf Limitleri (Koy. Ağr.(kg))		f_j (Koy.Say)
39	42	1
43	46	4
47	50	8
51	54	11
55	58	7
59	62	5
63	66	2
67	70	2

Sınıf sınırları bulunur ve sıralanır

Sınıf Limitleri (Koy. Ağr.(kg))		f_j (Koy.Say)	Sınıf Sınırları (Koy. Ağr.(kg))		Sınıf Sınırları (Koy. Ağr.(kg))
					38.5
39	42	1	38.5	42.5	42.5
43	46	4	42.5	46.5	46.5
47	50	8	46.5	50.5	50.5
51	54	11	50.5	54.5	54.5
55	58	7	54.5	58.5	58.5
59	62	5	58.5	62.5	62.5
63	66	2	62.5	66.5	66.5
67	70	2	66.5	70.5	70.5

Den daha az yığmalı frekanslar bulunur

Sınıf Limitleri (Koy. Ağr.(kg))		f_j (Koy.Say)	Sınıf Sınırları (Koy. Ağr.(kg))		Sınıf Sınırları (Koy. Ağr.(kg))	F
					38.5	0
39	42	1	38.5	42.5	42.5	1
43	46	4	42.5	46.5	46.5	5
47	50	8	46.5	50.5	50.5	13
51	54	11	50.5	54.5	54.5	24
55	58	7	54.5	58.5	58.5	31
59	62	5	58.5	62.5	62.5	36
63	66	2	62.5	66.5	66.5	38
67	70	2	66.5	70.5	70.5	40

Medyanın tanımından çıkılarak medyan sınıfı belirlenir

Sınıf Limitleri (Koy. Ağr.(kg))		f_j (Koy.Say)	Sınıf Sınırları (Koy. Ağr.(kg))		Sınıf Sınırları (Koy. Ağr.(kg))	F
					38.5	0
39	42	1	38.5	42.5	42.5	1
43	46	4	42.5	46.5	46.5	5
47	50	8	46.5	50.5	50.5	13
51	54	11	50.5	54.5	54.5	24
55	58	7	54.5	58.5	58.5	31
59	62	5	58.5	62.5	62.5	36
63	66	2	62.5	66.5	66.5	38
67	70	2	66.5	70.5	70.5	40

Medyanın tanımından çıkılarak medyan sınıfı belirlenir

Sınıf Limitleri (Koy. Ağr.(kg))		f_j (Koy.Say)	Sınıf Sınırları (Koy. Ağr.(kg))	F
			38.5	0
39	42	1	42.5	1
43	46	4	46.5	5
47	50	8	50.5	13
51	54	11	54.5	24
55	58	7	58.5	31
59	62	5	62.5	36
63	66	2	66.5	38
67	70	2	70.5	40

$$med = b_s + \frac{(\frac{n}{2} - F_{m-1})}{f_m} * c$$

$$= 50.5 + \frac{(\frac{40}{2} - 13)}{11} * 4$$

= 53.05 kg

Hangi Durumlarda medyan kullanılır?

Aritmetik ortalama,örneği oluşturan gözlemlerde aşırı değer veya değerler bulunduğu durumlarda çok etkilenir ve yanıltıcı sonuç verir. Örnekte anormal derecede düşük veya yüksek bir veya birkaç değer varsa, aritmetik ortalama beklenenden düşük veya yüksek bulunur.

Bu gibi aritmetik ortalamanın sağlıklı bir şekilde çalışmadığı durumlarda başka bir yer ölçüsü olan medyan kullanılır. Medyan, örneği meydana getiren gözlemlerde anormal değerlerden etkilenmez.

Ayrıca aritmetik ortalama, açık uçlu frekans tablolarından hesaplanamayacağından yer ölçüsü olarak medyan kullanılır.

Mod (Tepe Değeri):

Mod , herhangi bir örnekte en fazla tekrarlanan gözlem değeri olarak tanımlanabilir . İstatistikte çok az kullanılan bu değer , özellikle verilerin simetrik bir dağılış göstermedikleri durumlarda iyi bir yer ölçüsü olmaktadır . Aynı zamanda örnekteki aşırı değerlerden etkilenmez .

Sınıflandırılmamış Verilerde Mod'un Hesaplanması

Örnek: X={6,5,7,4,5,8,5,10} için modu bulunuz.
Çözüm: En fazla tekrarlanan gözlem değeri 5 olması nedeniyle mod=5 ‵ dir .

Eğer bir örnekte en fazla tekrarlanan değer yalnız bir tane ise örnek tek modlu , iki tane ise iki modlu (bimodal) , ikiden fazla olduğunda ise örnek çok modlu (multimodal) olur . Buna karşın örnekteki değerlerin hepsi birbirinden farklı ise veya gözlemler aynı sayıda tekrarlanıyorsa , o zaman mod yoktur .

Örnek: Aşağıda her bir veri grubu için modu bulunuz .

a) 6,15,7,8,15,14,15
b) 20,80,45,80,45,75,45,80,45,80
c) 20,15,8,12,5,16
d) 16,30,25,22,30,16,25,22

Çözüm: Modun tanımına göre her grubun modu şöyledir :

a)Grubun modu =15 olup 3 defa tekrarlanmıştır .
b)Grubun mod_1 =45 ve mod_2=80 olup 4 'er defa tekrarlanmışlardır .
c)Gruptaki değerlerin hepsi birbirinden farklı olduğundan veya bir defa tekrarlandığından mod yoktur .
d)Gruptaki bütün gözlemler aynı frekanslı olduklarından (eşit sayıda, yani iki defa) mod yoktur .

Sınıflandırılmış Verilerde Modun Hesaplanması

$$\mathrm{mod} = b_s + \frac{d_1}{d_1 + d_2} * c$$

Burada;
b_s:modun içinde bulunduğu sınıfın alt sınıf sınırını,
d_1: mod sınıfı ile bir önceki sınıfın frekansları arasındaki fark,
d_2: mod sınıfı ile bir sonraki sınıfın frekansları arasındaki fark,
c:sınıf aralığını

göstermektedir.

Örnek: 40 koyunun ağırlığı ile ilgili frekans tablosuna ait modu hesaplayınız

Sınıf Limitleri (Koy. Ağr.(kg))		f_j (Koy.Say)	Sınıf Sınırları (Koy. Ağr.(kg))	
39	42	1	38.5	42.5
43	46	4	42.5	46.5
47	50	8	46.5	50.5
51	54	11	50.5	54.5
55	58	7	54.5	58.5
59	62	5	58.5	62.5
63	66	2	62.5	66.5
67	70	2	66.5	70.5

Modun tanımından çıkılarak mod sınıfı kolayca belirlenir

Sınıf Limitleri (Koy. Ağr.(kg))		f_j (Koy.Say)	Sınıf Sınırları (Koy. Ağr.(kg))	
39	42	1	38.5	42.5
43	46	4	42.5	46.5
47	50	8	46.5	50.5
51	54	11	50.5	54.5
55	58	7	54.5	58.5
59	62	5	58.5	62.5
63	66	2	62.5	66.5
67	70	2	66.5	70.5

Mod formülündeki değerler yerine yazılırsa

Sınıf Limitleri (Koy. Ağr.(kg))		f_j (Koy.Say)	Sınıf Sınırları (Koy. Ağr.(kg))	
39	42	1	38.5	42.5
43	46	4	42.5	46.5
47	50	8	46.5	50.5
51	54	11	50.5	54.5
55	58	7	54.5	58.5
59	62	5	58.5	62.5
63	66	2	62.5	66.5
67	70	2	66.5	70.5

$$\text{mod} = b_s + \frac{d_1}{d_1 + d_2} * c$$

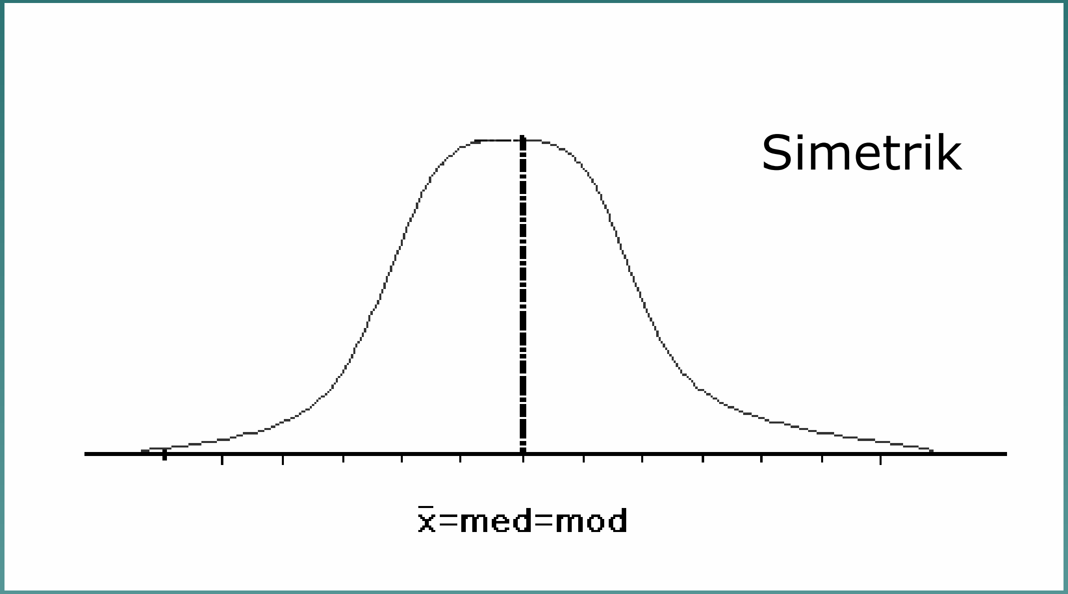
d_1=11-8=3
d_2=11-7=4

$$\text{mod} = 50.5 + \frac{3}{3+4} * 4$$

= 52.21 kg

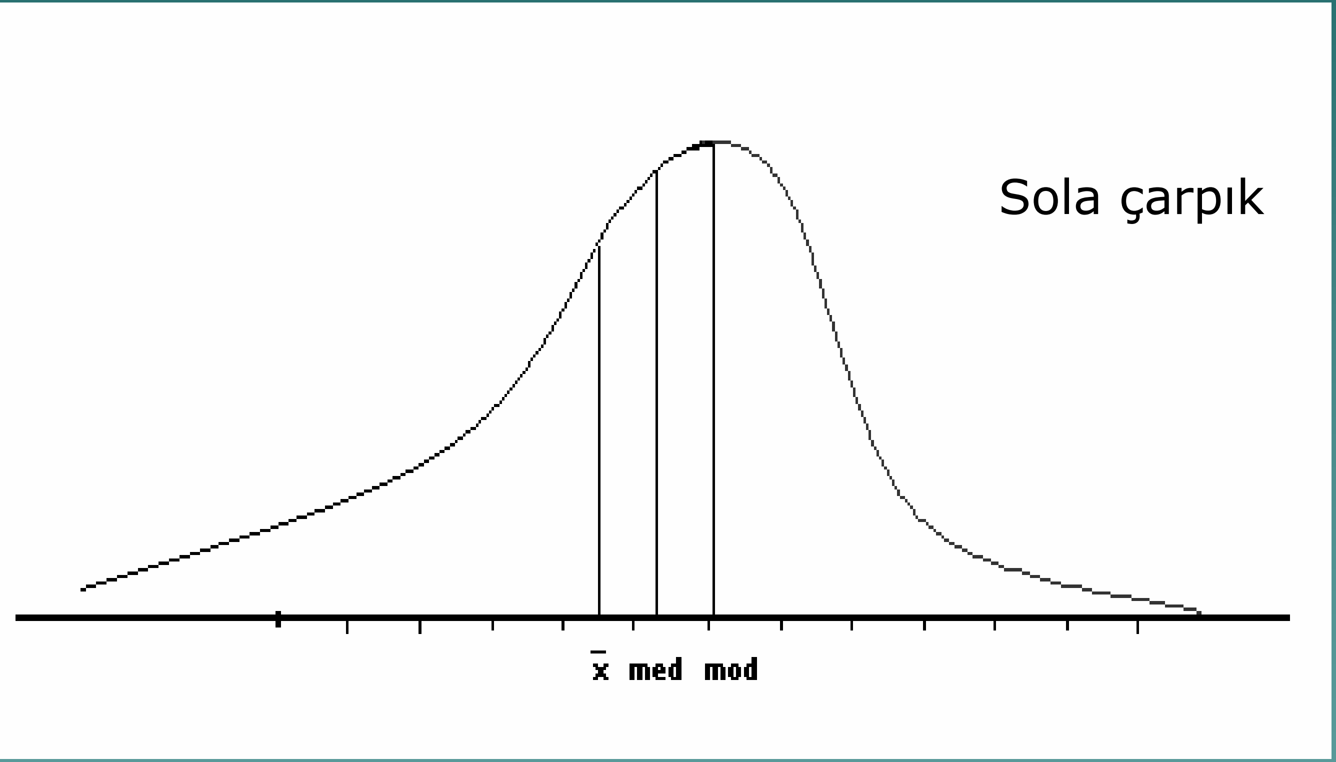
Aritmetik ortalama, medyan ve mod arasındaki ilişki

a) $\bar{x} = med = \mathrm{mod}$ ise frekans poligonunun şekli simetriktir.



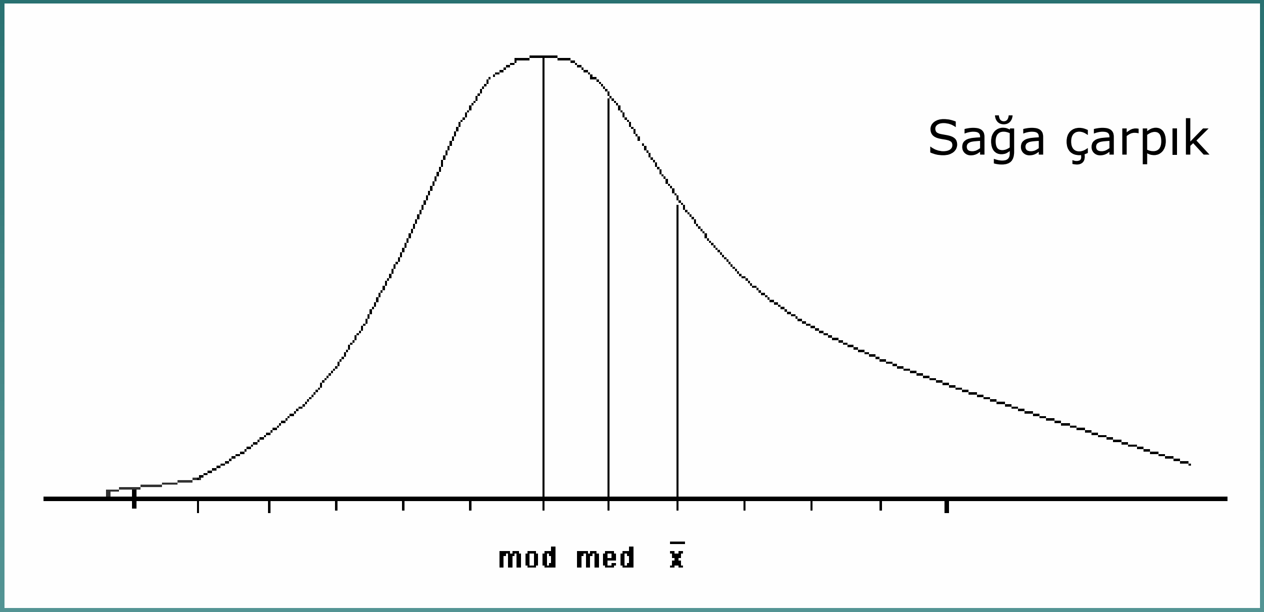
Aritmetik ortalama, medyan ve mod arasındaki ilişki

b) $\bar{x} < med < \mathrm{mod}$ ise frekans poligonunun şekli sola çarpıktır.



Aritmetik ortalama, medyan ve mod arasındaki ilişki

c) $\mathrm{mod} < med < \bar{x}$ ise frekans poligonunun şekli sola çarpıktır.



Aritmetik ortalama, medyan ve mod arasındaki bu ilişkiler biliniyorsa frekans poligonunu çizmeden dağılışın şekli hakkında fikir sahibi olunabilir.