3. Hafta

Merkezi Yayılım Ölçüleri

İstatistiksel hesaplamalarda genelde ortalama olarak aritmetik ortalama kullanıldığından aksi belirtilmediği sürece ortalama denildiğinde daima aritmetik ortalamayı anlayacağız.

Bu bölümde ilk olarak ortalamadan sapma anlayışını açıklayacağız.

X: Tesadüfi değişken

$X_1, X_2, \dots, X_n$: gözlem değerleri

$\bar{X}$: Aritmetik Ortalama

$X_1 - \bar{X}, X_2 - \bar{X}, \dots, X_n - \bar{X}$: Aritmetik ortalamadan sapmalar

ÖRNEK: X değişkeninin aldığı gözlem değerlerinin ortalamadan sapmaları aşağıdaki tabloda gösterilmiştir.

X	f	$X - \bar{X}$
3	10	-3,29
5	23	-1,29
6	35	-0,29
8	23	1,71
10	9	3,71
	n=100	$\sum(X_j - \bar{X}) = 0$

$\sum X = 629$

$\bar{X} = 6{,}29$

Tablonun üçüncü sütununda görüldüğü üzere ortalamadan sapmaların toplamı sıfır olarak hesaplanmıştır.

Ortalamadan sapmanın negatif işaretli olması verinin ortalamadan küçük (ortalamanın solunda), pozitif işaretli olması ise ortalamadan büyük (ortalamanın sağında) olduğunu gösterir. Sapmanın küçük olması verinin ortalamaya yakın (ortalama civarında), büyük olması ise verinin ortalamadan uzak olduğunu göstermektedir.

Aritmetik ortalamanın temsil gücünün yüksek olabilmesi verinin ortalama civarında olmasına bağlıdır. Bu durumda sapmaların küçük olması beklenmektedir. Bu ölçüye standart sapma deriz. σ

Varyans ve Standart Sapma

Bir verinin ortalamadan sapmaların karelerinin ortalamasına varyans (σ^2), karesel ortalamasına veya varyansın kareköküne standart sapma (σ) denir.

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

$$\sigma^2 = \frac{1}{n}\{\sum {X_i}^2 - 2\bar{X}\sum X_i + n\bar{X}^2\}$$

$$\sigma^2 = \frac{1}{n}\sum {X_i}^2 - \frac{2\bar{X}n\bar{X}}{n} + \bar{X}^2$$

$$\sigma^2 = \frac{1}{n}\sum {X_i}^2 - \bar{X}^2$$

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

Standart sapmanın yanı sıra ortalama sapma, değişim aralığı ve değişim katsayısı gibi sayısal göstergelerde verinin aritmetik ortalama etrafındaki yayılımının ölçümüdür.

$$Ortalama\ Sapma = \frac{\sum |X_i - \bar{X}|}{n}$$

$$Depişim\ Aralığı = Max(X_i) - Min(X_i)$$

$$Değişim\ Katsayısı = \frac{\sigma}{\bar{X}}$$ (Yüzdelik değer olarak yorumlanır)

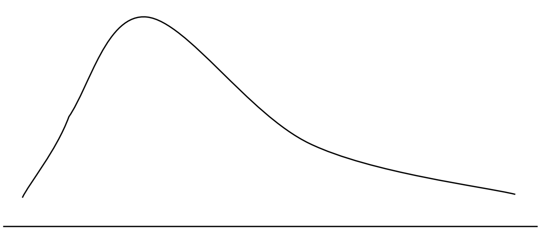
Momentler (Ortalamaya Göre)

Ortalamaya göre k-ıncı moment verinin ortalamaya göre sapmalarının k-ıncı kuvvetlerinin ortalamasıdır.

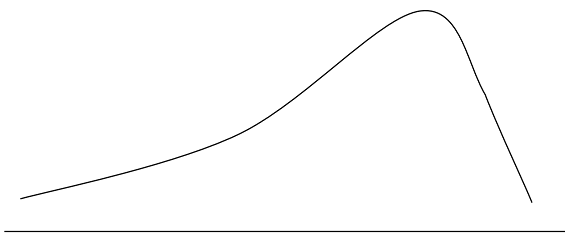
$$\mu_{\boldsymbol{k}} = \frac{\sum (X_i - \bar{X})^k}{n}$$

Çarpıklık ve Basıklık Katsayıları

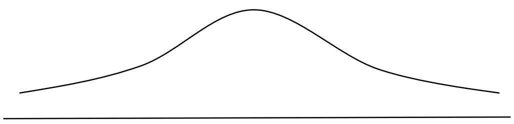
Örneklemden elde edilen histogramın simetrik olmasını ortalama etrafında toplanmasının optimal olduğunu yani histogramın normal eğri altında kalması gerektiğini hatırlayalım. Aşağıdaki şekiller histogramın normal eğri altında kalmayıp simetrisinin ve basıklığının bozulduğu durumları göstermektedir.



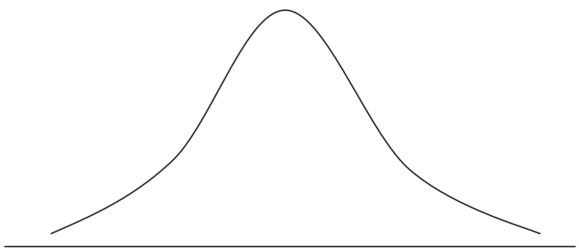
Sağa Çarpık



Sola Çarpık



Basık



Sivri

<u>Histogramdaki basıklık ve çarpıklığın nedenleri:</u>

Sağa çarpık; Veride az sayıda ortalamadan çok büyük aşırı uç değer olması

Sola çarpık; Veride az sayıda ortalamadan çok küçük aşırı uç değer olması

Basıklık; Verinin standart sapmasının çok büyük olması

Sivrilik; Verinin standart sapmasının çok küçük olması

<u>Çarpıklık ve basıklık formülleri:</u>

$$\alpha_3 = \frac{\mu_3}{\sigma^3} = \begin{cases} = 0 & : \quad Simetrik \\ < 0 & : Sola\ Çarpık \\ > 0 & : Sağa\ Çarpık \end{cases}$$

$$\alpha_4 = \frac{\mu_{\mathbf{4}}}{\sigma^4} = \begin{cases} = 3 & : Normal \\ < 3 & : Basık \\ > 3 & : Sivri \end{cases}$$

Pearson Çarpıklık Ölçüsü:

$$\frac{\bar{X} - Mod}{\sigma} \quad \text{veya} \quad \frac{3(\bar{X} - Medyan)}{\sigma}$$