OLASILIK VE İSTATİSTİK

İstatistik: Belirli bir amaç için veri toplama, toplanan verileri işleme, betimleme, çözümleme ve yorumlama yöntem ve teknikleri bilimidir.

Parametre bir ana kütleden hesaplanan aritmetik ortalama, standart sapma oran gibi özet değerlere denir.

İstatistik örneklemeye başvurulduğunda örneklemden hesaplanan aritmetik ortalama standart sapma oran gibi özet değerlere ise istatistik denir.

İstatistikin konusu kolektif olaylardır yani değişkenlik gösteren olaylardır. Nüfus en güzel örnektir.

Veri: Belirli amaçlarla toplanan benzer sayılara veri adı verilir.

İstatistik yöntem biliminde genellikle şu sıraya uyulur:

1) Verileri toplama
2) Toplanan verilerin işlenip düzenlenmesi
3) Düzenlenmiş verilerin tablo veya grafikler halinde sunulması
4) İstatistiksel çözümleme tahmin ve karar

NOT:

Parametre → Ana kütle N

İstatistik → Örneklem n

Birim: Ana kütle veya örneklemi oluşturan gözlem konusu olarak anlatan kolektif olaylardan her birine birim adı verilir.

Değişken: Birimin sahip olduğu ve bir diğerinden ayırt edilmesine yarayan özelliğe değişken adı verilir.

- Nicel (Sayısal): nüfus
- Nitel (Sözel): cinsiyet, medeni durum

Şık: Değişkenin ortaya çıkış biçimine şık denir.

Ana Kütle: Hakkında bilgi edinilmek istenen ve biçimsel homojenliğe sahip kolektif olay niteliğindeki birimlerin oluşturduğu topluluğa denir.

Örneklem: Gözlemlenmek üzere ana kütleden seçilen birimlerden oluşan ve ana kütlenin doğal bir parçası olan alt topluluğa örneklem denir.

DEĞİŞKENLERİN ÖLÇÜLMESİ

Ölçme:

Ölçme Düzeyleri:

1) Sınıflayıcı Ölçme Düzeyi →
   Sadece sınıflama yapılır
2) Sınırlayıcı Ölçme Düzeyi →
   Büyüklük eşitlik küçüklük

3) Eşit Aralıklı Ölçme Düzeyi →

Başlangıç 0 noktası yoktur
Başlangıç gelişigüzeldir
Toplama çıkartma var
Çarpma bölme var ama mantıksal değil
Takvim zekâ düzeyi hava sıcaklığı örnekleridir

4) Oranlı Ölçme Düzeyi →

4 işlem var
Sayı sırası anlamlı
Sayılar arası anlamlı
Başlangıç 0 noktasının bir anlamı vardır
Ağırlık hız boy elektrik faturası gibi örnekler verilebilir

Ortalama: İstatistikte bir seriyi temsil etmeye ve özetlemeye yarayan tek bir rakama ortalama denir. İstatistikte ortalama aynı zamanda merkezi eğilimi ölçüsü olarak da bilinmektedir.

$$X_{min} < Ort > X_{max}$$

DUYARLI ORTALAMALAR:

$$H < G < \bar{X} < K$$

Aritmetik Ortalama:

$$\frac{\sum Xi}{N} \rightarrow \text{Basit Seri}$$

$$\frac{\sum Ni\,.xi}{\sum Ni} \rightarrow \text{Sınıflanmış Seri}$$

$$\frac{\sum Ni\,.mi}{\sum Ni} \rightarrow \text{Gruplanmış Seri}$$ (mi = grup orta nokta)

Örnek: Aşağıdaki not serilerinin aritmetik ortalamasını hesaplayın

Basit Seri:

Xi

4
5
7
8
16

$$\frac{40}{5} = 8$$

Sınıflanmış Seri:

Xi	Ni
2	3
3	2
4	1
5	4

6+6+4+20 = 36

$$\frac{36}{10} = 3{,}6$$

Gruplanmış Seri:

Gruplar	Ni
1-3'ten az	3
3-5'ten az	3
5-7'den az	4

mi	Nimi
2	6
4	12
6	24

24+12+6 = 42

$$\frac{42}{10} = 4{,}2$$

Geometrik Ortalama:

$$\text{LogG} = \frac{\sum logXi}{N} \rightarrow \text{Basit Seri}$$

$$\text{LogG} = \frac{\sum NilogXi}{\sum Ni} \rightarrow \text{Sınıflanmış Seri}$$

$$\text{LogG} = \frac{\sum Nilogmi}{\sum Ni} \rightarrow \text{Gruplanmış Seri}$$

Örnek: Aşağıdaki not serilerinin geometrik ortalamasını hesaplayın

Basit Seri:

Xi	logXi
4	0,602
5	0,699
7	0,845
8	0,903
16	1,204

0,602+0,699+0,845+0,903+1,204

= 4,253

$\text{LogG} = \frac{4,253}{5} = 0,8506$ G = 7,09

Sınıflanmış Seri:

Xi	logXi	Ni
2	0,301	3
3	0,477	2
4	0,602	1
5	0,699	4

0,903+0,954+0,602+2,796

= 5,255

$\text{LogG} = \frac{5,255}{10} = 0,5255$ G = 3,35

Gruplanmış Seri:

Gruplar	Ni	mi	logmi
1-3'ten az	3	2	0,301
3-5'ten az	3	4	0,602
5-7'den az	4	6	0,778

0,903+1,806+3,112

= 5,821

$\text{LogG} = \frac{5,821}{10} = 0,5821$

G = 3,82

Harmonik Ortalama:

$$H = \frac{N}{\sum 1/Xi} \rightarrow \text{Basit Seri}$$

$$H = \frac{\sum Ni}{\sum Ni/Xi} \rightarrow \text{Sınıflanmış Seri}$$

$$H = \frac{\sum Ni}{\sum Ni/mi} \rightarrow \text{Gruplanmış Seri}$$

Örnek: Aşağıdaki not serilerinin harmonik ortalamasını hesaplayınız.

Basit Seri:

Xi	1/Xi
4	0,25
5	0,20
7	0,14
8	0,12
16	0,06

0,25+0,20+0,14+0,12+0,06

= 0,77

$$H = \frac{5}{0{,}77} = 6{,}49$$

Sınıflanmış Seri:

Xi	Ni	Ni/Xi
2	3	1,5
3	2	0,67
4	1	0,25
5	4	0,80

1,5+0,67+0,25+0,80

= 3,22

$H = \frac{10}{3,22} = 3,11$

Gruplanmış Seri:

Gruplar	Ni	mi	Ni/mi
1-3'ten az	3	2	1,5
3-5'ten az	3	4	0,75
5-7'den az	4	6	0,67

1,5+0,75+0,67

= 2,92

$H = \frac{10}{2,92} = 3,42$

Kareli Ortalama:

$$K = \sqrt{\frac{\sum(Xi).(Xi)}{N}}$$ → Basit Seri

$$K = \sqrt{\frac{\sum(Xi).(Xi).(Ni)}{\sum Ni}}$$ → Sınıflanmış Seri

$$K = \sqrt{\frac{\sum(mi).(mi).(Ni)}{\sum Ni}}$$ → Gruplanmış Seri

Örnek: Aşağıdaki not serilerinin kareleri ortalamalarını hesaplayınız.

Basit Seri:

Xi	Xi^2
4	16
5	25
7	49
8	64
16	256

16+25+49+64+256

= 410

$$K = \sqrt{\frac{410}{5}} = \sqrt{82} = 9{,}06$$

Sınıflanmış Seri:

Xi	Ni	Xi^2	$Ni.Xi^2$
2	3	4	12
3	2	9	18
4	1	16	16
5	4	25	100

12+18+16+100

= 146

$$K = \sqrt{\frac{146}{10}} = \sqrt{14{,}6} = 3{,}82$$

Gruplanmış Seri:

Gruplar	Ni	mi	mi^2
1-3’ten az	3	2	4
3-5’ten az	3	4	16
5-7’den az	4	6	36

12+48+144 = 204

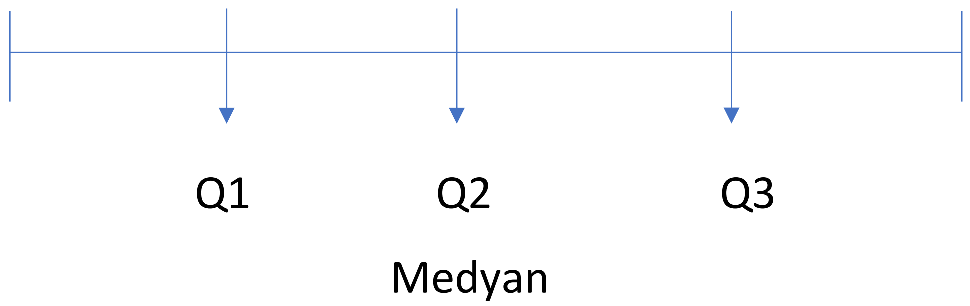
$$K = \sqrt{\frac{204}{10}} = 4{,}51$$

Desil (Sentil): Bir seriyi 10 eşit parçaya bölenlere desil denir.

Medyan: Sıralı bir veri setindeki orta değer medyandır.

Kartil: Bir seriyi 4 eşit parçaya bölenlere kartil denir.

Q1 = 1.kartil Q2 = 2.kartil Q3 = 3.kartil



Q1 = $\frac{(N+2)}{4}$. Sıradaki terim

Q2 = $\frac{(N+1)}{2}$. Sıradaki terim

Q3 = $\frac{(3N+2)}{4}$. Sıradaki terim

Örnek: Aşağıdaki basit tüketim serisinin medyan ve kartillerini hesaplayın.

Xi
21
42
67
88

Q1 = $\frac{4+2}{4}$ = 1,5.terim $\frac{21+42}{2}$ = 31,5

Q2 = $\frac{4+1}{2}$ = 2,5.terim $\frac{67+42}{2}$ = 54,5

Q3 = $\frac{12+2}{4}$ = 3,5.terim $\frac{67+88}{2}$ = 77,5

Medyan = Q2 = 54,5

Örnek: Aşağıdaki basit tüketim serisinin medyan ve kartillerini hesaplayın.

Xi
17
42
55
62
87

Q1 = $\frac{5+2}{4} = \frac{7}{4}$ = 1,75.terim 0,25.17+0,75.42 = 35,75

Q2 = $\frac{5+1}{2} = \frac{6}{2}$ = 3.terim = 55

Q3 = $\frac{15+2}{4} = \frac{17}{4}$ = 4,25.terim 0,75.62+0,25.87 = 68,25

Medyan = Q2 = 55

Örnek: Aşağıdaki basit tüketim serisinin medyan ve kartillerini hesaplayın.

Xi
10
23
38
57
75
90

Q1 = $\frac{6+2}{4}$ = 2.terim = 23

Q2 = $\frac{6+1}{2}$ = 3,5.terim $\frac{57+38}{2}$ = 47,5

Q3 = $\frac{18+2}{4}$ = 5.terim = 75

Medyan = Q2 = 47,5

Örnek: Aşağıdaki basit tüketim serisinin medyan ve kartillerini hesaplayın.

Xi
12
25
37
43
66
82
94

Q1 = $\frac{7+2}{4}$ = 2,25.terim = 0,75.25+0,25.37 = 28

Q2 = $\frac{7+1}{2}$ = 4.terim = 43

Q3 = $\frac{21+2}{4}$ = 5,75.terim = 0,25.66+0,75.82 = 78

Medyan = Q2 = 43

Değişkenlik Ölçüleri:

Ortalama sapma, standart sapma, oransal değişkenlik ölçüsü niteliğindeki değişim katsayısının hesabında bütün gözlem değerleri dikkate alınır. Ancak değişim genişliği hesabında bütün gözlem değeri hesaba alınmaz.

1) Ortalama Sapma: Ortalama mutlak sapma adıyla da anılan bu ölçü teriminin aritmetik ortalamadan mutlak sapmalarının aritmetik ortalamasıdır.

O.S = $\frac{\sum|Xi-\bar{x}|}{N}$ → Basit Seri

O.S = $\frac{\sum Ni|Xi-\bar{x}|}{\sum Ni}$ → Sınıflanmış Seri

O.S = $\frac{\sum Ni|mi-\bar{x}|}{\sum Ni}$ → Gruplanmış Seri

Örnek: Aşağıdaki basit satış serisinin ortalama sapmasını hesaplayınız.

Xi
21,000
23,000
24,000
26,000
30,000
32,000

21,000+23,000+24,000+26,000+30,000+32,000

= 156,000

$$\bar{X} = \frac{156{,}000}{6} = 26{,}000$$

|Xi- $\bar{X}$|
5,000
3,000
2,000
0
4,000
6,000

5,000+3,000+2,000+4,000+6,000

= 20,000

$$\text{O.S.} = \frac{20{,}000}{6} = 3333{,}33$$

Örnek: Aşağıdaki sınıflanmış satış serisinin ortalama sapmasını hesaplayınız.

Xi	Ni	Ni.Xi	|Xi- $\bar{X}$ |	Ni.|Xi- $\bar{X}$ |
1,000	2	2,000	5200	10400
5,000	7	35,000	1200	8400
7,000	9	63,000	800	7200
12,000	2	24,000	5800	11600

10400+ 8400+7200+11600 = 37600

$$\text{O.S.} = \frac{37600}{20} = 1880$$

Standart Sapma: Kareli ortalama sapma adıyla da anılan bu ölçü terimlerin aritmetik ortalamalarının cebirsel sapmalarının kareli ortalamasıdır. Standart sapma ortalama sapma ile aynı temele dayanmakla beraber ortalama sapmanın mutlak sapmalarla hesaplanması sakıncasını taşımadığı için matematiksel işlemlerle en uygun kullanım alanına sahip önemli bir değişkenlik ölçüsüdür. Öte yandan standart sapma veya onun karesi olan varyans normal dağılımın önemli bir parametresi olduğu gibi ayrıca örnekleme regresyon ve korelasyon konularındaki istatistiksel analizlere temel oluşturur.

σ → standart sapma → S

σ → varyans → S^2

$$\sigma = \sqrt{\frac{\sum(Xi - \bar{x}).(Xi - \bar{x})}{N}} \rightarrow \text{Basit Seri}$$

$$\sigma = \sqrt{\frac{\sum Ni(Xi - \bar{x}).(Xi - \bar{x})}{\sum Ni}} \rightarrow \text{Sınıflanmış Seri}$$

$$\sigma = \sqrt{\frac{\sum Ni(mi - \bar{x}).(mi - \bar{x})}{\sum N\dot{I}}} \rightarrow \text{Gruplanmış Seri}$$

ÖDEV:

Standart Sapma Özellikleri:

1) Standart sapma eksi bir değer alamaz.
2) Standart sapma değişkene ait ölçü birimini içerir.
3) Standart sapma istatistik analiz raporlarımızda normal dağılıma uygun veriler için kullanılabilir.
4) Standart sapma verilerin aritmetik ortalamaya göre nasıl bir yayılım gösterdiğini anlatır.
5) Aritmetik ortalamaları aynı olan farklı veri gruplarından; açıklığı küçük olanın standart sapması büyüktür.
6) Standart sapmasının küçük olması bir veri grubundaki değerlerin birbirine yakın olduğunu gösterir. Standart sapmanın büyük olması ise veri grubundaki değerlerin birbirinden uzak olduğunu gösterir.
7) Standart sapmanın küçük olması, ortalamadan sapmanın ve riskin azlığının; büyük olması ise ortalamadan sapmanın ve riskin çokluğunun bir göstergesidir.

Örnek: Aşağıdaki basit satış serisinin standart sapmasını hesaplayınız.

Xi	(Xi- $\bar{X}$)	(Xi- $\bar{X}$)^2
21,000	-5,000	25.10^6
23,000	-3,000	9.10^6
24,000	-2,000	4.10^6
26,000	0	0
30,000	4,000	16.10^6
32,000	6,000	36.10^6

$$\sigma = \sqrt{\frac{90.000.000}{6}}$$

$$= 3872,98$$

Örnek: Aşağıdaki sınıflanmış satış serisinin standart sapmasını hesaplayınız.

Xi	Ni	(Xi- $\bar{X}$)	(Xi- $\bar{X}$)^2	Ni(Xi- $\bar{X}$)^2
1,000	2	-5,200	27,040,000	54,080,000
5,000	7	-1,200	1,440,000	10,080,000
7,000	9	800	640,000	5,760,000
12,000	2	5,800	33,640,000	67,280,000

54,080,000+10,080,000+5,760,000+67,280,000

= 137,200,000

$$\sigma = \sqrt{\frac{137{,}200{,}000}{20}} = 2619{,}16$$

Örnek: Aşağıdaki gruplanmış satış serisinin standart sapmasını hesaplayınız.

Gruplar	Ni	mi	mi- $\bar{x}$	Ni(mi- $\bar{x}$)^2
0-2’den az	3	1	-2,625	20,67
2-4’ten az	7	3	-0,625	2,73
4-6’dan az	4	5	1,375	7,56
6-8’den az	2	7	3,375	22,78

20,67+2,73+7,56+22,78

$$\sigma = \sqrt{\frac{53{,}74}{16}} = \sqrt{3{,}358} = 1{,}83$$

Değişim Katsayısı:

Oransal değişkenlik ölçüsü niteliğindeki değişim kat sayısı herhangi bir serinin standart sapmasının aritmetik ortalamaya bölünmesi sonucunda elde edilmektedir. Değişim kay sayısı formülü;

D.K. $\rightarrow \frac{\sigma}{\bar{x}}$

Değişim kat sayısı küçük olan serilerin diğerine göre daha az değişken oldukları söylenir. Bunun anlamı ise değişim kat sayısı küçük olan serinin daha homojen birimlerden oluştuğudur.

Örnek:

A (istatistik sınavı) aritmetik ortalamasının 50 puan standart sapmasının 19,20 puan

B (işletme sınavı) aritmetik ortalamasının 60 puan standart sapmasının 27,36 puan olarak hesaplandığını varsayarak değişim kat sayılarını hesaplayıp kıyaslayalım.

$D.K_A = \frac{19{,}20}{50} = 0{,}384$

$D.K_B = \frac{27{,}36}{60} = 0{,}456$

Yorum: A serisinin değişim kat sayısı B serisininkinden küçük olduğuna göre A serisi B serisinden daha az değişkendir. Diğer bir değişle A serisi daha homojen birimlerden oluşmaktadır.

Örnek:

C (ağırlık) aritmetik ortalamasının 62 kg standart sapmasının 10 kg

D (boy uzunluğu) aritmetik ortalamasının 170 cm standart sapmasının 20 cm olduğunu varsayarsak değişim kat sayılarını hesaplayıp kıyaslama yapınız.

$$D.K_c = \frac{10}{62} = 0{,}161$$

$$D.K_D = \frac{20}{170} = 0{,}117$$

Yorum: D serisinin değişim kat sayısı C serisininkinden küçük olduğuna göre D serisi C serisinden daha az değişkendir. Diğer bir değişle D serisi daha homojen birimlerden oluşmaktadır.

OLASILIK (İHTİMAL de denebilir)

Deney: Bir aktivitenin gözlenlenmesi ve ölçüm yapma şekilleridir.

Örnek Uzay (S): Bir deneyde çıkabilecek bütün sonuçların kümesine örnek uzay denir.

Sonuç (Çıktı): Bir deneyde elde edilebilecek sonuçların her birine çıktı denir.

- Deneyin bir tane sonucu varsa basit olay denir.
- Deneyin birden fazla sonucu varsa birleşik olay denir.

S = {Yazı, Tura} S = {1,2,3,4,5,6}

$\bar{A}$ → Tümleyen

AuB → A birleşim B

AnB → A kesişim B

P(A) → A olayının olma olasılığı

P(A) ≥ 0

0'a eşit olması imkânsız olaydır.

P(A) ≤ 1

1'e eşit olması kesin olaydır.

0 ≤ P(A) ≤ 1

Örnek:

Bir madeni para iki kez atılıyor. En az 1 yazı gelmesi olasılığı nedir.

S {YY, YT, TY, TT}

P(A) = $\frac{3}{4}$

- Olasılık için;

AnB = Ø → Ayrık olay

P(AuB) = P(A) + P(B) – P(AnB)

Koşullu Olasılık ve Bağımsızlık:

A ve B olayları meydana gelme olasılıkları 0 olmayan iki olay olup bir S örnek uzayı içinde tanımlansın. B olayı olduktan sonra A olayının olması olasılığı A'nın koşullu olasılığı olarak adlandırılır.

"$P(A/B) = \frac{P(AnB)}{P(B)}$; $P(B)>0$"

"$P(B/A) = \frac{P(AnB)}{P(A)}$; $P(A)>0$"

Örnek:

İki zar birlikte atılıyor toplam 6 gelmişse zarlardan birinin 2 gelmesi olasılığı nedir?

A = {(2,1) (2,3) (2,4) (2,5) (2,6) … (6,2)}

B = {(1,5) (2,4) (3,3) (4,2) (5,1)}

AnB = {(2,4) (4,2)}

$P(B) = \frac{5}{36}$

$P(A) = \frac{10}{36}$

$P(AnB) = \frac{2}{36}$

$P(A/B) = \frac{2}{5}$

Not:

P(AnB) = P(A/B). P(B)

veya olasılığın çarpım kuralı

P(AnB) = P(B/A). P(A)

Örnek:

Bir kutuda 30 tane disket vardır. Bunların 5 tanesi arızalıdır. Kutudan 4 tane rasgele disket arka arkaya çekiliyor. 4 disketinde sağlam olması olasılığı nedir?

$P(A_1 n A_2 n A_3 n A_4) = P(A_1).\ P(A_2/A_1).\ P(A_3/A_1 u A_2).\ P(A_4/A_1 u A_2 u A_3)$

$$\frac{25}{30}.\frac{24}{29}.\frac{23}{28}.\frac{22}{27} = 0{,}46$$

Örnek:

Bir çantada 6 beyaz ve 7 siyah top diğer bir çanta da ise 8 beyaz ve 5 siyah top vardır. İlk çantadan rastgele bir top seçilip ikinci çantaya atılıyor. İkinci çantadan bir top seçiliyor. Seçilen bu topun siyah gelme olasılığı nedir.

1. İlk torbadan beyaz çekilirse;

$$\frac{6}{13}.\frac{5}{14} = \frac{30}{182}$$

2. İlk torbadan siyah çekilirse;

$$\frac{7}{13}.\frac{6}{14} = \frac{42}{182}$$

Sonuçları toplayalım:

$$\frac{30}{182} + \frac{42}{182} = \frac{72}{182} = 0{,}395$$

- $P(A/B) = P(A)$

Bağımsız iki olaydır. $P(AnB) = P(A)P(B)$

$P(B/A) = P(B)$

Örnek:

Bir fabrikada üretilen malların %50'si A makinasında, %30'u B makinasında, %20'si C makinasında üretilmektedir. Bu makinalardaki üretimden A'dakinin %3'ü, B'dekinin %4'ü, C'dekinin %5'i arızalı olduğu bilinmektedir. Üretilen mallardan rastgele alınan bir tanesinin arızalı olması olasılığı nedir?

X ile gösterelim

A.

B.

C.

$$P(X) = P(A)P(X/A) + P(B)P(X/B) + P(C)P(X/C)$$

$$= \frac{50}{100}.\frac{3}{100} + \frac{30}{100}.\frac{4}{100} + \frac{20}{100}.\frac{2}{100} = 0{,}037$$

Örnek:

A, B, C gibi 3 atlet yarışıyor. A'nın kazanma olasılığı B'nin kazanma olasılığına eşit fakat C'ninkinin iki katıdır. A, B, C kazanma olasılıkları nedir?

A'nın kazanma olasılığı $\frac{40}{100}$ = %40

B'nın kazanma olasılığı $\frac{40}{100}$ = %40

C'nın kazanma olasılığı $\frac{20}{100}$ = %20

Örnek:

15 tane disketin 5 tanesinin bozuk olduğu biliniyor. Bunlardan 3 tane örnek alınırsa;

A) Hepsinin sağlam olması

(15,3) kombinasyonu → $\frac{15.14.13}{3.2.1}$ = 455

(10,3) kombinasyonu → $\frac{10.9.8}{3.2.1}$ = 120

P(A) = $\frac{120}{455} = \frac{24}{91}$

B) 1 tanesinin bozuk olması

(5,1). (10,2) kombinasyonu → 5.45 = 225

C) En az 1 tanesinin bozuk olması

P(C) = 1- $\frac{24}{91} = \frac{67}{91}$

Örnek:

1 çift zar birlikte atılıyor. Aşağıdaki koşullar altında toplamı 10 veya daha büyük olması olasılığını hesaplayınız.

A) 1.zar 5 gelmişse

A {(56), (51),(52),(55),(53),(54)}

P(A) = $\frac{1}{3}$

B) Zarlardan biri 5 ise

B {(51), (15), (52), (25), (53), (35), (54), (45), (55), (65), (56)}

P(B) = $\frac{3}{11}$

Örnek:

A torbasında 3 kırmızı 5 beyaz, B torbasında 2 kırmızı 1 beyaz, C torbasında 2 kırmızı 3 beyaz top vardır. Rastgele seçilen bir torbadan rastgele de bir top çekilmiştir. Çekilen topun kırmızı olduğu bilindiğine göre bunun A torbasından çekilmiş olması olasılığı nedir?

$$P(A/K) = \frac{P(A).P(K\backslash A)}{P(A)P(K\backslash A)+P(B)P(K\backslash B)+P(C)P(K\backslash C)}$$

$$= \frac{\frac{1}{3}.\frac{3}{8}}{\frac{1}{3}.\frac{3}{8}+\frac{1}{3}.\frac{2}{3}+\frac{1}{3}.\frac{2}{5}} = 0{,}26$$

Örnek:

A ve B iki olay olup A'nın olasılığı 3/8, B'nin olasılığı 5/8 ve AuB'nin olasılığı 3/4 olarak verildiğine göre P(A\B) ve P(B\A) olasılıklarını hesaplayınız.

P(AuB) = P(A) + P(B) – P(AnB)

3/4 = 3/8 + 5/8 – P(AnB)

P(AnB) = 2/8

$$P(A/B) = \frac{P(AnB)}{P(B)} = \frac{2/8}{5/8} = 0{,}4$$

$$P(B/A) = \frac{P(AnB)}{P(A)} = \frac{2/8}{3/8} = 2/3$$

Örnek:

Bilgisayar Mühendisliği bölümündeki öğrencilerin %25'i elektronik dersinden %30'u istatistik dersinden %10'u da her iki dersten kalmışlardır.

100x öğrenci olsun

15xi sadece elektronikten kalmıştır

20xi sadece istatistikten kalmıştır

10xi ikisinden de kalmıştır

A) Öğrenci elektronik dersinden kalmışsa istatistik dersinden de kalma olasılığı

$$= \frac{10}{25} = \frac{2}{5} = 0{,}4$$

B) Öğrenci istatistik dersinden kalmışsa elektronik dersinden de kalma olasılığı

$$= \frac{10}{30} = \frac{1}{3}$$

C) Öğrencinin istatistik veya elektronik dersinden kalma olasılığı

$$= \frac{45}{100} = 0{,}45$$

Örnek:

A ve B iki olay olsun. A'nın olasılığı 1/2, B'nin olasılığı 1/3 P(AnB) = 1/4 ise P(A/B), P(B/A), P(AuB) ve P(Ā/B) nedir?

P(A/B) = $\frac{P(AnB)}{P(B)}$ = $\frac{1/4}{1/3}$ = 0,75

P(B/A) = $\frac{P(AnB)}{P(A)}$ = $\frac{1/4}{1/2}$ = 0,5

P(AuB) = 1/2 + 1/3 – 1/4 = 9/12 = 0,75

P(Ā/B) = $\frac{1-P(AuB)}{1-P(B)}$ = $\frac{0,25}{2/3}$ = 5/8

Örnekleme ve Dağılımları:

1.aşama: planlama

2.aşama: veri toplama

3.aşama: verilerin çözümlenmesi ve sonuçların elde edilmesi

Örnekleme: Bir ana kütleden rassal olarak seçilmiş ve daha az sayıda birimden oluşan bir örneklemi inceleyerek bu ana kütle hakkında genelleme yapma işlemidir.

Ana Kütle: Hakkında bilgi edinilmek istenen ve biçimsel bir homojenliğe sahip (belirli tanıma uyan) birimlerin oluşturduğu topluluğa ana kütle; bu ana kütleden seçilen birimlerden oluşan ve ana kütlelerin doğal bir parçası olan alt topluluğa ise örneklem adı verilir.

Çerçeve:

1) Basit Rassal Örnekleme:

• Oluşturulan evren listesinden örnekleme birimlerinin ölçüt olmaksızın çekilmesidir.

Örneğin; 1.000 kişilik bir hasta listesinden 100 hasta seçileceğini varsayalım. Rastgele sayılar tablosunu kullanarak ya da günümüzde daha sıklıkla kullanılan rastgele sayı üreten yazılım veya internet siteleri yardımıyla 1'den 1.000'e kadar 100 rastgele sayı seçiliyor. Bu şekilde elde edilen sayıların temsil ettiği hastalar çalışmaya dahil edilir.

$$ {}_{N}C_{n} = \frac{N!}{n!(N-n)!} $$

2) Tabakalı Örnekleme:

• Tabakalama yöntemi ile örnekleme yönteminde popülasyon üyeleri ortak özelliklere göre 2 veya daha fazla sayıda katmana ayrılırlar. Daha sonra her bir katman için elde edilen sonuçlar birleştirilerek popilasyon parametreleri ile ilgili tahmin yapılır.

→ Ankara üniversitesi, Siyasi Bilgiler Fakültesi öğrencilerinin mezuniyet sonrası ne tür işlerde çalıştıklarına ilişkin bir araştırma yapılmak istendiğinde mezunlar önce bitirdikleri bölümlere göre katmanlara ayrılırlar daha sonra her bir katmandan rastsal yöntemle örnek oluştururlar.

3) Çok Kademeli Örnekleme:

• Çok kademeli örnekleme olasılığı taban olarak kullanarak herhangi bir popülasyon hakkında fikir veya bilgi elde edinmek için kullanılır.

4) Kümeleme Örneklemesi:

• Bu yöntemle örnekleme de popülasyon üyeleri ortak özelliklere göre çok sayıda kümeye ayrılır ve her bir kümeden rastsal örnekleme ile istenen sayıda eleman elde edilerek örnek oluşturulur. Oluşturulan her bir örnek için elde edilen sonuçlar kullanılarak popülasyon parametrelerine ilişkin tahmin yapılır.

• Kümeleme yöntemi ile örnekleme, bir ülkenin çeşitli bölgeler veya bir ülkenin bir bölgesinin çeşitli yerel oluşumlar itibariyle istatistiksel analizinde yaygın olarak kullanılır.

→ Ailelerin geliri ile ilgili araştırmada; mahalle, sokak veya binaların, öğrencilerle ilgili araştırmada okulların bir kısmı örnek olarak seçilir.

5) Çok Aşamalı Örnekleme:

- Örneklem seçiminde tabakalı ve küme örnekleme yöntemleri bir arada kullanılır.

Örneğin; Şehirleri tabaka mahalleleri küme olarak kabul edersek. 81 ilden 10 ili seçip, sonrasında 10 ildeki 8.000 mahalleden 800'ünü seçmemiz durumunda çok aşamalı tabaklı-küme örnekleme yapmış oluruz.

Örneklem Seçimi:

İadeli Seçim: Genelde kullanılır.

İadesiz Seçim: Yerine koymadan

Örneklemede Hata Kavramı:

Rassal Hata: Bir ana kütleden çok sayıda "n" hacimli örneklem çektiğimizi ve bunların söz gelimi aritmatik ortalamalarını hesapladığımız zaman ortalamaların ana kütle ortalamasında genellikle farklı olduğunu görmekteyiz. Bu farklara rassal hata ya da örnekleme hatası adı verilir.

Standart Hata: İstatistiğin örnekleme bölünmesinin standart sapmasından başka bir şey değildir. Kısacası standart hata örneklem istatistiğinin ana kütle parametresine yaklaşık olarak ne kadar yakın veya uzak olduğunu gösteren karel, ortalama temelinde bir ölçüdür.

σ_X → Örneklem standart hatası

İadesiz Seçim → $\frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$

n = örneklem hacmi

N = Anakütle hacmi

σ = Anakütle standart sapması

İadesiz olduğunu gösterir → $\frac{n}{N} \geq 0{,}05$

İadeli Seçim → $\frac{\sigma}{\sqrt{n}}$

İadeli olduğunu gösterir → $\frac{n}{N} < 0{,}05$

Örnek:

Sonsuz büyük bir anakütlenin standart sapmasının 200, örneklem hacminin ise 25 olduğu varsayalım. Örneklem hacmi 25'ten 100'e yükseltildiğinde ortalamanın standart hatasında ne gibi bir değişiklik olur.

$$\sigma_X = \frac{\sigma}{\sqrt{n}} = \frac{200}{\sqrt{25}} = \frac{200}{5} = 40$$

$$\sigma_X = \frac{\sigma}{\sqrt{n}} = \frac{200}{\sqrt{100}} = \frac{200}{10} = 20$$

Yorum: Örneklem hacminin artmasıyla standart hata azalmıştır.

Not: Standart sapmayı azaltmanın en iyi yolu örnek birim sayısını (hacmini) artırmaktır.

Örneklem Hacminin Belirlenmesi:

$$n = \sigma^2 / \sigma_x^2$$

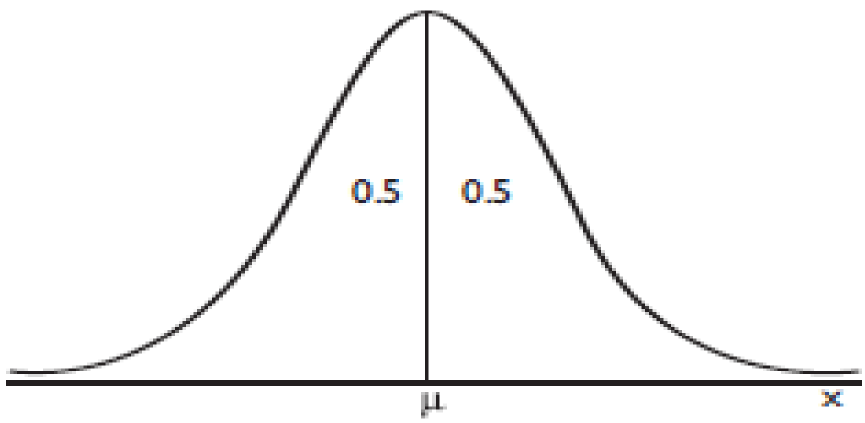
Örnek:

500 hacimlik bir anakütlenin standart sapması 85 cm'dir. Ortalamanın standart sapmasının 15 cm'den fazla olmaması için örneklemin kaç birimden oluşması gerektiğini iadeli seçim durumuna göre hesaplayınız.

$$n = \frac{85.85}{15.15} = 32{,}11$$

Normal Dağılım:

Bir anakütleden rassal olarak seçilmiş çok sayıda örneklemden elde edilen istatistiklerinin dağılımlarının normale uyması belirli olasılıklar çerçevesinde anakütle parametreleri hakkında tahmin yapmamızı ve karara varmamızı sağlamakta ve dolayısıyla istatistiksel tekniklerinin temelini oluşturmaktadır.



μ = normal dağılımın aritmetik ortaması

Olasılık yoğunluk fonksiyonları

$$f(x) = \frac{1}{\sigma\sqrt{2\,\pi}}.\ e^{-(x-\mu)(x-\mu)/2ss.ss}$$

$$-\infty < x < \infty$$

X rassal değişimi normal dağılıma sahiptir demek için:

X N(μ, σ) e = 2,71828 π = 3,14159

$$P(a<x<b) = \int_a^b \frac{1}{2\sigma(\sqrt{2\,\pi})}\ .\ e^{-(x-\mu)(x-\mu)/2ss.ss}\ dx$$

çözmek gerekir. (Normal Dağılım Tablosu)

(İntegrali almak yerine normal dağılım tablosundan değeri yerine koyacağız)

Not:

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\,\pi}}.\ e^{-(x-\mu)(x-\mu)/2ss.ss}\ dx = 1$$

- Eğri altında kalan alan 1’dir.

Özellikler:

1) $\int_{-\infty}^{\infty} f(X)dx = 1$

2) Normal dağılım simetriktir.
$\int_{-\infty}^{\mu} f(x)dx = \int_{\mu}^{\infty} f(x)dx = 0{,}5$

3) x = μ olursa
$f(x) = \frac{1}{\sigma\sqrt{2\pi}}$

Not:

- Normal dağılımın aritmetik ölçüsü $a_3 = 0$
- Normal dağılımın basıklık ölçüsü $a_4 = 3$

Standart Normal Dağılım:

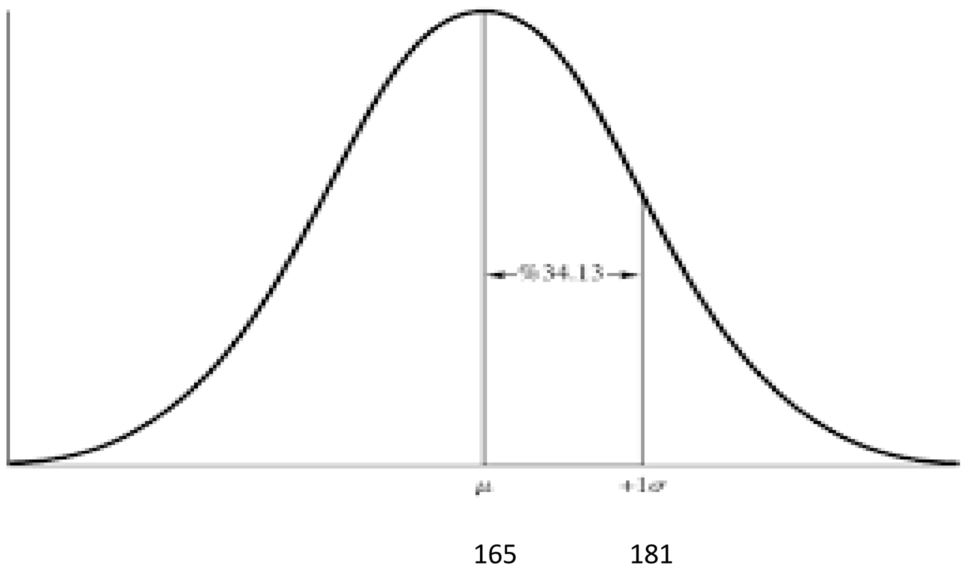
Gerçek uygulamalarda standart olmayan sürekli bir rassal değişken sıfırdan farklı ortalama ve birden farklı standart sapma değerleri ile normal dağılım gösterir. Böylesi durumlarda tablonun kullanılabilmesi için normal dağılımı bir rassal değişkenin bir dönüştürme neticesinde standart normal dağılımı bir rassal değişkene dönüştürmesi gerekir. Bu amaçla normal dağılımı gösteren Z rassal değişkenine dönüştürmesi yapılacaktır ve buna standartlaştırma adı verilir.

Örnek:

Belli bir yaş grubunda bulunan 50.000 erkeğin boy uzunluklarının 165 cm aritmetik ortalaması ve 165 σ normal dağılıma sahip olduğunu varsayarak boy uzunlukları

a) 165-181 cm arasında
b) 173 cmden fazla
c) 153-185 cm arasında
d) 157 cmden az veya 177 cmden fazla

a)

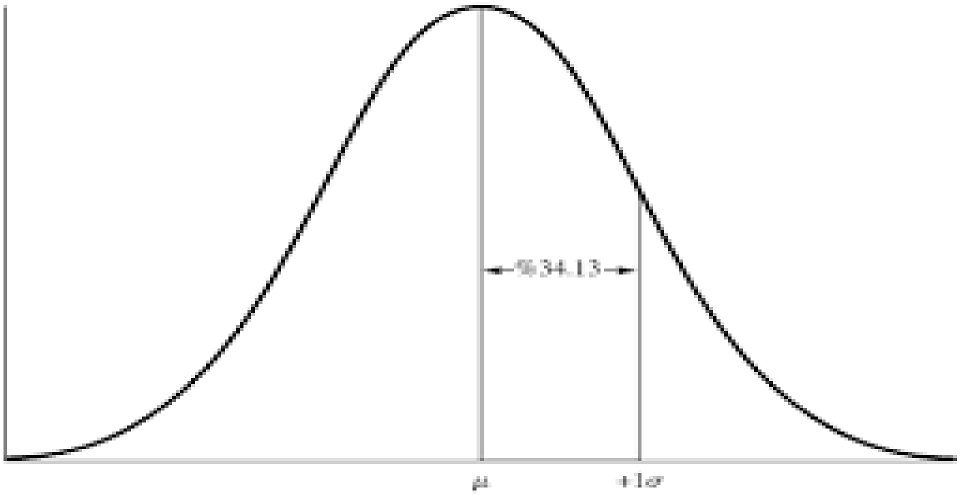


$Z_1 = \frac{181-165}{16} = +1{,}00$

P(a<z<+1,00) = 0,3413

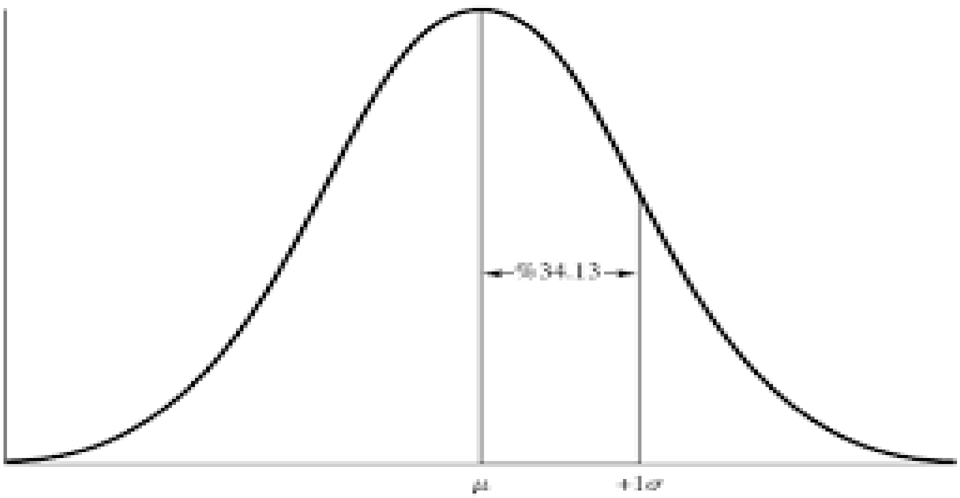
50,000(0,3413) = 17065

b)



165 173

$$Z_1 = \frac{173-165}{16} = +0{,}50$$



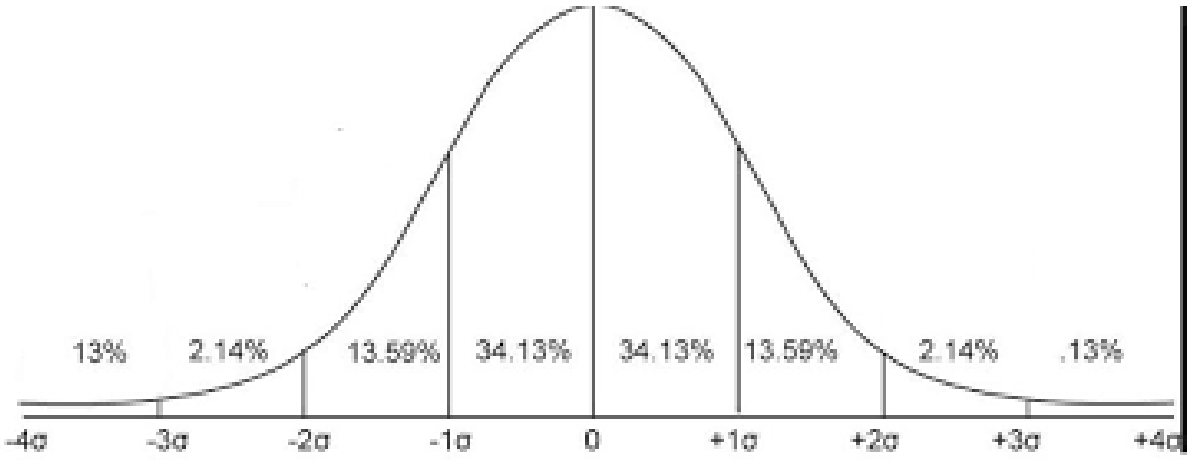
0,50

$P(+0{,}50<z) = 0{,}3085$

$P(173<x) = 0{,}3085$

$50{,}000(0{,}3085) = 15425$

c)



153 165 185

$Z_1 = \frac{153-165}{16} = -0{,}75$

P(-0,75<Z<0) = 0,2734

$Z_2 = \frac{185-165}{16} = 1{,}25$

P(0<Z<1,25) = 0,3944

P(-0,75<Z<1,25) = 0,2734+0,3944 = 0,6678

d)

Bulmak için 157 177 arasını bulup toplam değer olan 1'den çıkartırız.

$Z_1 = \frac{157-165}{16} = -0{,}50$

P(-0,50<Z<0) = 0,3085

$Z_2 = \frac{177-165}{16} = 0{,}75$

P(0<Z<0,75) = 0,2734

P(-0,50<Z<0,75) = 0,2734+0,3085 = 0,6819

İstenilen Alan = 1-0,6819 = 0,3181

İstatistiksel Tahmin:

Tahmin “^” ile gösterilir.

$\mu_1 = \bar{x}$ π = p (örneklem oranı)

$\mu_1 - \mu_2 = \bar{x}_1 - \bar{x}_2$

$\pi_1 - \pi_2 = p_1 - p_2$

Nokta Tahmini:

1) Sapmasızlık (Yansızlık):
2) Tutarlılık
3) Etkinlik
4) Yeterlilik

“$E(T) = \theta$”

1)Yansızlık:

Eğer bir istatistiğin beklenen değeri tahmin edilmek istenen anakütle parametresine eşitse söz konusu istatistik anakütle parametresinin sapmasız ya da yansız bir tahminleyicisidir.

2)Tutarlılık:

Örnek genişliği verilen bir sayıdan daha büyük alındığında tahmin ile parametre arasındaki farkın düşünülebilen en küçük pozitif bir sayıdan daha küçük kalma olasılığı 1 ise, o tahmin tutarlıdır denir.

3)Etkinlik:

Sapmasız ve tutarlı tahmin ediciler içinde minimum varyansa sahip olan tahmin edici etkindir.

4)Yeterlilik:

Bu bölümde bir amaç parametresi hakkında bir örneklemin içinde bulundurduğu bütün bilgileri özetleyen istatistiklerin bulunması için metotlar verilecektir. Bu tür istatistiklerin yeterlilik özelliğine sahip oldukları söylenir. Daha sade bir biçimde yeterli istatistikler olarak isimlendirilirler. Yeterli istatistikler, bütün yansız tahmin ediciler arasından en küçük varyanslı olanının belirlenmesi için kullanılır

Anakütle Parametresinin Aralık Tahmini:

Anakütle ortalamasının tahmini:

Güven Düzeyi:

%95 ise anlamlılık düzeyi %5’tir

%99 ise anlamlılık düzeyi %1’dir

Tahmin doğruluğundan ne kadar emin olduğunu belirten “Güven Düzeyi (Olasılık Düzeyi)” güven sınırlarının belirlenmesinde işe yarar.

Aralık tahmininin söz gelimi %95 güvenle yapılmış olması ana kütleden çekilecek her biri n hacimli çok sayıdaki örneklemden hesaplanan her 100 istatistikten 95’inin güven sınırları içinde kalacağı 5’inin ise dışına çıkacağı anlamını taşır.

P(Alt S. < μ < Üst S.) = 1 – α = Güven Düzeyi

α = anlamlılık düzeyi

%95 → $\bar{x} + (1{,}96).S_X$ Üst Sınır

%95 → $\bar{x} - (1{,}96).S_X$ Alt Sınır

%99 → $\bar{x} + (2{,}58).S_X$ Üst Sınır

%99 → $\bar{x} - (2{,}58).S_X$ Alt Sınır

$S_X =$ Ortalamanın Standart Sapması

S = Örneklem Standart Sapması

n/N ≥ 0,05 seçim iadesiz ise

$\sigma_x \rightarrow S_X = \frac{S}{\sqrt{n}} . \sqrt{\frac{N-n}{N-1}}$

n/N < 0,05 seçim iadeli ise

$\sigma_x \rightarrow S_X = \frac{S}{\sqrt{n}}$

Örnek:

Bir elektrik dağıtım şirketi 10.000 abonesinin aylık ortalama tüketim maliyetinin belirlemek amacıyla 625 abonesini rassal olarak seçmiştir. Bunlar için hesaplanan aylık ortalama tüketim 120 kwh ana kütlenin aylık ortalama tüketiminin güven sınırlarını %95 olasılıkla belirleyin.

$1 - \alpha = 0,95$

$N = 10.000$

$n = 625$

$\bar{x} = 120$ kwh

$S = 40$ kwh

$n/N = \frac{625}{10.000} = 0,0625 > 0,05$

$S_X = \frac{40}{\sqrt{625}} . \sqrt{\frac{10.000-625}{10.000-1}} = 1,55$ kwh

$\bar{x} + S_X\,(1,96) \rightarrow$ üst sınır

$\bar{x} - S_X\,(1,96) \rightarrow$ alt sınır

$120 + (1,55).(1,96) = 123,04$

$120 - (1,55).(1,96) = 116,96$

$P(116,96 < \mu < 123,04) = 0,95$ Güven Düzeyi

Yorum:

Anakütlede aylık ortalama elektrik tüketimi %95 olasılıkla 116,96 – 123,04 kwh aralığındadır.

Örnek:

Korona virüs hastalığından şikayetçi olarak bir hastaneye kabul edilenlerin ortalama olarak kaç gün kaldıkları tahmin edilmek isteniyor. Rassal olarak seçilen 400 hastanın ortalama kalış süresi 6 gün ve standart sapması 2.2 gündür. Söz konusu hastalıktan ortalama kalış süresinin güven sınırları %95 olasılıkla belirleyiniz.

$\alpha = 0,05$

$1 - \alpha = 0,95$

$n = 400$

$\bar{x} = 6$

$S = 2.2$

$S_X = \frac{S}{\sqrt{n}} = \frac{2.2}{\sqrt{400}} = 0,11$ gün

$6+(0,11).(1,96) = 6,22$

$6-(0,11).(1,96) = 5,78$

$P(5,78 < \mu < 6,22) = 0,95$ Güven Düzeyi

Yorum:

Anakütlede söz konusu hastalıktan ortalama kalış süresi %95 olasılıkla 5,78 – 6,22 gün arasındadır.

"$\mu_1 - \mu_2$" → anakütle ortalamaları arasındaki fark

$n_1 > 30$ $n_2 > 30$;

%95 $(\bar{x}_1 - \bar{x}_2) + 1{,}96 \cdot S_{X1-X2}$ %95 $(\bar{x}_1 - \bar{x}_2) - 1{,}96 \cdot S_{X1-X2}$

%99 $(\bar{x}_1 - \bar{x}_2) + 2{,}58 \cdot S_{X1-X2}$ %95 $(\bar{x}_1 - \bar{x}_2) - 2{,}58 \cdot S_{X1-X2}$

$$S_{X1-X2} = \sqrt{\frac{S1.S1}{n1} + \frac{S2.S2}{n2}}$$

S_{X1-X2} → ortalamalar arasındaki farkın standart hatası

S_1 = 1.örneklem standart sapması

S_2 = 2.örneklem standart sapması

n_1 = 1.örneklem hacmi

n_2 = 2.örneklem hacmi

Örnek:

İki farklı tür buğday tohumunun verimlerin, karşılaştırmak istenmektedir. X türü tohumun atıldığı rassal olarak seçilen 400 tarlada hektar başına ortalama 575 kg ürün alınmış ve standart sapması 140 kg olarak hesaplanmıştır. Y türü tohumun atıldığı rassal olarak seçilen 100 tarlada ise heltar başına ortalama 500 kg ürün alınmış. Standart sapması 110 kg olarak görülmüştür. Her iki tür buğday tohumunun verimleri arasındaki farkı %95 olasılıkla belirleyiniz.

Güven Düzeyi = 1 – α = 0,95

$n_1 = 400$

$n_2 = 100$ $\bar{x}_1 = 575$

$S_1 = 140$ $\bar{x}_2 = 500$

$S_2 = 110$ $S_{X1-X2} = \sqrt{\frac{140.140}{400} + \frac{110.110}{100}} = 13,04$

(575-500) + (1,96).(13,04) = 100,56 kg üst sınır

(575-500) - (1,96).(13,04) = 49,44 kg alt sınır

$P(49,44 < \mu_1 - \mu_2 < 100,56) = 0,95$

Yorum:

Her iki tür buğday tohumunun verimleri arasındaki fark %95 olasılıkla 49,44 – 100,56 kg arasındadır.

Tek Uçlu Hipotez Testi:

- $H_0 = \theta = \theta_1$ → Doğruluğunu kabul ettiğimiz hipotezdir.
  $H_1 = \theta > \theta_1$ → Doğruluğundan şürhe duyulan hipotezdir.

Bunlara tek taraflı hipotez testi ya da tek kuyruklu hipotez testi denir.

- $H_0 = \theta = \theta_1$
  $H_1 = \theta \neq \theta_1$

Bunlara ise çift taraflı hipotez testi ya da çift kuruklu hipotez testi denir.

Hipotez Testi:

Örneklem istatistiklerinden yararlanarak 0 hipotezinin (H_0) geçerli olup olmadığını bakılır.

0 Hipotezi:

H_0 ile gösterilir. Doğruluğuna inandığımız hipotezdir.

Alternatif Hipotez (Karşıt Hipotez):

H_1 ile gösterilir. Doğruluğundan kuşku duyulan hipotezdir. $_0$

Hipotez Testindeki Adımlar:

1) Hipotezlerin İfade Edilmesi:

Bizim için önemli olan alternatif hipotezin nasıl oluştuğudur.

$H_0 = \theta = \theta_0$ $\qquad$ $H_0 = \theta = \theta_0$

$H_1 = \theta > \theta_0$ $\qquad$ $H_1 = \theta \neq \theta_1$

$H_1 = \theta < \theta_0$

2) Anlamlılık Düzeyinin (α) Belirlenmesi:

$\alpha = 0{,}05$

$\alpha = 0{,}01$

Bir hipotez testinde 0,05 anlamlılık düzeyinin şeçilmesi doğru bir kararın %95 güvenle verilmiş olmasından emin olunabileceği anlamını taşır. Bu durumda 0 Hipotezini aslında doğru olduğu halde reddetmeyi 100 defada 5 defa hoş görmüş olmaktayız.

Anlamlılık Düzeyi: 1.tip hatayı işleme riskine razı olabileceğimiz maksimum olasılığa anlamlılık düzeyi adı verilir.

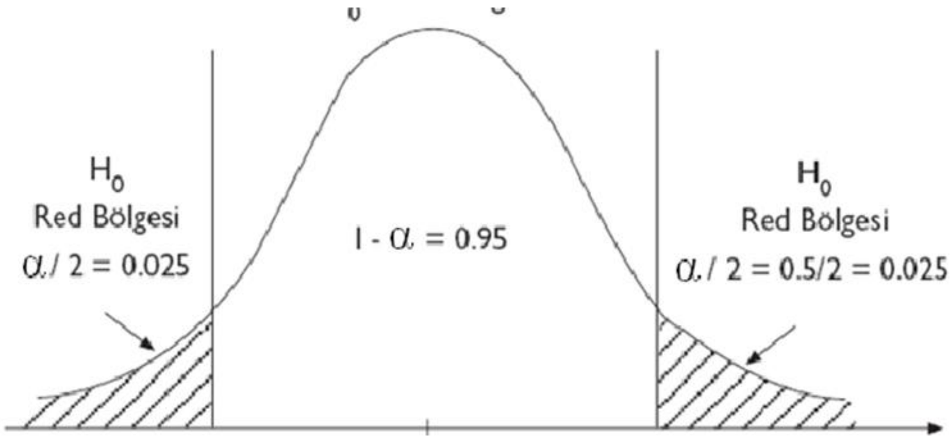
3) Örneklem Hacminin Belrilenmesi:

4) 0 Hipotezinin Red Bölgesinin Belirlenmesi:

→ Hipotez testi çiftine göre red bölgesi değişir.

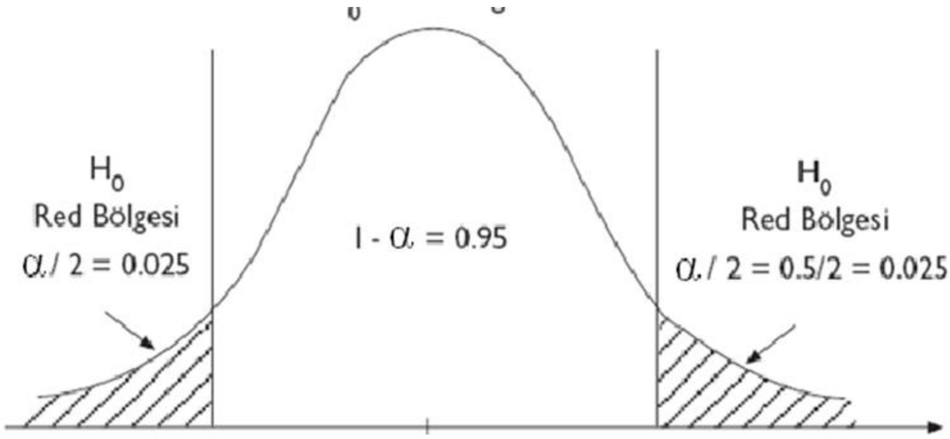
→ $H_0 = \theta = \theta_0$

$H_1 = \theta > \theta_0$ tek taraflı kuyruk testi



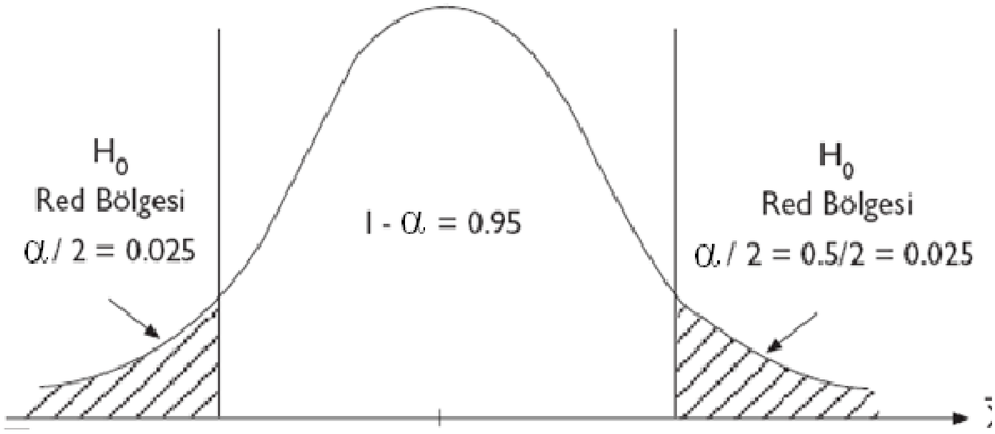
→ $H_0 = \theta = \theta_0$

$H_1 = \theta < \theta_0$ tek taraflı kuyruk testi



→ $H_0 = \theta = \theta_0$

$H_1 = \theta \neq \theta_1$ çift taraflı kuyruk testi



5) Örneklem İstatistiğinin Hesaplanması:

6) Test İstatistiğinin Hesaplanması:

$$Z = |\,(T - \theta_0) / \sigma_T|$$

7) İstatistiksel Kararın Alınması:

Hipotez Testi Çeşitleri:

1) Parametrik Hipotez Testleri

Z Testleri, T Testleri, F Testleri

2) Parametrik Olmayan Hipotez Testleri:

X^2 Testleri, Monn Testleri, V Testleri, Kruskal Wallis Testi

Tek Anakütle Parametreleriyle İlgili Hipotez Testi:

Test İstatistiği

$n \geq 30$

$Z = Z_h = |\ (\bar{x} - \mu_0) / \sigma_X\ |$

$\sigma_X = \frac{\sigma}{\sqrt{n}}$

$S_X = \frac{S}{\sqrt{n}}$

μ_0 = anakütle bilinen ortalaması

$\bar{x}$ = örneklem ortalaması

σ_X = örneklem ortalamasının standart sapması

n = örneklem hacmi

Not:

- $Z_h < Z_k \rightarrow H_0$ Hipotezi kabul edilir.
- $Z_h > Z_k \rightarrow H_0$ Hipotezi reddedilir.

Z_K Değer Tablosu:

Z_K	Anlamlılık Düzeyi = 0,05	Anlamlılık Düzeyi = 0,01
Tek Taraflı Kuyruk T.	1,64	2,33
Çift Taraflı Kuyruk T.	1,96	2,58

Örnek:

Tatil amacıyla ülkemize gelen yabancıların ortalma konaklama sürelerinin 15 günden farklı olduğu iddia edilmektedir. Rassal olarak seçilen 100 yabancı turistin ortalama konaklama süresi 15,5 gün ve standart sapması 1,5 gündür. %1 anlamlılık düzeyine göre karar veriniz.

$H_0 = \mu = 15$ gün

$H_1 = \mu \neq 15$ gün

$\alpha = 0{,}01$

$GD = 1 - \alpha = 0{,}99$

$\bar{x} = 15{,}5$ gün

$S = 1{,}5$ gün

$Z_K = 2{,}58$

$Z_h > Z_K$ kontrolü yapalım

$S_X = \frac{S}{\sqrt{n}} = \frac{1{,}5}{\sqrt{100}} = 0{,}15$

$Z_H = |\ (15{,}5 - 15) / 0{,}15\ | = 3{,}33$

$Z_h > Z_K$ sağlandığı için H_0 reddedilir.

Yorum:

Tatil amacıyla ülkemize gelen yabancıların ortalama kalma süresi 15 günden farklıdır.

Örnek:

A ve B şehirleri arasında yapılan telefon görüşmelerinin 6 dakika fazla olduğu iddia edilmektedir. Şehirler arası servisine gidilip 144 telefon konuşması rassal olarak seçildikten sonra bu konuşmaların ortalama süresinin 6,3 dakika ve standart sapmasının 2,64 dakika olduğu hesaplanmıştır. %5 anlamlılık düzeyine göre karar veriniz.

$H_0 = \mu = 6$ dakika

$H_1 = \mu > 6$ dakika

$\alpha = 0,05$

$GD = 1 - \alpha = 0,95$

$n = 144$

$\bar{x} = 6,3$ dakika

$S = 2,64$ dakika

$Z_K = 1,64$

$Z_h > Z_K$ kontrolü yapalım

$S_X = \frac{S}{\sqrt{n}} = \frac{2,64}{\sqrt{144}} = 0,22$ dakika

$Z_H = |\ (6,3 - 6) / 0,22\ | = 1,36$

$Z_h < Z_K$ sağlandığı için H_0 kabul edilebilir.

Yorum:

A ve B şehirleri arasında yapılan telefon görüşmeleri ortalama 6 dakikadan fazla sürmektedir.

Örnek:

XY firmasının ithal ettiği yeni makinaların parça başına ortalama üretim süresinin 25 dakika altına düşürdüğü iddia edilmektedir. Rassal olarak seçilen 121 işçinin yeni makinalarla yaptıkları üretimde ortalama sürenin 22 dakika ve standart sapmanın 6,6 dakika olduğu belirlenmiştir. %1 anlamlılık düzeyine göre karar veriniz.

$H_0 = \mu = 25$ dakika

$H_1 = \mu < 25$ dakika

$\alpha = 0{,}01$

$GD = 1 - \alpha = 0{,}99$

$n = 121$

$\bar{x} = 22$ dakika

$S = 6{,}6$ dakika

$Z_K = 2{,}33$

$Z_h > Z_K$ kontrolü yapalım

$S_X = \frac{S}{\sqrt{n}} = \frac{6{,}6}{\sqrt{121}} = 0{,}6$ dakika

$Z_H = |\ (22 - 25) / 0{,}6\ | = 5$

$Z_h > Z_K$ sağlandığı için H_0 reddedilir.

Yorum:

XY firmasında ithal edilen makine parça başına ortalama üretim süresi 25 günden daha az sürmektedir.

Anakütle Oranına İlişkin Testler:

Test İstatistiği

$n \geq 30$

$Z = Z_h = | (P - \pi_0) / \sigma_P |$

$$\sigma_P = \sqrt{\frac{\pi 0(1-\pi 0)}{n}}$$

$P = a / n$

π_0 = Anakütlenin bilinen oranı

a = Örneklemde ilgilenilen türden birim sayısı

P = Örneklem oranı

σ_P = Oranın Standart Hatası

Not:

$Z_h < Z_K$ olursa H_0 kabul edilir, H_1 reddedilir.

$Z_h > Z_K$ olursa H_0 reddedilir, H_1 kabul edilir.

Örnek:

Bir fakültenin belli bir bölüme isteyerek giren öğrencilerin oranı %40'tan farklı olduğu iddia edilmektedir. Bölüm öğrencileri arasından seçilen 250 öğrenciden 110'unun bölüme isteyerek girdiği belirlenmiştir. %5 anlamlılık düzeyine göre karar veriniz.

$H_0 = \pi = 0,40$

$H_1 = \pi \neq 0,40$

$\alpha = 0,05$

$n = 250$

$a = 110$

$Z_K = 1,96$

$\pi_0 = 0,40$

$1 - \pi_0 = 0,60$

$$\sigma_P = \sqrt{\frac{\frac{4}{10} \cdot \frac{6}{10}}{250}} = \sqrt{\frac{24}{25000}} = 0,03$$

$$P = a / n = \frac{110}{250} = 0,44$$

$$Z_H = |\ (0,44 - 0,40) / 0,03\ | = \frac{0,04}{0,03} = 1,33$$

$Z_h < Z_K$ olduğu için H_0 reddedilemez. H_1 reddedilir.

Karar:

Söz konusu bölüme isteyerek giren öğrenci sayısı %40'tan farklı değildir.

Örnek:

H şehrinde satılan ekmeklerden belediyenin belirlediği gramajın altında olan oranın %20'den fazla olduğu iddia edilmektedir. Rassal olarak seçilen 100 ekmekten 32'sinin belirlenen gramajın altında olduğu belirlenmiştir. 0,01 anlamlılık düzeyine göre karar veriniz.

$H_0 = \pi = 0,20$

$H_1 = \pi > 0,20$

$\alpha = 0,01$

$n = 100$

$a = 32$

$Z_K = 2,33$

$\pi_0 = 0,20$

$1 - \pi_0 = 0,80$

$$\sigma_P = \sqrt{\frac{\frac{2}{10} \cdot \frac{8}{10}}{100}} = \sqrt{\frac{16}{10000}} = 0,04$$

$$P = a / n = \frac{32}{100} = 0,32$$

$$Z_H = |\ (0,32 - 0,20) / 0,04\ | = \frac{0,12}{0,04} = 3$$

$Z_h > Z_K$ olduğu için H_0 reddedilir. H_1 kabul edilir.

Karar:

H şeklinde satılan ekmeklerde belediyenin belirlediği gramajın altında olanların oranı %20'den fazladır.

Örnek:

Büyük bir şehrimizde köprü geçişlerinde OGS yararlananların oranının %60'tan az olduğu iddia edilmektedir. Rassal olarak seçilen 200 sürücüden 110'unun OGS'den yararlandığı belirlenmiştir. %5 anlamlılık düzeyine göre karar veriniz.

$H_0 = \pi = 0{,}60$

$H_1 = \pi < 0{,}60$

$\alpha = 0{,}05$

$n = 200$

$a = 110$

$Z_K = 1{,}64$

$\pi_0 = 0{,}60$

$1 - \pi_0 = 0{,}40$

$$\sigma_P = \sqrt{\frac{\frac{6}{10} . \frac{4}{10}}{200}} = \sqrt{\frac{24}{20000}} = 0{,}034$$

$$P = a / n = \frac{110}{200} = 0{,}55$$

$$Z_h = |\ (0{,}55 - 0{,}60) / 0{,}034\ | = \frac{0{,}050}{0{,}034} = 1{,}47$$

$Z_h < Z_K$ olduğu için H_0 kabul edilir. H_1 reddedilir.

Karar:

Seçilen şehirde köprü geçişlerinde OGS'den yararlananların oranı %60'tan az değildir.

İKİ ANAKÜTLENİN PARAMETRELERİ İLE İLGİLİ HİPOTEZ TESTLERİ:

Anakütle Ortalamaları Arasındaki Farkla İlişkin Testler:

$Z = | (\bar{x}_1 - \bar{x}_2) / (\sigma_{X1-X2}) |$

$$\sigma_{X1-X2} = \sqrt{\frac{(\sigma 1).(\sigma 1)}{n1} + \frac{(\sigma 2).(\sigma 2)}{n2}}$$

$$S_{X1-X2} = \sqrt{\frac{S1.S1}{n1} + \frac{S2.S2}{n2}}$$

S_1 = 1.örneklem standart sapması

S_2 = 2.örneklem standart sapması

n_1 = 1.örneklem hacmi

n_2 = 2.örneklem hacmi

Örnek:

Bir fakültenin bilgisayar mühendisliği bölümündeki erkek öğrencilerin derse devamsızlıklarının kız öğrencilerinkinden farklı olduğu idiia edilmektedir. Rassal olarak söz konusu bölümden seçilen 60 erkek öğrencinin devamsızlık 30 gün/dönem ve standart sapması 9 gün/dönem; 40 kız öğrencinin devamsızlık ortalamsı 24 gün/dönem ve standart sapması 6 gün/dönem olarak hesaplanmıştır. 0,01 anlamlılık düzeyine göre karar veriniz.

$H_0 = \mu_1 = \mu_2$

$H_0 = \mu_1 \neq \mu_2$

$n_1 = 60 \quad n_2 = 40$

$\bar{x}_1 = 30$ gün/dönem

$\bar{x}_2 = 24$ gün/dönem

$S_1 = 9$ gün/dönem

$S_2 = 6$ gün/dönem

$Z_K = 2{,}58$

$$S_{X1-X2} = \sqrt{\frac{81}{60} + \frac{36}{40}} = 1{,}5$$

$Z_h = |\,(\bar{x}_1 - \bar{x}_2) / S_{X1-X2}\,| = 4$

$Z_h > Z_K$ olduğu için H_0 reddedilir, H_1 kabul edilir.

Karar:

Söz konusu bölümdeki erkek öğrencilerin derse devamsızlıkları kız öğrencilerinkinden farklıdır.

Örnek:

Bir otomobil fabrikasının ürettiği otomobiller için satın alacağı A marka lastiklerin B marka lastiklerden daha dayanaklı olduğu iddia edilmektedir. Rassal olarak seçilen A marka 125 lastiğin ortalaması ömür süresi 502000 km ve standart sapması 2200 km, B marka 150 lastiğin ortalama ömür süresi 49800 km ve standart sapma 2400 km'dir. %5 anlamlılık düzeyine göre karar veriniz.

$H_0 = \mu_1 = \mu_2$

$H_0 = \mu_1 > \mu_2$

$\alpha = 0{,}05$

$n_1 = 125 \quad n_2 = 150$

$\bar{x}_1 = 50200$ km

$\bar{x}_2 = 49800$ km

$S_1 = 2200$

$S_2 = 2400$

$Z_K = 1{,}64$

$$S_{X1\text{-}X2} = \sqrt{\frac{2200.2200}{125} + \frac{2400.2400}{150}} = 277{,}70 \text{ km}$$

$Z_h = |\ 400 / 277{,}70\ | = 1{,}44$

$Z_h < Z_K$ olduğu için H_0 kabul edilir, H_1 reddedilir.

Karar:

A marka lastikler B marka lastiklerden daha dayanaklı değildir.

Örnek:

Aynı sektörde çalışan yabancı firmalardan PR işçilerinin ST işçilerinden daha az ücret aldıkları iddia edilmektedir. Rassal olarak seçilen 100 PR işçisinin haftalık ortalama ücreti 87,5 $ ve standart sapması 22,5 $, 125 ST işçisinin haftalık ortalama ücreti 97,5 $ ve standart sapması 27,5 $'dır. %1 anlamlılık düzeyine göre karar veriniz.

$H_0 = \mu_1 = \mu_2$

$H_0 = \mu_1 < \mu_2$

$\alpha = 0,01$

$n_1 = 100 \quad n_2 = 125$

$\bar{x}_1 = 87,5$

$\bar{x}_2 = 97,5$

$S_1 = 22,5$

$S_2 = 27,5$

$Z_K = 2,33$

$$S_{X1-X2} = \sqrt{\frac{22,5.22,5}{100} + \frac{27,5.27,5}{125}} = 3,33$$

$Z_h = |\ -10 / 3,33\ | = 3$

$Z_h > Z_K$ olduğu için H_0 reddedilir, H_1 abul edilir.

Karar:

Söz konusu firmada PR işçilerinin ST işçilerinden aldıkları ücret daha azdır.

Anakütle Oranları Arasındaki Farka İlişkin Testler:

Test İstatistiği

$n_1 \geq 30$ $n_2 \geq 30$

$Z_h = | (P_1 - P_2) / S_{P1-P2} |$

$S_{P1-P2} = \sqrt{\frac{P1(1-P1)}{n1} + \frac{P2(1-P2)}{n2}}$

$P_1 = \frac{a1}{n1}$

$P_2 = \frac{a2}{n2}$

P_1 = 1.örneklem oranı

P_2 = 2.örneklem oranı

Örnek:

İstatislik dersini aynı öğretim üyesinden alan A ve B bölümleri öğrencilerinin başarı oranlarının farklı olduğu iddia edilmektedir. Rassal olarak A bölümden şeçilen 40 öğrenciden 30'unun ve B bölümünden seçilen 50 öğrenciden 35'inin başarılı olduğu biliniyor. %5 anlamlılık düzeyine göre karar veriniz.

$H_0 = \pi_1 = \pi_2$

$H_1 = \pi_1 \neq \pi_2$

$\alpha = 0,05$

$n_1 = 40 \quad n_2 = 50$

$a_1 = 30 \quad a_2 = 35$

$Z_K = 1,96$

$P_1 = \frac{30}{40} = 0,75$

$P_2 = \frac{35}{50} = 0,70$

$S_{P1-P2} = \sqrt{\frac{0,75.(1-0,75)}{40} + \frac{0,70.(1-0,70)}{50}} = 0,094$

$Z_h = |\ (0,75-0,70)\ /\ 0,094\ | = \frac{0,05}{0,094} = 0,53$

$Z_h < Z_K$ olduğu için H_0 reddedilemez, H_1 reddedilir.

Karar:

İstatistik dersini aynı öğretim üyesinden alan A ve B bölümleri öğrencilerinin başarı oranları farklı değildir.

Örnek:

Televizyonda futbol maçlarının daha sık yayınlanmasını isteyen erkeklerin oranının kadınlarınkinden yüksek olduğu iddia edilmektedir. Rassal olarak seçilen 100 erkekten 60'ının ve 140 kadından 49'unun daha sık yayından yana oldukları belirlenmiştir. %1 anlamlılık düzeyine göre karar veriniz.

$H_0 = \pi_1 = \pi_2$

$H_1 = \pi_1 > \pi_2$

$\alpha = 0{,}01$

$n_1 = 100 \quad n_2 = 140$

$a_1 = 60 \quad a_2 = 49$

$Z_K = 2{,}33$

$P_1 = \frac{60}{100} = 0{,}6$

$P_2 = \frac{49}{140} = 0{,}35$

$S_{P1-P2} = \sqrt{\frac{0{,}6.(1-0{,}6)}{100} + \frac{0{,}35.(1-0{,}35)}{140}} = 0{,}063$

$Z_h = |\ (0{,}6 - 0{,}35) / 0{,}063\ | = \frac{0{,}25}{0{,}063} = 3{,}97$

$Z_h > Z_K$ olduğu için H_0 reddedilir, H_1 reddedilemez.

Karar:

Televizyondaki futbol maçlarının daha sık yayınlanmasını isteyen erkeklerin oranı kadınların oranından yüksektir.

Örnek:

Klasik lise mezunlarının üniversiteye girme oranının Anadolu lisesi mezunlarınınkinden düşük olduğu iddia edilmektedir. Rassal olarak seçilen 50 Klasik Lise mezunun 15'inin ve 60 Anadolu lisesi mezununun 24'ünün üniversiteye girdikleri belirlenmiştir. %5 anlamlılık düzeyine göre karar veriniz.

$H_0 = \pi_1 = \pi_2$

$H_1 = \pi_1 < \pi_2$

$\alpha = 0{,}05$

$n_1 = 50 \quad n_2 = 60$

$a_1 = 15 \quad a_2 = 24$

$Z_K = 1{,}64$

$P_1 = \frac{15}{50} = 0{,}3$

$P_2 = \frac{24}{60} = 0{,}4$

$S_{P1-P2} = \sqrt{\frac{0{,}3.(1-0{,}3)}{50} + \frac{0{,}4.(1-0{,}4)}{60}} = 0{,}090$

$Z_h = |\ (0{,}3 - 0{,}4) / 0{,}090\ | = \frac{0{,}1}{0{,}090} = 1{,}1$

$Z_h < Z_K$ olduğu için H_0 reddedilemez, H_1 reddedilir.

Karar:

Klasik lise mezunlarının üniversiteye girme oranının Anadolu lisesi mezunlarınınkinden düşük değildir.

Anakütle Oranının Tahmini:

Alt Sınır → $P - Z.S_P$

Üst Sınır → $P + Z.S_P$

$S_P = \sqrt{\frac{P.(1-P)}{n}}$ n/N < 0,05 ise bu formül kullanılır

$S_P = \sqrt{\frac{P.(1-P)}{n}}$ n/N > 0,05 ise bu formül kullanılır

P (Alt Sınır < π < Üst Sınır) = 1 – α

Not:

Z değeri:

%95 → 1,96

%99 → 2,58

Örnek:

Bir şehirdeki 60.000 genç erkekten yüzde kaçının düzenli olarak spor yaptığı belirlenmek isteniyor ve bu amaçla rassal olarak 4.000 genç erkekten oluşan bir örneklem seçilip bunlardam 1.200'ünün düzenli olarak spor yaptığı anlaşılıyor. Anakütlede düzenli olarak spor yapanların oranının güven sınırlarını %99 olasılıkla belirleyiniz.

$1 - \alpha = 0{,}99$

$N = 60.000$

$n = 4.000$

$a = 1.200$

$n/N = \frac{4.000}{60.000} = 0{,}067 > 0{,}05$

$P = \frac{1200}{4000} = 0{,}3$

$S_P = \sqrt{\frac{0{,}3.(1-0{,}3)}{4{,}000}} = 0{,}007$

Alt sınır → $P - Z.S_P = 0{,}3 - 2{,}58.0{,}007 = 0{,}2819$

Üst sınır → $P + Z.S_P = 0{,}3 + 2{,}58.0{,}007 = 0{,}3181$

$P\ (0{,}2819 < \pi < 0{,}3181) \rightarrow 1 - \alpha$

Yorum:

Anakütlede düzenli olarak spor yapanların oranı %99 olasılıkla %28,19 ile %31,81 arasındadır.

Örnek:

Bir şehirde oturan aileler arasından rassal 200 ailelik bir örneklem seçilmiştir. Bunlardan 40'ının aylık gelirinin yüksek düzeyde olduğu anlaşılmıştır. Anakütlede aylık gelirleri yüksek düzeyde olanların güven sınırlarını %99 olasılıkla belirleyiniz.

$1 - \alpha = 0{,}99$

$N = 200$

$n = 40$

$a = 1.200$

$n/N = \frac{4.000}{60.000} = 0{,}067 > 0{,}05$

$P = \frac{40}{200} = 0{,}2$

$S_P = \sqrt{\frac{0{,}2.(1-0{,}2)}{200}} = 0{,}028$

Alt sınır → $P - Z.S_P = 0{,}2 - 2{,}58.0{,}028 = 0{,}1278$

Üst sınır → $P + Z.S_P = 0{,}2 + 2{,}58.0{,}028 = 0{,}2722$

$P\ (0{,}1278 < \pi < 0{,}2722) \rightarrow 1 - \alpha$

Yorum:

Anakütlede aylık geliri yüksek düzeyde olanların oranı %99 olasılıkla %12,78 ile %27,22 arasındadır.

Anakütle Oranları Arasındaki Farkın Tahmini:

$S_{P1-P2} = \sqrt{\frac{P1(1-P1)}{n1} + \frac{P2(1-P2)}{n2}}$

$P_1 = a_1 / n_1$

$P_2 = a_2 / n_2$

Alt Sınır → $(P_1 – P_2) – Z.S_{P1-P2}$

Üst Sınır → $(P_1 – P_2) + Z.S_{P1-P2}$

P(Alt sınır < $\pi_1 - \pi_2$ < Üst sınır) = $1 – \alpha$

Not:

Z değeri:

%95 → 1,96

%99 → 2,58

Örnek:

Aynı dersi veren ik öğretim üyesinin dönem sonu başarıları karşılaştırılmak istenmektedir. Söz konusu öğretim üyelerinden A'nın 500 öğrenciden 275'i ve B'nin ders verdiği 400 öğrenciden 300'ü dönem sonu sınavlarında başarılı olmuştur. Öğretim üyelerinin dönem sonu başarı oranları arasındaki farkın güven sınırlarını %99 olasılıkla belirleyiniz.

$1 - \alpha = 0{,}99$

$n_1 = 500$

$n_2 = 400$

$a_1 = 275$

$a_2 = 3000$

$P_1 = \frac{275}{500} = 0{,}55$

$P_2 = \frac{300}{400} = 0{,}75$

$S_{P1-P2} = \sqrt{\frac{0{,}55(1-0{,}55)}{500} + \frac{0{,}75(1-0{,}75)}{400}} = 0{,}031$

Alt Sınır → (0,55-0,75) – 2,58.0,031 = -0,28

Üst Sınır → (0,55-0,75) + 2,58.0,031 = -0,12

$P\,(-0{,}28 < \pi_1 - \pi_2 < -0{,}12) = 1 - \alpha = 0{,}99$

Yorum:

Söz konusu Öğretim üyelerinden A'nın başarı oranı B'nin başarı oranından daha düşüktür.

Küçük Örneklem için T Dağılımı:

n < 30 durumlarında kullanılır.

$$T = \frac{X - \mu}{S}$$

V = n – 1; selbestlik derecesi

Selbeslik Derecesi: Anakütle de parametrelerinin bilinmediği durumlarda bunları örneklem örneklem istatistiklerine dayanarak tahmin ederiz. Bu tahminleri sapmasız bir biçimde yapabilmek için örneklem hacminin herkesçe kullanılabilen kısmına selbestlik derecesi adı verilir.

Anakütle ortalama aralık tahmini:

$\bar{x} + T_{tablo}.S_X$ → üst sınır

$\bar{x} - T_{tablo}.S_X$ → alt sınır

$$S_X = \frac{S}{\sqrt{n-1}}$$

S_X = Örneklem ortalaması standart hatası

qSS

Örnek:

Bir su dağıtım şirketinin abonelerinin aylık ortalama su tüketimini belirlemek amacıyla 10 abonesini rassal olarak seçmiştir. Bunlar için hesaplanan aylık ortalam tüketim 15 m^3, standart sapması 4,5 m^3'tür. Anakütlenin aylık ortalama su tüketiminin güven sınırlarını %95 olasılıkla belirleyiniz.

$\bar{x} = 15\ m^3$

$S = 4,5\ m^3$

s.d.; $V = n - 1 = 10 - 1 = 9$

$T_{tablo}\ (1 - \alpha;\ 9) = 2,62$

$S_X = \frac{S}{\sqrt{n-1}} = \frac{4,5}{3} = 1,5\ m^3$

$X + T.S_X = 15 + 2,62.1,5 = 18,93 \rightarrow$ üst sınır

$X - T.S_X = 15 - 2,62.1,5 = 11,07 \rightarrow$ alt sınır

$P(11,07 < \mu < 18,93) = 0,95 = \alpha$

$\mu_1 - \mu_2$ için aralık tahmini:

$n_1 < 30\ n_2 < 30$ olduğu durumlarda kullanılır.

$(\bar{x}_1 - \bar{x}_2) + T_{tablo}.S_{X1-X2} \rightarrow$ üst sınır

$(\bar{x}_1 - \bar{x}_2) - T_{tablo}.S_{X1-X2} \rightarrow$ alt sınır

$$S_{X1-X2} = \sqrt{\frac{n1.S1.S1+n2.S2.S2}{n1+n2-2}} \cdot \sqrt{\frac{1}{n1} + \frac{1}{n2}}$$

Örnek:

İki ayrı firma tarafından üretilen akülerin ortalama ömür süreleri karşılaştırılmak istenmektedir. Rassal olarak seçilen A marka aküden örneklem hacmi 15, örneklem ortalamsı 1830, standart sapması 122; B marka akünün örneklem hacmi 11, örneklem ortalamsı 1560, standart sapması, 108'dir. A ve B marka aküye ilişkin ömür sürelerini 0,99 güven düzeyine göre inceleyiniz.

s.d. → 11+15-2 = 24

T_{tablo} (0,99; 24)

$$S_{X1-X2} = \sqrt{\frac{15.122.122+11.108.108}{15+11-2}} \cdot \sqrt{\frac{1}{15}+\frac{1}{11}} = 48{,}04$$

(1830 - 1560) + 2,797.48,04 = 404,37

(1830 - 1560) - 2,797.48,04 = 135,63

$P(135{,}63 < \mu_1 - \mu_2 < 404{,}37) = 1 - \alpha = 0{,}99$

Yorum:

A marka akülerin ortalama ömür süresi B marka akülerden fazla olup ortalamalar arasındaki fark %99 olasılıkla ve A marka akülerin lehine 135,63 – 404,37 arasındadır.

T testinde ortalamaya ilişkin testler:

$T = |(\bar{x} - \mu_0) / S_X|$

$S_X = \frac{S}{\sqrt{n-1}}$

Örnek:

Standart bir bilim sınavını uygulayan bir okulda yıllardır 66 puanlık bir ortalama elde edildiği halde bu ortalamanın değiştiği iddia edilmektedir. Bu yıl öğrenciler arasından 26'sı rassal olarak seçilmiş ve bu standart sınava alınmıştır. Söz konusu öğrencilerin elde ettikleri not ortalaması 69 puan ve standart sapması 20 puandır. %5 anlamlılık düzeyine göre karar veriniz. ($T_T = 2{,}06$)

$H_0 = \mu_0 = 66$ puan

$H_1 = \mu_1 \neq 66$ puan

$\alpha = 0{,}05$

$n = 26$

$V = n - 1 = 26 - 1 = 25$

$\bar{x} = 69$

$S_X = \frac{10}{\sqrt{26-1}} = 2$

$T_h = |(69 - 66) / 2| = 1{,}5$

$T_h < T_T$ olduğundan H_0 reddedilemez, H_1 reddedilir.

Karar:

Söz konusu okulda uygulanan standart bilim sınavından elde edilen ortalama puan değişmemiştir.

Örnek:

Belli bir mesafenin bayan atletler tarafından 20 dakikanın üzerinde koşulduğu iddia edilmektedir. Rassal olarak seçilen 17 bayan atletin ortalama derecesi 20,9 dakika ve standart sapması 1,2 dakikadır. %1 anlamlılık düzeyine göre karar veriniz. (T_T = 2,583)

$H_0 = \mu_0 = 20$ dakika

$H_1 = \mu_1 > 20$ dakika

$\alpha = 0,01$

$n = 17$

$V = n - 1 = 17 - 1 = 16$

$\bar{x} = 20,9$ dakika

$S = 1,2$ dakika

$S_X = \frac{1,2}{\sqrt{17-1}} = 0,3$ dakika

$T_h = |(20,9 - 20) / 0,3| = 3$

$T_h > T_T$ olduğundan H_0 reddedilir, H_1 reddedilemez.

Karar:

Söz konusu mesafe bayan atletler tarafından 20 dakikanın üzerinde koşulmaktadır.

Örnek:

Bir paketleme makinesinin yaptığı paketlerin ağırlığının ortalama 1000 gram olması gerektiği halde piyasaya daha hafif paketlerin sürüldüğü iddia edilmektedir. Paketleme sırasında rassal olarak seçilen 10 paketin ortalama ağırlığı 940 gram ve standart sapması 120 gramdır. %5 anlamlılık düzeyine göre karar veriniz. (T_T = 1,833)

$H_0 = \mu_0 = 1000$ gram

$H_1 = \mu_1 < 1000$ gram

$\alpha = 0,05$

$n = 10$

$V = n - 1 = 10 - 1 = 9$

$\bar{x} = 940$ gram

$S = 120$ gram

$S_X = \frac{120}{\sqrt{10-1}} = 40$ gram

$T_h = |(940 - 1000) / 40| = 1,5$

$T_h < T_T$ olduğundan H_0 reddedilemez, H_1 reddedilir.

Karar:

Söz konusu paketleme makinesinin yaptığı ortalama paketleme 1000 gramdır.

T testi için ortalamalar arasındaki farka ilişkin testler:

$T = |(\bar{x}_1 - \bar{x}_2) / (S_{X1} - S_{X2})$

$$S_{X1} - S_{X2} = \sqrt{\frac{n1.S1.S1+n2.S2.S2}{n1+n2-2}} \cdot \sqrt{\frac{1}{n1} + \frac{1}{n2}}$$

Örnek:

UV ve YZ marka ampüllerin ömür sürelerinin farklı olduğu iddia edilmektedir. Rassal olarak seçilen UV marka 10 ampulün ortalama ömür süresi 850 saat ve standart sapması 100 saattir. YZ marka 15 ampulün ortalama ömür saati 650 ve standart sapması 150 saattir. %1 anlamlılık düzeyine göre karar veriniz.

$H_0 = \mu_1 = \mu_2$

$H_1 = \mu_1 \neq \mu_2$

$\alpha = 0{,}01$

$n_1 = 10; n_2 = 15$

$\bar{x}_1 = 850; \bar{x}_2 = 650$

$S_1 = 100; S_2 = 150$

$$S_{X1} - S_{X2} = \sqrt{\frac{10.100.100 + 15.150.150}{10+15-2}} \cdot \sqrt{\frac{1}{10} + \frac{1}{15}} = 56{,}31 \text{ saat}$$

$T = |(850 - 650) / 56{,}31| = 3{,}552$

$T > T_K$ olduğu için H_0 reddedilir, H_1 kabul edilir.

Karar:

UV ve YZ marka ampüllerin ömür süreleri farklıdır.

Örnek:

Aynı öğretim üyesinden olasılık ve istatistik dersi alan X bölümü öğrencilerinin Y bölümü öğrencilerinden daha başarılı olduğu iddia edilmektedir. Rassal olarak seçilen X bölümünden 10 öğrencinin not ortalaması 65 puan ve standart sapmsı 10 puandır; Y bölümden 11 öğrenin not ortalaması 45 puan ve standart sapmsı 30 puan olarak hesaplanmıştır. %5 anlamlılık düzeyine göre karar veriniz.

(T_T = 1,79)

$H_0 = \mu_1 = \mu_2$

$H_1 = \mu_1 > \mu_2$

$\alpha = 0{,}05$

$n_1 = 10;\ n_2 = 11$

$\bar{x}_1 = 65;\ \bar{x}_2 = 45$

$S_1 = 10;\ S_2 = 30$

$$S_{X1} - S_{X2} = \sqrt{\frac{10.10.10 + 11.30.30}{10+11-2}} \cdot \sqrt{\frac{1}{10} + \frac{1}{11}} = 16{,}12$$

$T = |(65 - 45) / 16{,}12| = 1{,}24$

$T < T_T$ olduğu için H_0 kabul edilir, H_1 reddedilir.

Karar:

X ve Y bölümü öğrencilerinden olasılık ve istatislik dersindeki başarıları birbirine eşittir.