

Bölüm 3

Çarpıklık ve Basıklık Katsayıları

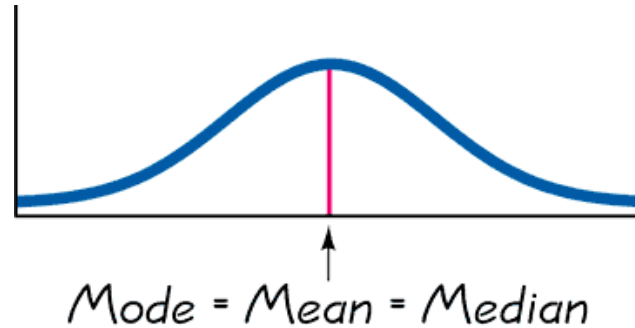
Tanımlayıcı İstatistikler

- Bir veri setini tanımak veya birden fazla veri setini karşılaştırmak için kullanılan ve ayrıca örnek verilerinden hareket ile frekans dağılımlarını sayısal olarak özetleyen değerlere **tanımlayıcı istatistikler** denir.
- Analizlerde kullanılan veri tiplerine (*basit, gruplanmış, sınıflanmış*) göre hesaplamalarda kullanılacak formüller değişmektedir.

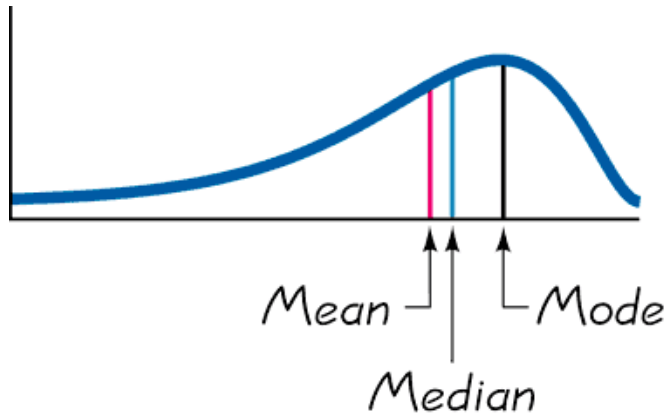
Tanımlayıcı İstatistikler



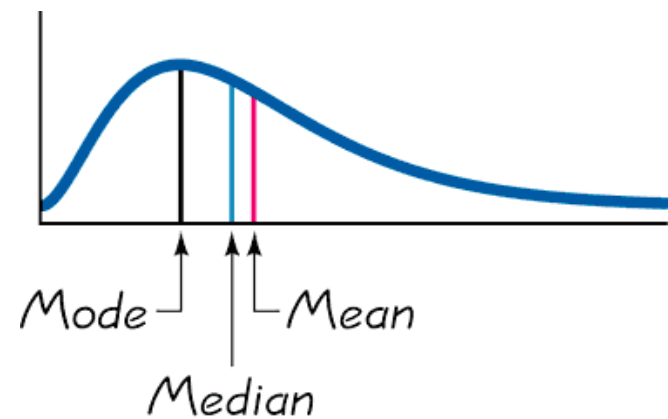
Çarpıklık



(b) Symmetric



(a) Skewed to the Left
(Negatively)



(c) Skewed to the Right
(Positively)

Çarpıklık (Asimetri) Ölçüleri

- Anakütleleri birbirinden ayırmak için her zaman yalnızca yer ve yayılım ölçüleri yeterli olmayabilir. Aşağıda iki farklı anakütleden alınmış örnekler için oluşturulan histogramlar verilmiştir.

13) Asimetri Ölçüleri

PEARSON ÇARPIKLIK ÖLÇÜSÜ

$$Sk_p = \frac{\bar{x} - \text{mod}}{s} \quad \text{veya}$$

$$Sk_p = \frac{3(\bar{X} - med)}{s}$$

$Sk_p < 0 \rightarrow$ Negatif çarpık(Sola)

$Sk_p > 0 \rightarrow$ Pozitif Çarpık(Sağa)

$Sk_p = 0$ ise dağılış simetrik

BOWLEY ÇARPIKLIK ÖLÇÜSÜ

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

$Sk_b < 0 \rightarrow$ Negatif çarpık(Sola)

$Sk_b > 0 \rightarrow$ Pozitif Çarpık(Sağa)

$Sk_b = 0$ ise dağılış simetrik

Örnek: Aşağıdaki tabloda 30 günlük süre içinde bir restoranın kullandığı et miktarının dağılımından elde edilen bazı tanımlayıcı istatistikler verilmiştir. Buna göre pearson ve bowley asimetri ölçülerini hesaplayıp yorumlayınız.

Aritmetik Ort.	Mod	Medyan	Q_1	Q_3	s^2
46,6	45,4	46,2	41,5	51,9	54,46

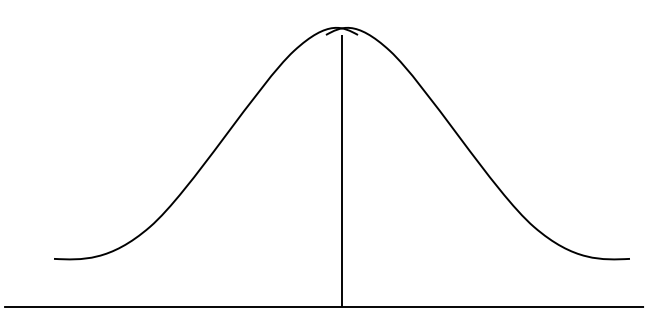
$$Sk_p = \frac{3(\bar{X} - med)}{s} = \frac{3(46,6 - 46,2)}{\sqrt{54,46}} \approx 0,16 > 0 \quad \text{Sağa Çarpık , Pozitif Asimetri}$$

$$Sk_p = \frac{\bar{x} - \text{mod}}{s} = \frac{46,6 - 45,4}{\sqrt{54,46}} \approx 0,16 > 0 \quad \text{Sağa Çarpık, Pozitif Asimetri}$$

$$Sk_b = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{Q_3 - Q_1} = \frac{(51,9 - 46,2) - (46,2 - 41,5)}{51,9 - 41,5}$$

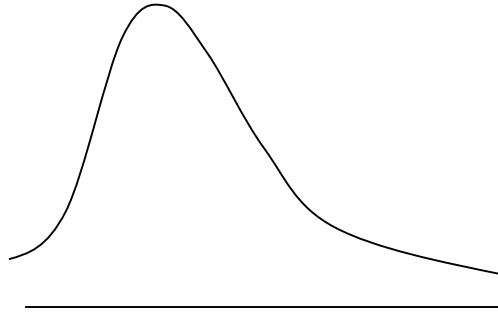
$$= \frac{1}{10,4} \approx 0,10 > 0$$

**Sağa Çarpık ,
Pozitif Asimetri**



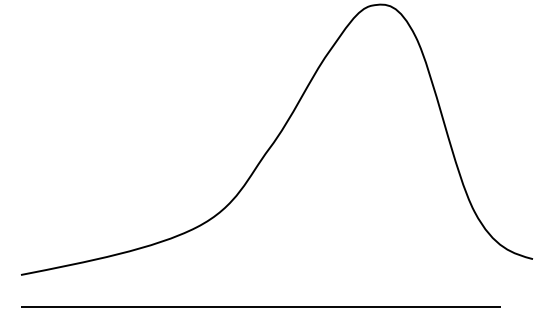
Simetrik Dağılım

$$A.O = Med = Mod$$



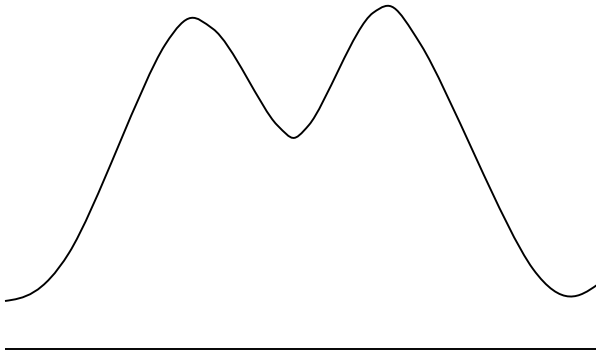
Sağa çarpık dağılım

$$A.O > Med > Mod$$

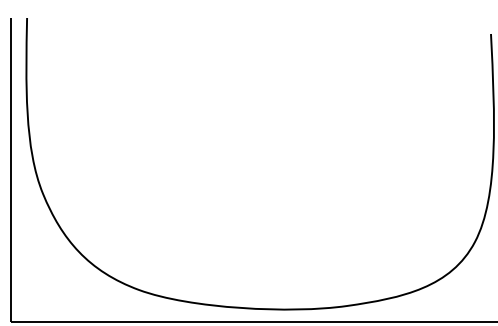


Sola çarpık dağılım

$$A.O < Med < Mod$$



İki modlu simetrik dağılım

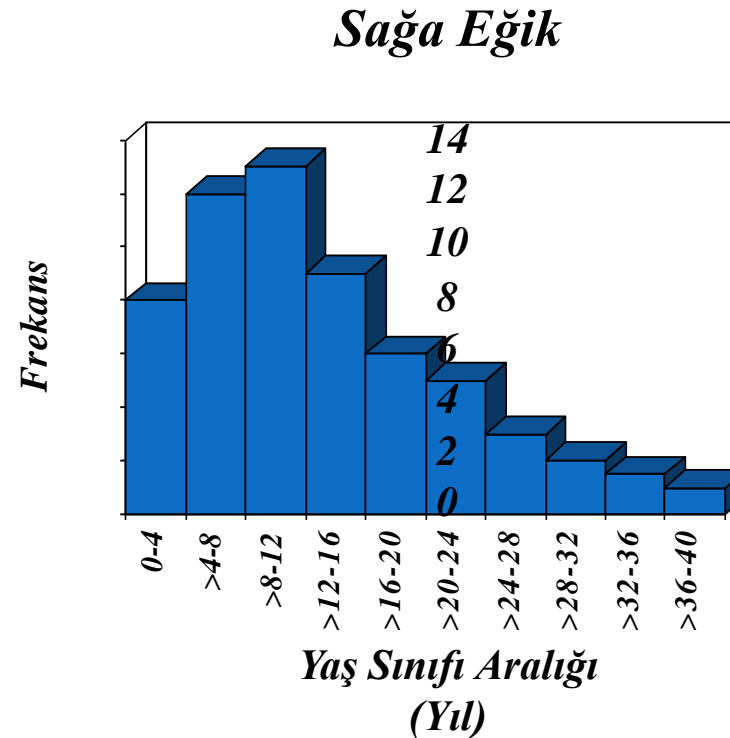


Modu olmayan dağılım



Tekdüzen dağılım

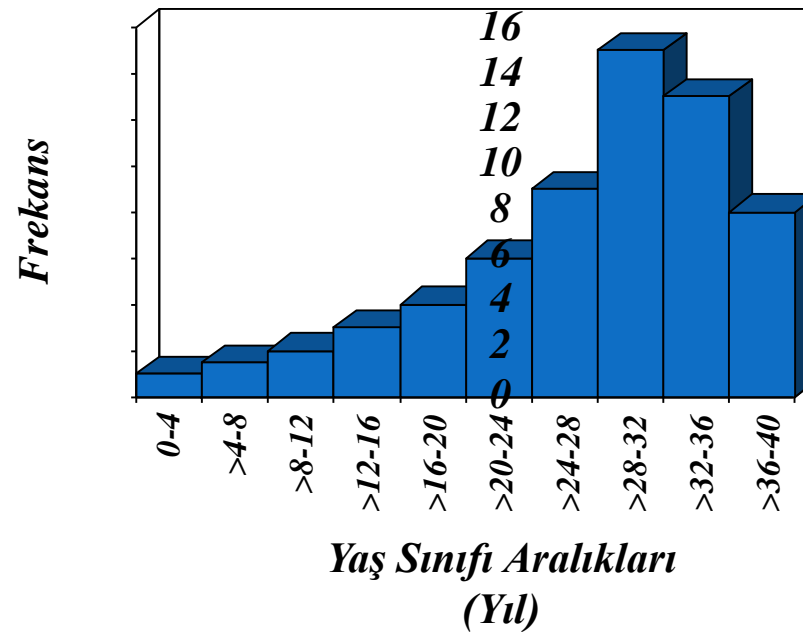
Ortalama, Medyan ve Modun karşılaştırılması



Ortalama > Medyan > Mod

Ortalama, Medyan ve Modun karşılaştırılması

Sola Eğik



Ortalama < Medyan < Mod

14) Sapan Gözlemler

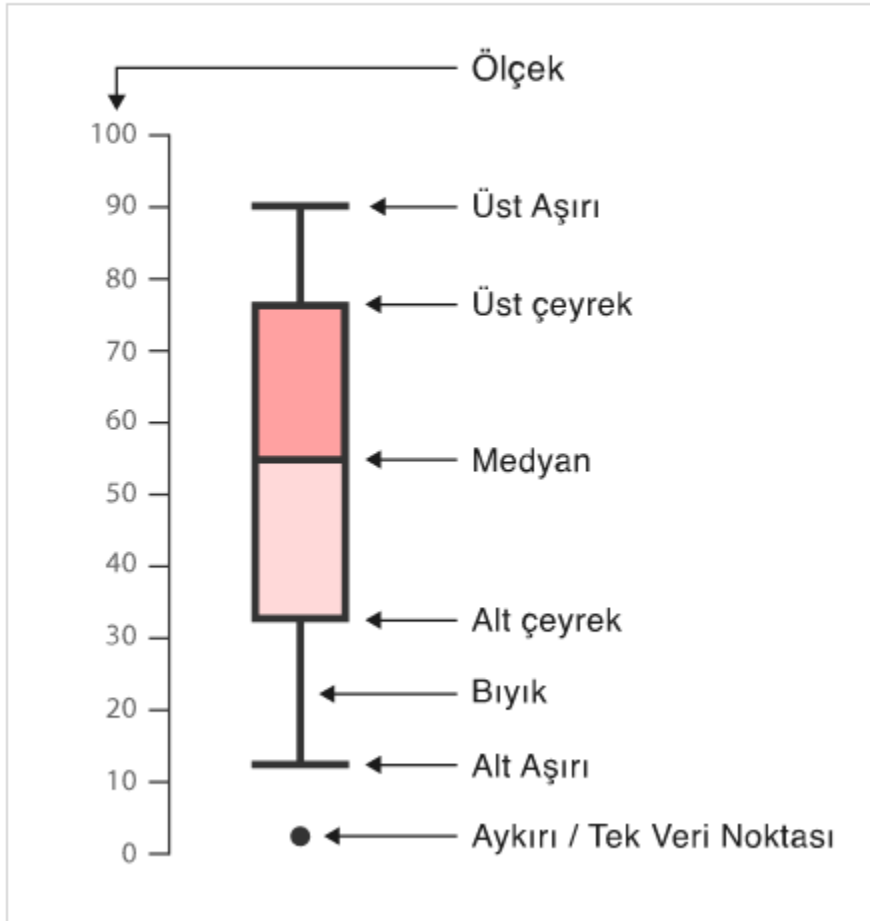
- ❖ Sapan gözlem, diğer bütün gözlemlerden uzakta bulunan gözlemdir.
- ❖ Sapan gözlem ortalama üzerinde önemli bir etkiye sahip olabilir.
- ❖ Sapan gözlem standart sapma üzerinde önemli bir etkiye sahip olabilir.
- ❖ Sapan gözlem, dağılımın gerçek histogramının ölçeği üzerinde önemli bir etkiye sahip olabilir.

15) 5 Sayı Özeti

- ❖ 5 sayı özeti, bir veri setinde minimum değer, 1.Kartil, 2.Kartil(medyan), 3.Kartil'i ve maksimum değeri içerir.
- ❖ Kutu grafiği(veya kutu ve bıyık grafiği) bir veri seti için, sınırları maksimum ve minimum değer olmak üzere, içinde 1.Kartil, 2.Kartil(medyan) ve 3.Kartil'i bulunduran kutu şeklindeki grafikdir.

Kutu Grafiği

Anatomi



Kutu grafikleri, verilerin dörtte birinin dağılımını görsel olarak göstermenin kolay bir yoludur.

Çeyreğin daha üstünde veya altında olan veriler arasındaki değişkenliği belirtmek için kullanılmaktadır.

Her ne kadar kutu grafikleri histogram grafiği ile karşılaştırıldığında daha basit görünse de daha az yer kaplama gibi bir avantajları vardır ve bu da birçok grup veya veri seti dağılımını karşılaştırırken oldukça kullanışlı bir özelliktir.

Kutu grafiği hazırlama

- Q1:Kutunun sol kenarı
- Q3:Kutunu sağ kenarı
- Q2:Kutunun ortasındaki çizgi
- Sapan hariç min.: Sol bıyık
- Sapan hariç max.: Sağ bıyık
- Sapan değer kontrolü

$$Q1 - 1.5(Q3 - Q1)$$

$$Q3 + 1.5(Q3 - Q1)$$

bu değerleri aşan veriler * ile gösterilir.

- **Örnek:**

Yazlık ürünler satan bir mağazada haftalık satılan t-shirt sayıları yandaki tabloda verilmiştir.

Verilen tablodan beş sayı özetini bulunuz ve kutu grafiğini çiziniz.

27	22	20
17	18	18
22	21	29
20	32	17
30	19	28
25	20	31
22	23	21
28	22	24
18	18	32
25	18	44
		17

- **Çözüm:**

Öncelikle veriler yandaki gibi sıralanırsa;

$Q_1 = (31+1)/4 = 8$.sıraya karşılık gelen veri olur.

$Q_1 = 18$

$Q_3 = 3(31+1)/4 = 24$. sıraya karşılık gelen veri olur.

$Q_3 = 28$

Minimum değer=17,

Maksimum değer=44 ve

Medyan(Q_2)=22 olur.

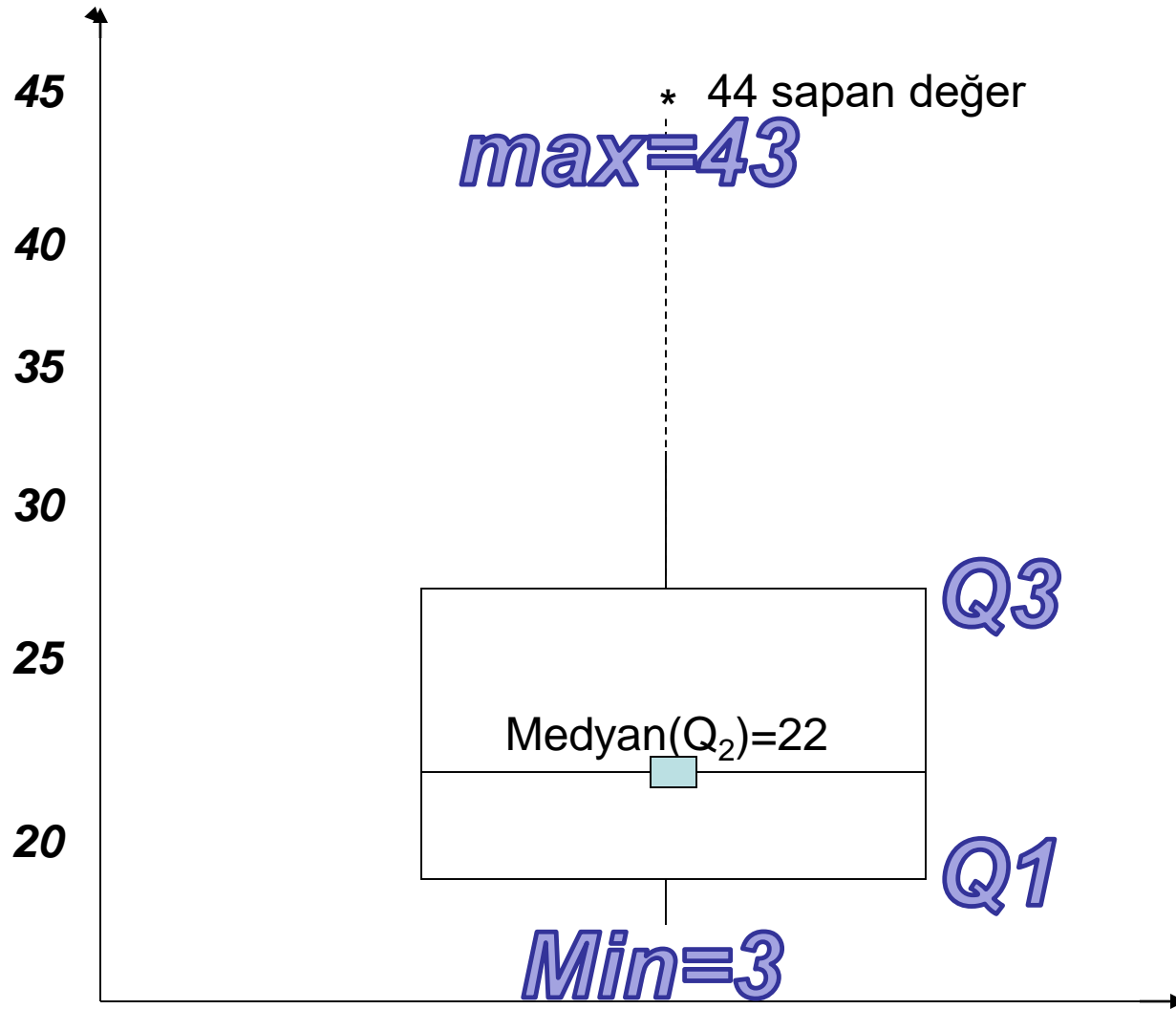
Sapan değerleri kontrol etmek için;

$Q_1 - 1,5(Q_3 - Q_1) = 18 - 1,5(28 - 18) = 3$

$Q_3 + 1,5(Q_3 - Q_1) = 28 + 1,5(28 - 18) = 43$

bulunur. Bu durumda elimizdeki 44 değeri sapan değerdir ve * ile gösterilir..

17	20	25
17	20	25
17	21	27
18	21	28
18	22	28
18	22	29
18	22	30
18	22	31
19	23	32
20	24	32
		44



Kutu Grafiği

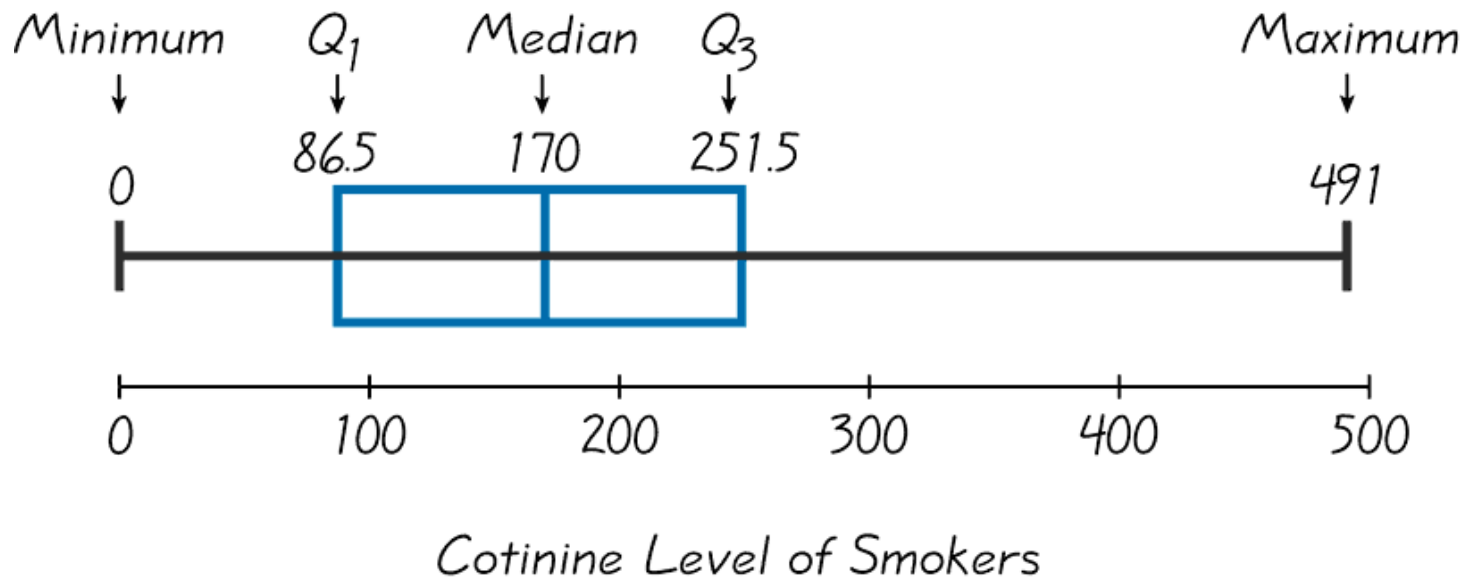
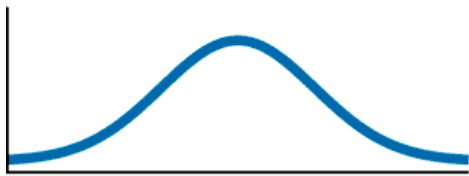


Figure
2-16

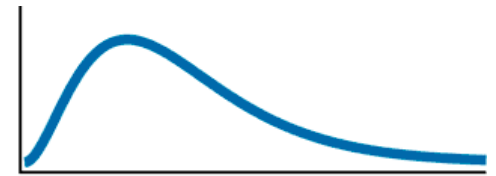
Kutu Grafiği



Bell-shaped

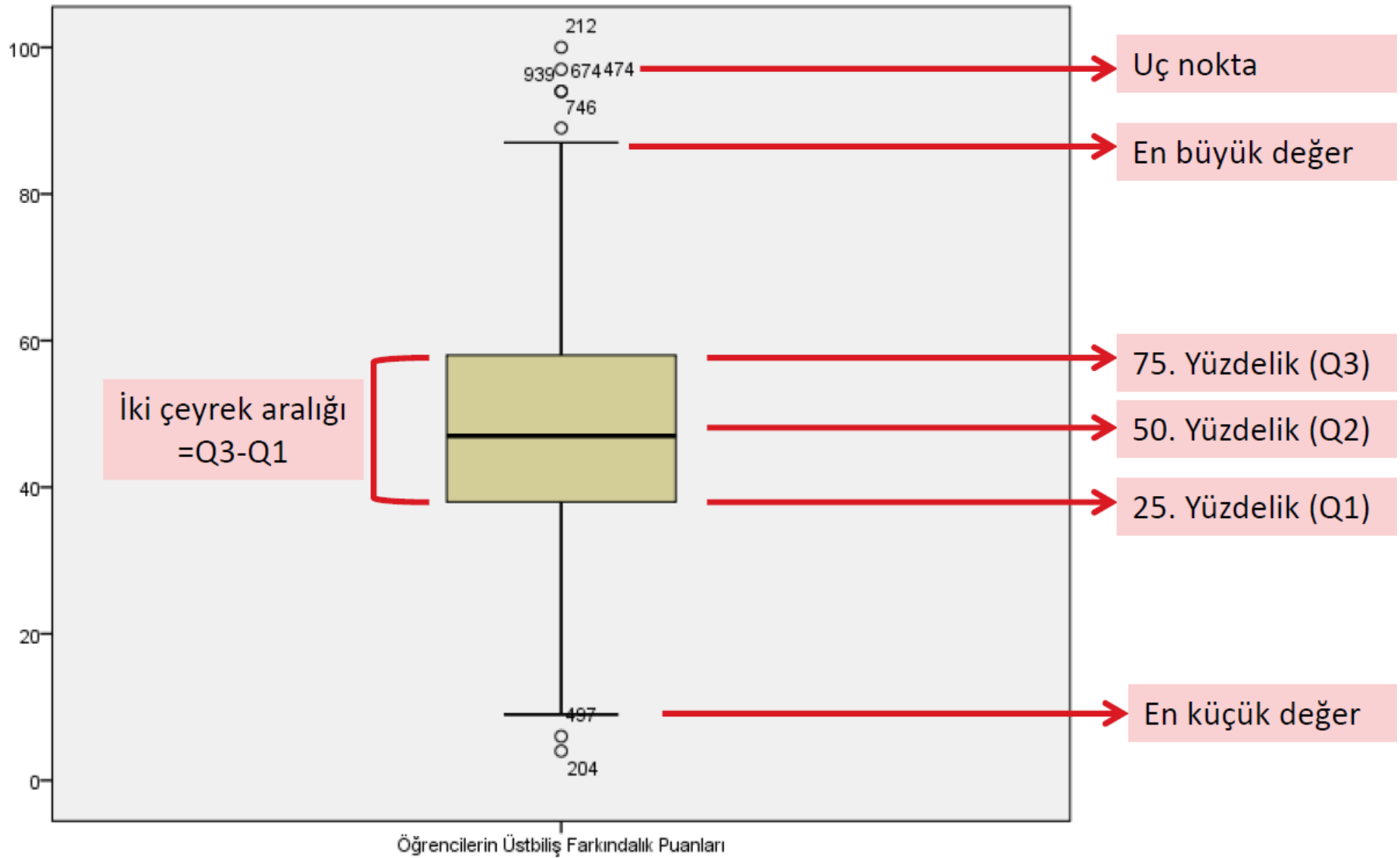


Uniform



Skewed

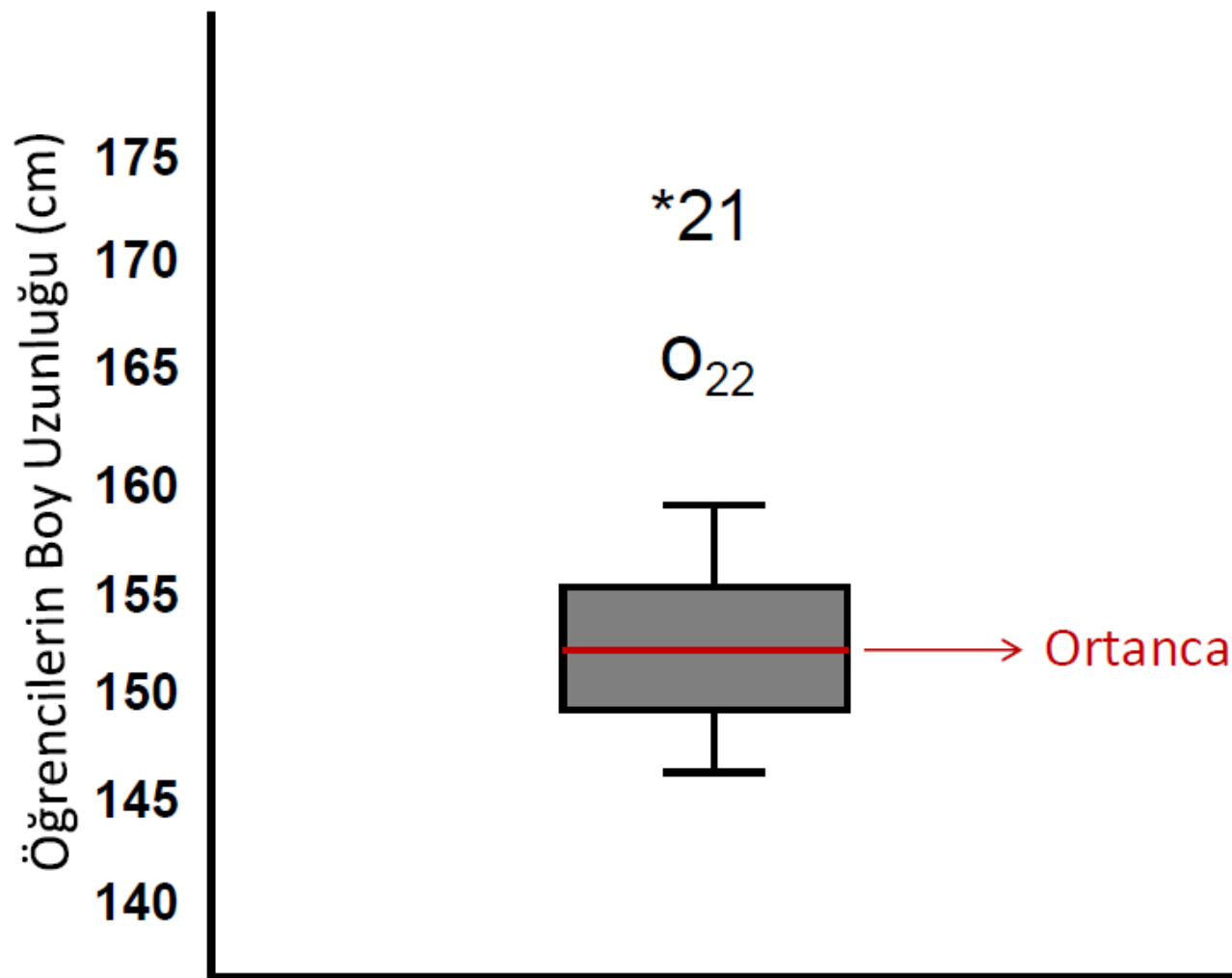
Figure 2-17



1000 öğrencinin üstbilis puanlarının box-plot grafiği

Simetrik Dağılım

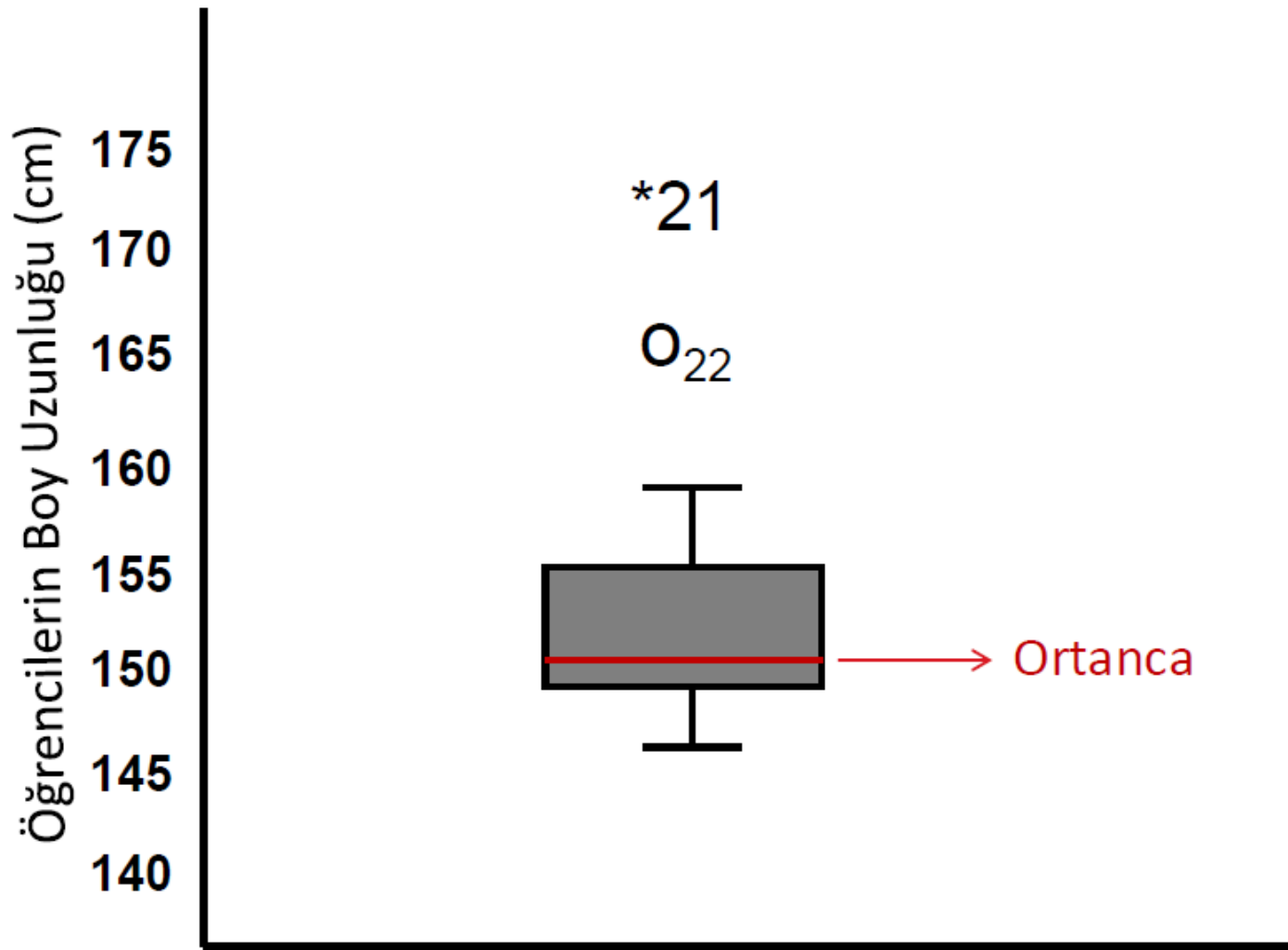
Ortalama=Ortanca=Tepe Değeri



Bir sınıftaki öğrencilerin boy uzunluklarının box-plot grafiği

Sağa Çarpık Dağılım

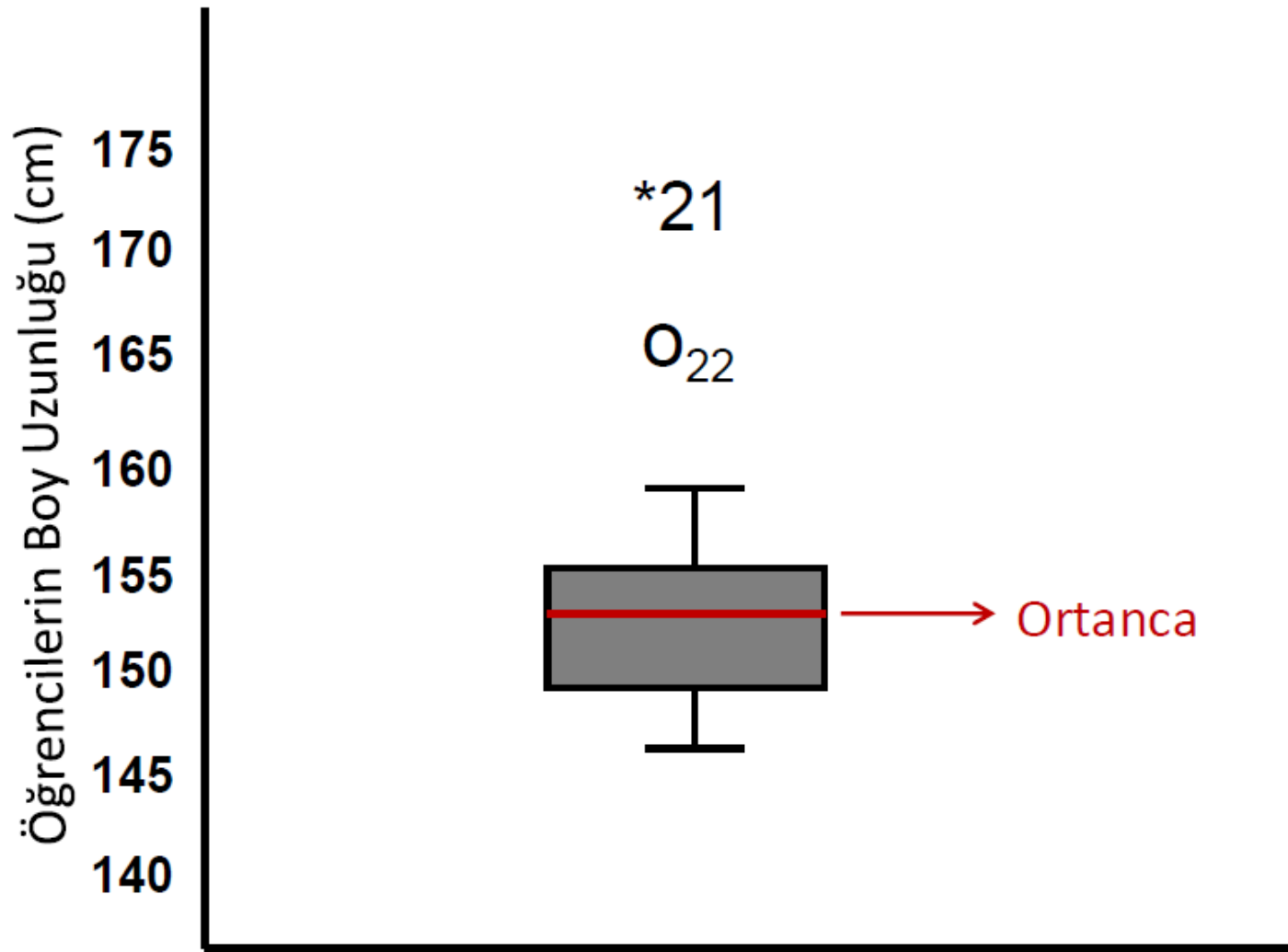
Tepe Değeri < Ortanca < Ortalama



Bir sınıftaki öğrencilerin boy uzunluklarının box-plot grafiği

Sola Çarpık Dağılım

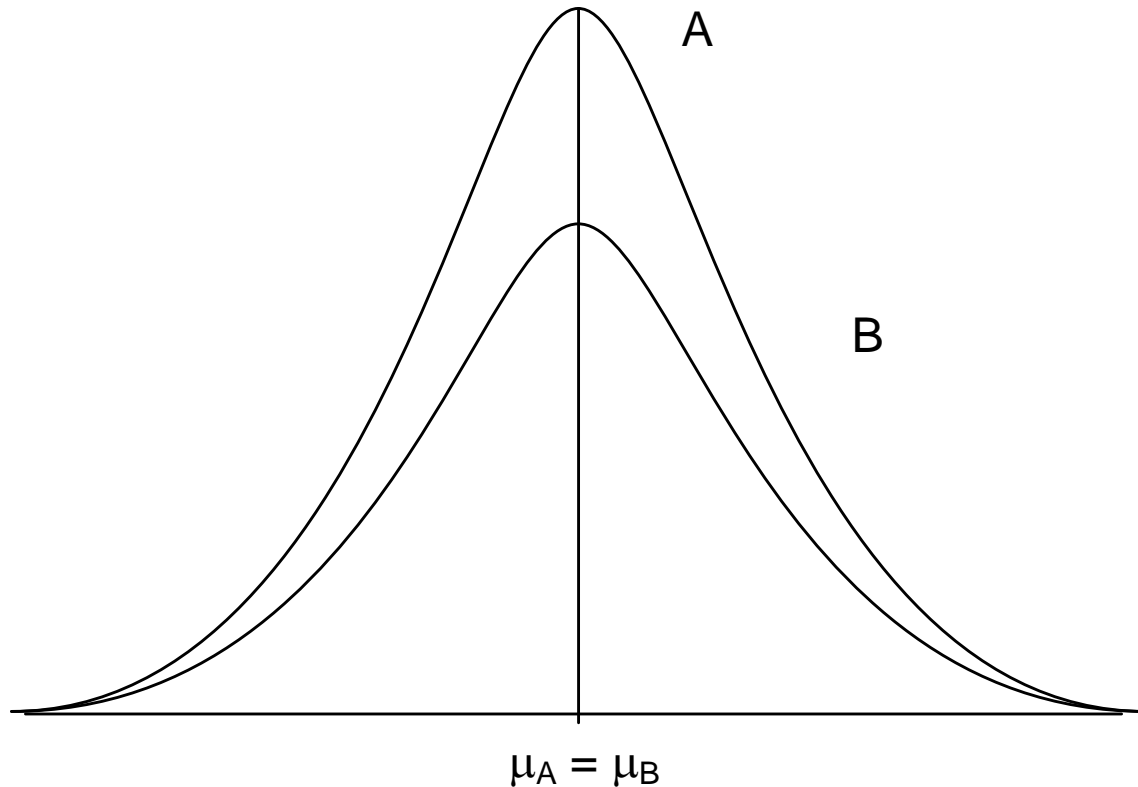
Ortalama < Ortanca < Tepe Değeri



Bir sınıftaki öğrencilerin boy uzunluklarının box-plot grafiği

16) Basıklık Ölçüsü

Aşağıdaki A ve B dağılımlarının ortalamaları, değişkenlik ölçülerinin aynı olmasından dolayı ve hatta ikisinin de simetrik olmalarından dolayı bu iki dağılışı ayırt etmek için Basıklık Ölçüsü kullanılır.



Herhangi bir olasılık fonksiyonunun şekli ile ilgili parametrelerden bir tanesi de basıklık ölçüsüdür. **Basıklık Ölçüsü** ortalamaya göre dördüncü momentten gidilerek hesaplanır ve α_4 olarak gösterilir.

$$\alpha_4 = \frac{\mu_4}{\sigma^4} \quad \text{Basit Seri İçin} \quad \mu_4 = \frac{\sum_{i=1}^n (x_i - \mu)^4}{n}$$

$\alpha_4 = 3$ ise **Seri Normal**

$\alpha_4 < 3$ ise **Seri Basık**

$\alpha_4 > 3$ ise **Seri Sivri Ya da Yüksek**

THE END.



Dinlediğiniz İçin Teşekkür Ederim...

Dr. Gökhan AKSU