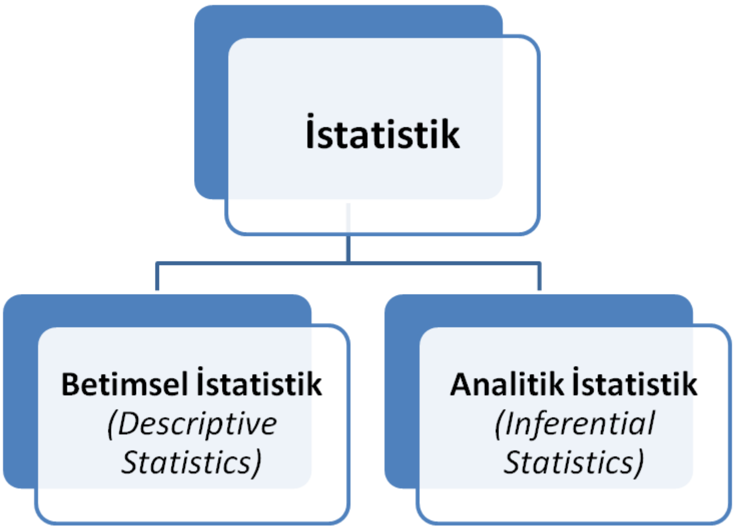
Betimsel İstatistik 1

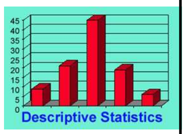
Doç. Dr. Meryem ULUSKAN

2021



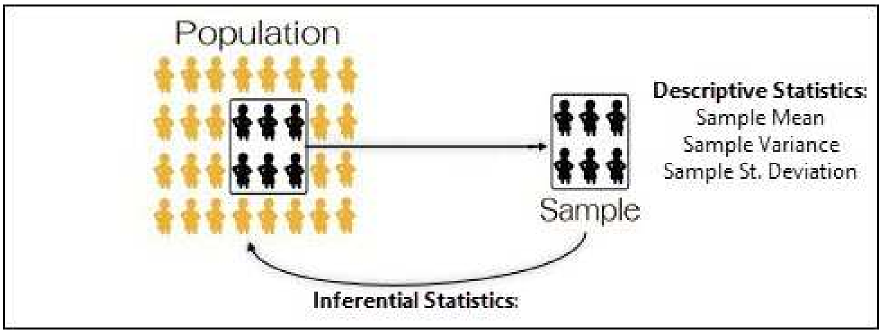
İstatistiksel yöntemler kullanılış amaçlarına göre iki grupta toplanabilir:

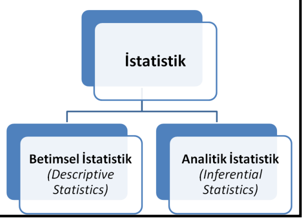
- **Betimsel İstatistik** *(descriptive statistics)*
  - Gözlenmiş durumları bazı istatistiksel ölçülerle açıklamaya (tasvir etmeye) yarayan yöntemler



- **Analitik İstatistik** *(inferential statistics)*
  - Gözlenmiş durumlardan elde edilen bilgilerden gözlenmemiş durumlar hakkında yorumlamalar, genellemeler yapılmasında yararlanılan yöntemler

Betimsel İstatistik vs. Analitik İstatistik

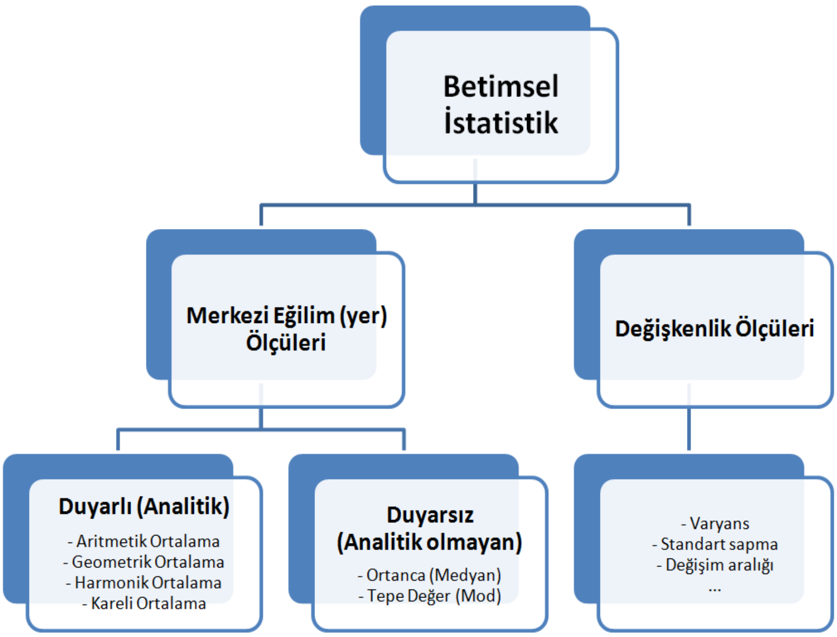




"Data don't make any sense,
we will have to resort to statistics."

Toplanan verinin anlam ifade etmesi için ne yapılmalı?
Betimsel İstatistik için?

Doç. Dr. Meryem ULUSKAN

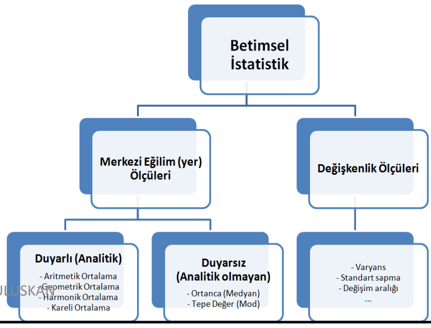


Doç. Dr. Meryem ULUSKAN

Betimsel İstatistik

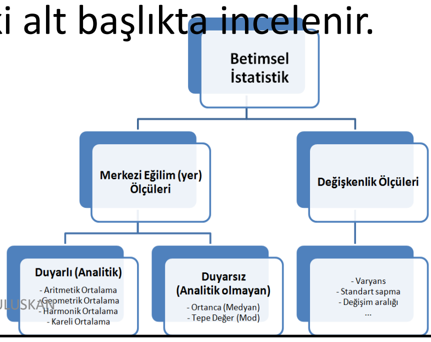
- Anakütlenin bazı özelliklerinin belirlenmesi
- Örnekten türetilen istatistiklerden hareketle karşı gelen anakütle parametresi tahmin edilir.
- Betimsel istatistikler
  - Merkezi Eğilim (yer) Ölçüleri (Measures of Central Tendancy)
  - Değişkenlik Ölçüleri

  (Measures of Variability)



Merkezî Eğilim (Yer) Ölçüleri

- Anakütlenin elemanlarının ya da yapılan gözlemlerin nerede yoğunlaştığını ya da yığıldığını gösterir.
- Gözlemlerin x-ekseni üzerindeki durumu
- Duyarlı (Analitik) yada Duyarsız (Analitik olmayan) olmak üzere iki alt başlıkta incelenir.



Önemli eşitlikler:

- X_i : gözlem değeri; i=1, 2, 3, . . . , n olsun.

i- $$\sum_{i=1}^{n} X_i = X_1 + X_2 + \cdots + X_n$$

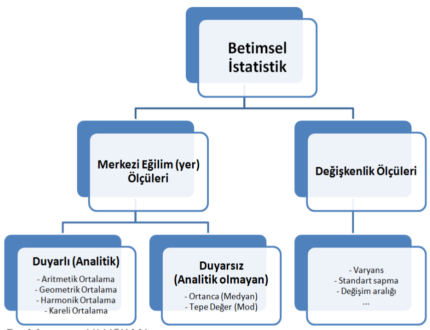
ii- $$\sum_{i=1}^{n} X_i^2 = X_1^2 + X_2^2 + \cdots + X_n^2$$

iii- $$\sum_{i=1}^{n} c_i X_i = c_1 X_1 + c_2 X_2 + \cdots + c_n X_n$$

iv- $$\sum_{i=1}^{n} (X_i \pm Y_i \pm Z_i) = \sum_{i=1}^{n} X_i \pm \sum_{i=1}^{n} Y_i \pm \sum_{i=1}^{n} Z_i$$

1. Duyarlı Yer Ölçüleri

- Bir anakütleden alınan gözlem değerleri $X_1, X_2, : . . , X_n$
- <u>*Gözlemlerin tümünün katılımıyla*</u> hesaplanan bu yer ölçüleri genelde ortalama olarak adlandırılır.



1. Duyarlı Yer Ölçüleri

- **1.1. Aritmetik Ortalama**
- Gözlemlerin toplamının gözlem sayısına bölünmesiyle elde edilir.

$\bar{X}$ sembolüyle gösterilir.

Veriler ne şekilde toplamış olabilir?

1.1. Aritmetik Ortalama

- Basit serilerde,

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

- Frekans serilerinde,
- i-inci gözlem değeri X_i , karşı gelen frekansı n_i i=1, 2, . . . ,k

$$\sum_{i=1}^{k} n_i = n$$

$$\bar{X} = \frac{\sum_{i=1}^{k} n_i X_i}{n}$$

1.1. Aritmetik Ortalama

Sınıflandırılmış serilerde ise,

- i-inci sınıfın orta değeri X_i^*
- sınıfın frekansı n_i , i=1, 2, . . . ,m,

$\sum_{i=1}^{m} n_i = n$ olmak üzere,

$$\bar{X} = \frac{\sum_{i=1}^{m} n_i X_i^*}{n}$$

şeklinde hesaplanır.

1.1. Aritmetik Ortalama

Özellikler:

- Gözlemlerin aritmetik ortalamadan sapmalarının (farklarının) cebirsel toplamı sıfırdır.

1.1. Aritmetik Ortalama
Özellikler:

- Gözlemlerin aritmetik ortalamadan sapmalarının (farklarının) cebirsel toplamı sıfırdır.

$$\sum_{i=1}^{n}\left(X_i-\bar{X}\right)=0$$

$$\sum_{i=1}^{n}\left(X_i-\bar{X}\right)=\sum X_i-\sum \bar{X}=\sum X_i-n\bar{X}=\sum X_i-n\frac{\sum X_i}{n}=0$$

1.1. Aritmetik Ortalama
Özellikler:

Gözlemlerin bir $a \neq 0$ sayısından sapmalarının karelerinin toplamı ancak ve ancak $\boldsymbol{a} = ??$ ise, enküçük olur.

1,2,3,4,5, gözlemlerimiz olsun

1.1. Aritmetik Ortalama

iii- Gözlemlerden

- n_1 tanesinin aritmetik ortalaması $\overline{Y}_1$,
- n_2 tanesinin aritmetik ortalaması $\overline{Y}_2$ ve
- ...
- n_k tanesinin aritmetik ortalaması $\overline{Y}_k$ ise,

tüm gözlemlerin aritmetik ortalaması,

$$\overline{Y} =$$

Doç. Dr. Meryem ULUSKAN

1.1. Aritmetik Ortalama

Özellikler:

iii- Gözlemlerden

- n_1 tanesinin aritmetik ortalaması $\overline{Y}_1$,
- n_2 tanesinin aritmetik ortalaması $\overline{Y}_2$ ve
- ...
- n_k tanesinin aritmetik ortalaması $\overline{Y}_k$ ise,

tüm gözlemlerin aritmetik ortalaması,

$$\overline{Y} = \frac{\sum_{i=1}^{k} n_i \overline{Y}_i}{\sum_{i=1}^{k} n_i}$$

olur ve *ağırlıklı (weighted) aritmetik ortalama* olarak adlandırılır.

Doç. Dr. Meryem ULUSKAN

1.1. Aritmetik Ortalama
Özellikler:

iv- Bir $a \neq 0$ ve $Y_i=X_i-a$, i=1. 2, . . . , n ise,

$$\bar{X} = a + \frac{\sum Y_i}{n}$$ olur.

1.1. Aritmetik Ortalama
Özellikler:

v) Uç değerler aritmetik ortalamayı etkiler ve kendine doğru kaymasına neden olur.

Aritmetik ortalamanın uç değerlerden etkilenmemesi için,

- sıralanan verilerden eşit oranda olmak üzere, enküçükler ve enbüyükler çıkarılır.
- Geriye kalan verilerin hesaplanan aritmetik ortalaması *düzeltilmiş aritmetik ortalama* (trimmed arithmetic average) olarak adlandırılır.

1.1. Aritmetik Ortalama

Özellikler:

vi) Bir değişkenin aritmetik ortalaması, bu değişken için yapılan gözlemlerin ağırlıklı merkezini gösterir.

vii) aritmetik ortalamanın gözlem değerlerinden birine eşit olması gereklidir??;

fakat gözlemlerin …………. yer alması gerekir.

1.1. Aritmetik Ortalama

- **ÖRNEK-1:** Rassal olarak seçilen 50 öğrencinin dakikadaki kalp atış sayıları (dk) ölçülmüş ve izleyen değerler bulunmuştur

  88 67 72 62 81 60 73 70 72 75 72 75 80 77 80
  83 80 72 65 88 65 70 72 60 75 67 88 77 70 87
  63 64 67 70 77 63 83 64 72 80 72 85 64 75 73
  57 63 64 64 73.

a) Kalp atışlarının frekans serisini oluşturunuz ve aritmetik ortalamasını hesaplayınız.
b) Sınıf aralığı 5 olmak üzere, -den az şeklinde sınıflandırınız ve aritmetik ortalamasını hesaplayınız.
c) Yukarıda hesapladığınız ortalamaları karşılaştırınız.

AO- Örnek 1- a

Öğrencilerin dakikadaki kalp atışlarının frekans serisi

Kalp Atışı, dk	n_i		Kalp Atışı, dk	n_i		Kalp Atışı, dk	n_i
57	*1*		67	*3*		80	*4*
60	*2*		70	*4*		81	*1*
62	*1*		72	*7*		83	*2*
63	*3*		73	*3*		85	*1*
64	*5*		75	*4*		87	*1*
65	*2*		77	*3*		88	*3*

AO- Örnek 1 - a

- Bu Frekans serisinin aritmetik ortalaması ??

AO- Örnek 1 - b

- Sınıf aralığı 5 olarak verildiğinden, enküçük değerden başlanarak –den az şeklinde sınıflandırma yapılırsa,
- Kaçtan başlayacak?
- …

AO- Örnek 1 - b

- Sınıf aralığı 5 olarak verildiğinden, enküçük değerden başlanarak –den az şeklinde sınıflandırma yapılırsa aşağıdaki sınıflandırılmış seri elde edilir.

AO- Örnek 1 - b

- Sınıf orta değerleri ile

AO- Örnek 1-b

- Aritmetik ortalama ?? m=?

AO- Örnek 1-c

- Hesaplanan ortalamalar karşılaştırılırsa ??
- Frekans serisinin aritmetik ortalaması = 72,32
- Sınıflandırılmış serinin aritmetik ortalaması = 73,30
- Neden??
- Hangisi daha gerçek ortalama?

AO- ÖRNEK-2:

- Bir işlemede saat ücretiyle çalışanlardan rassal olarak seçilenlerin saat ücretleri (TL) ve sayıları tabloda verilmiştir. Çalışanların saat ücreti ortalamasını hesaplayınız.

Sınıflar	n_i
10-15	7
15-20	12
20-25	15
25-30	16
30-40	14
40-50	15
50-60	20
60-80	18
80-100	10
100-120	5
120-150	5
150-200	4
200-250	3

Çözüm

AO- ÖRNEK-2:

- Farklı sınıf aralıkları, aritmetik ortalaması?

Sınıflar	n_i		X_i^*
10-15	7		12,5
15-20	12		17,5
20-25	15		22,5
25-30	16		27,5

Toplam=50

AO- ÖRNEK-2:

Sınıflar	n_i		X_i^*
30-40	14		35
40-50	15		45
50-60	20		55

Toplam=49

AO- ÖRNEK-2:

Sınıflar	n_i		X_i^*
60-80	18		70
80-100	10		90
100-120	5		110

Toplam=33

AO- ÖRNEK-2:

Sınıflar	n_i		X_i^*
120-150	5		135

Toplam=5

AO- ÖRNEK-2:

Sınıflar	n_i		X_i^*
150-200	4		175
200-250	3		225

Toplam=7

$$\bar{X}_5 = \frac{\sum_{i=1}^{2} n_i X_i^*}{\sum_{i=1}^{2} n_i} = \frac{4 \cdot 175 + 3 \cdot 225}{4+3} = 196{,}43$$

AO- ÖRNEK-2:

- Ortalama saat ücreti de belirlenen alt sınıf ortalamalarının ortalaması olarak,

$$\bar{X} = \frac{\sum_{i=1}^{5} n_j \bar{X}_j}{\sum_{i=1}^{5} n_j} =$$

1.2. Geometrik Ortalama

- Duyarlı ortalama
- $X_1, X_2, \ldots, X_n$ n-tane sayı olmak üzere, bunların geometrik ortalaması, çarpımlarının n-inci dereceden kökü olarak tanımlanır.
- Sembolik gösterimle

$$\bar{X}_G = \sqrt[n]{X_1 * X_2 * \ldots * X_n} \qquad = \sqrt[n]{\prod_{i=1}^{n} X_i}$$

1.2. Geometrik Ortalama

- Geometrik ortalamanın hesabında logaritma işleminin özelliklerinden yararlanılır

$$\bar{X}_G = \sqrt[n]{X_1 * X_2 * \ldots * X_n} = (X_1 * X_2 * \ldots * X_n)^{\frac{1}{n}}$$

Doç. Dr. Meryem ULUSKAN

Logaritma hatırlatma

$$\log_a x = y$$

Doç. Dr. Meryem ULUSKAN

Logaritma hatırlatma

$$\log_a x = y \rightarrow x = a^y$$

$$\log_a x^m =$$

Logaritma hatırlatma

$$\log_a x = y \rightarrow x = a^y$$

$$\log_a x^m = \mathrm{m} * \log_a x$$

$$\log_a(x * y) = \log_a x + \log_a y$$

$$\log_a(x/y) = \log_a x - \log_a y$$

1.2. Geometrik Ortalama

- Geometrik ortalamanın hesabında logaritma işleminin özelliklerinden yararlanılır

$$\bar{X}_G = \sqrt[n]{X_1 * X_2 * \ldots * X_n} = (X_1 * X_2 * \ldots * X_n)^{\frac{1}{n}}$$

- Her iki tarafın logaritması alınırsa

$$\log \bar{X}_G =$$

Doç. Dr. Meryem ULUSKAN

1.2. Geometrik Ortalama

- Geometrik ortalamanın hesabında logaritma işleminin özelliklerinden yararlanılır

$$\bar{X}_G = \sqrt[n]{X_1 * X_2 * \ldots * X_n} = (X_1 * X_2 * \ldots * X_n)^{\frac{1}{n}}$$

- Her iki tarafın logaritması alınırsa

$$\log \bar{X}_G = \frac{1}{n} \log(X_1 * X_2 * \ldots * X_n) =$$

Doç. Dr. Meryem ULUSKAN

1.2. Geometrik Ortalama

- Geometrik ortalamanın hesabında logaritma işleminin özelliklerinden yararlanılır

$$\bar{X}_G = \sqrt[n]{X_1 * X_2 * \ldots * X_n} = (X_1 * X_2 * \ldots * X_n)^{\frac{1}{n}}$$

- Her iki tarafın logaritması alınırsa

$$\log \bar{X}_G = \frac{1}{n}\log(X_1 * X_2 * \ldots * X_n) = \frac{1}{n}\sum_{i=1}^{n} \log X_i$$

- İşlemler sonrasında her iki tarafın antilogaritması alınarak geometrik ortalama elde edilir.

1.2. Geometrik Ortalama

- İlgilenilen bir olayda gözlem değerleri bir önceki değere bağlı olarak değişiyor ve değişimin hızı (oranı) belirlenmek isteniyorsa geometrik ortalama hesaplanır.
- Örneğin, nüfus, milli gelir, ekonomik değişim, finans ve benzeri konularda geometrik ortalamadan yararlanılır.

1.2. Geometrik Ortalama
Özellikler

- i) Gözlem değerlerinin geometrik ortalamaya oranları çarpımı 1'e eşittir.

$$\frac{X_1}{\bar{\bar{X}}_G} \cdot \frac{X_2}{\bar{\bar{X}}_G} \cdot \;\cdot\; \cdot \frac{X_n}{\bar{\bar{X}}_G} =$$

Doç. Dr. Meryem ULUSKAN

1.2. Geometrik Ortalama
Özellikler

- i) Gözlem değerlerinin geometrik ortalamaya oranları çarpımı 1'e eşittir.

$$\frac{X_1}{\bar{\bar{X}}_G} \cdot \frac{X_2}{\bar{\bar{X}}_G} \cdot \;\cdot\; \cdot \frac{X_n}{\bar{\bar{X}}_G} = \frac{\prod_{i=1}^{n} X_i}{\left(\bar{\bar{X}}_G\right)^n} = \frac{\left(\bar{\bar{X}}_G\right)^n}{\left(\bar{\bar{X}}_G\right)^n} = 1$$

Doç. Dr. Meryem ULUSKAN

1.2. Geometrik Ortalama
Özellikler

- ii) Gözlem değerlerinin logaritmalarıyla, geometrik ortalamanın logaritması arasındaki farkların cebirsel toplamı sıfırdır.

$$\sum_{i=1}^{n}\left(\log X_i - \log \bar{X}_G\right) = 0 \text{ olur.}$$

1.2. Geometrik Ortalama
Özellikler

- iii) Gözlemlerin, $X_1, X_2, \ldots, X_k$, sıklıkları (frekansları) sırasıyla, $n_1, n_2, \ldots, n_k$ ve $n = \sum_{i=1}^{k} n_i$ ise,

$$\bar{X}_G = \sqrt[n]{(X_1 \cdot X_1 \cdot \cdots X_1)(X_2 \cdot X_2 \cdot \cdots X_2) \cdots (X_k \cdot X_k \cdot \cdots X_k)}$$

$$\bar{X}_G = \sqrt[n]{X_1^{n_1} \cdot X_2^{n_2} \cdots X_k^{n_k}}$$

$$\log \bar{X}_G = \frac{1}{n}\sum_{i=1}^{k} n_i \log X_i$$

- Sınıflandırılmış serilerde de sınıf orta değeri ve sınıftaki gözlem sayısı dikkate alınarak, yukarıdaki eşitlikle geometrik ortalama hesaplanır.

1.2. Geometrik Ortalama
Özellikler

- Aritmetik ortalama gibi uç değerlere karşı duyarlı değildir.
- Gözlemlerden birinin sıfır ya da negatif olması durumunda geometrik ortalama hesaplanamaz.