HACETTEPE ÜNİVERSİTESİ

İSTATİSTİK BÖLÜMÜ

İST156 İSTATİSTİĞE GİRİŞ II

DERS 5- HİPOTEZ TESTLERİ ve KİTLE ORTALAMASI İÇİN HİPOTEZ TESTLERİ

Ders sorumluları: Prof.Dr.Serpil AKTAŞ ALTUNAY (01 Şubesi)
Doç.Dr. Ayten YİĞİTER (02 Şubesi)

HİPOTEZ TESTLERİNE GİRİŞ

Çıkarsama ne demek?

Örneklemden elde dilen bilgiler yardımıyla kitle parametreleri hakkında bir karara varma işlemine "çıkarsama (inference)" denir.

Hipotez ne demek ?

Doğruluğu henüz kanıtlanmamış iddialara (önermelere) "hipotez" denir. Kitle parametresine ilişkin iddialara "istatistiksel hipotez" denir.

Hipotez testi ne demek ?

Örneklem istatistiklerinden yararlanarak bir hipotezin geçerli olup olmadığını ortaya koyma işlemine istatistiksel "hipotez testi" veya denir.

Test edilmek istenen ifadenin yazıldığı hipoteze “Yokluk Hipotezi (Null Hypothesis)” denir ve H_0 ile gösterilir.

Yokluk hipotezine karşılık bir şeçenek olmalıdır o da “Seçenek Hipotezi (Alternative Hypothesis)” H_A, H_S ya da H_1 ile gösterilir.

Seçenek hipotezinin yönü $\neq$, > ya da < olarak ifade edilir.

Örneğin:

H_0: $\mu=5$ hipotezine karşılık

H_S: $\mu \neq 5$, H_S: $\mu > 5$, H_S: $\mu < 5$ seçenek hipotezlerinden birisi kurulur.

Bunun anlamı kitle ortalaması 5’den farklıdır; 5’den büyüktür; 5’den küçüktür. Bunlardan sadece bir tanesi seçilebilir.

Örneğin bir bölgede yaşayanların yaş ortalamasının 30 olduğu iddia edilsin. Bu iddia nasıl test edilecek? O zaman hipotezimiz **$H_0: \mu=30$** olacaktır.

KİTLE

ÖRNEKLEM ÇEK

ÖRNEKLEM

Eğer kitleyi temsil edecek şekilde doğru örnekleme yöntemi ile yeterli sayıda bir örneklem oluşturduysak, o zaman yaptığımız testte bulduğumuz sonuç kitlenin gerçek değerini tahmin edebilir.

H_0: μ=30 şeklinde hipotez kurduk (sadece iddia ediyoruz ve o bölgede yaşayanların tamamının yaş ortalamasını hiçbir şekilde bilmiyoruz) ve çektiğimiz örneklemdeki kişi sayılarının (n) yaş ortalaması $\bar{X} = 29{,}7$ çıktıysa o zaman kitledeki gerçek yaş ortalaması da 30'dur diyebilir miyiz?
Bu konuda elimizde yeterince kanıt var mı?

İşte bunun cevabını istatistiksel hipotez testleri ile veriyoruz.

Örneğin bir spor kulübündeki sporcuların kilolarının ortalama 76 kg olup olmadığı araştırılmak istensin,

H_0: μ=76 hipotezine karşılık

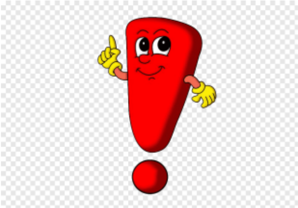
H_S: μ ≠ 76,

H_S: μ > 76 ,

Hs: μ < 76

seçenek hipotezlerinden birisi kurulur.

H_0: μ ≥ 76



H_S: μ < 76

<76 diye kurulan seçenek hipotezi, aslında yokluk hipotezinin ≥ olduğu anlamına gelir.

Anlamlılık düzeyi(significance level)

α : I.Tip hata ya da anlamlılık düzeyi (significance level) olarak bilinir. Hipotez testlerinde karşılaşabileceğimiz bir risk, gerçekte H_0 hipotezi doğru iken, hipotez test sonucunda H_0 hipotezini doğru olmadığını bulmaktır . Buna hata istatistikte I. Tip hata denir. Araştırmacılar bunu çalışmanın başında kontrol altında tutmak isterler bu nedenle uygulamada genelde α=0,01, 0,05 ya da 0,10 kabul edilir.

	Kitledeki Durum	
Test Sonucu	H_0 Doğru	H_0 Yanlış
H_0 Kabul	Doğru Karar	II.Tip Hata (β)
H_0 Ret	I.Tip Hata (α)	Doğru Karar

Güven aralığını hesaplarken 1-α'ı güven düzeyi olarak adlandırılır.

β : II.Tip Hata
1-β : Testin gücü (Power of test) olarak adlandırılır.

İstatistiksel hipotez testlerinin tümü H_0 hipotezinin doğru olduğu varsayımı altında gerçekleştirilir.

Bu hipotezler için olası dört durumu inceleyelim.

Örneğin bir e-posta koruma sistemi yazılıyor ve sistemdeki izinsiz gönderileri (spam mail) tespit etmek istiyoruz. Yazılımımız bu gönderileri tespit edip engelleyecek.

H_0: Gelen posta spam değildir.
H_s: Gelen posta spamdır.

Durum 1: Gelen posta spam değil ve bizim yazılım o postaya normal dediyse…………………………DOĞRU KARAR
Durum 2: Gelen posta spam değil ama bizim yazılım onu spam'a atıyorsa…………………………………I.TİP HATA
Durum 3: Gelen posta spam ve ve yazılım onu spam'a atıyorsa……………………………DOĞRU KARAR
Durum 4: Gelen posta spam ancak yazılım onu normal postaya atıyorsa…………………………………II.TİP HATA

Makine Öğrenmede nasıl kullanılır?

Karmaşa Matrisi (Confusion Matrix)

		Tahmin Edilen	
		Pozitif	**Negatif**
Gerçek Değer	**Pozitif**	DP	YN
	Negatif	YP	DN

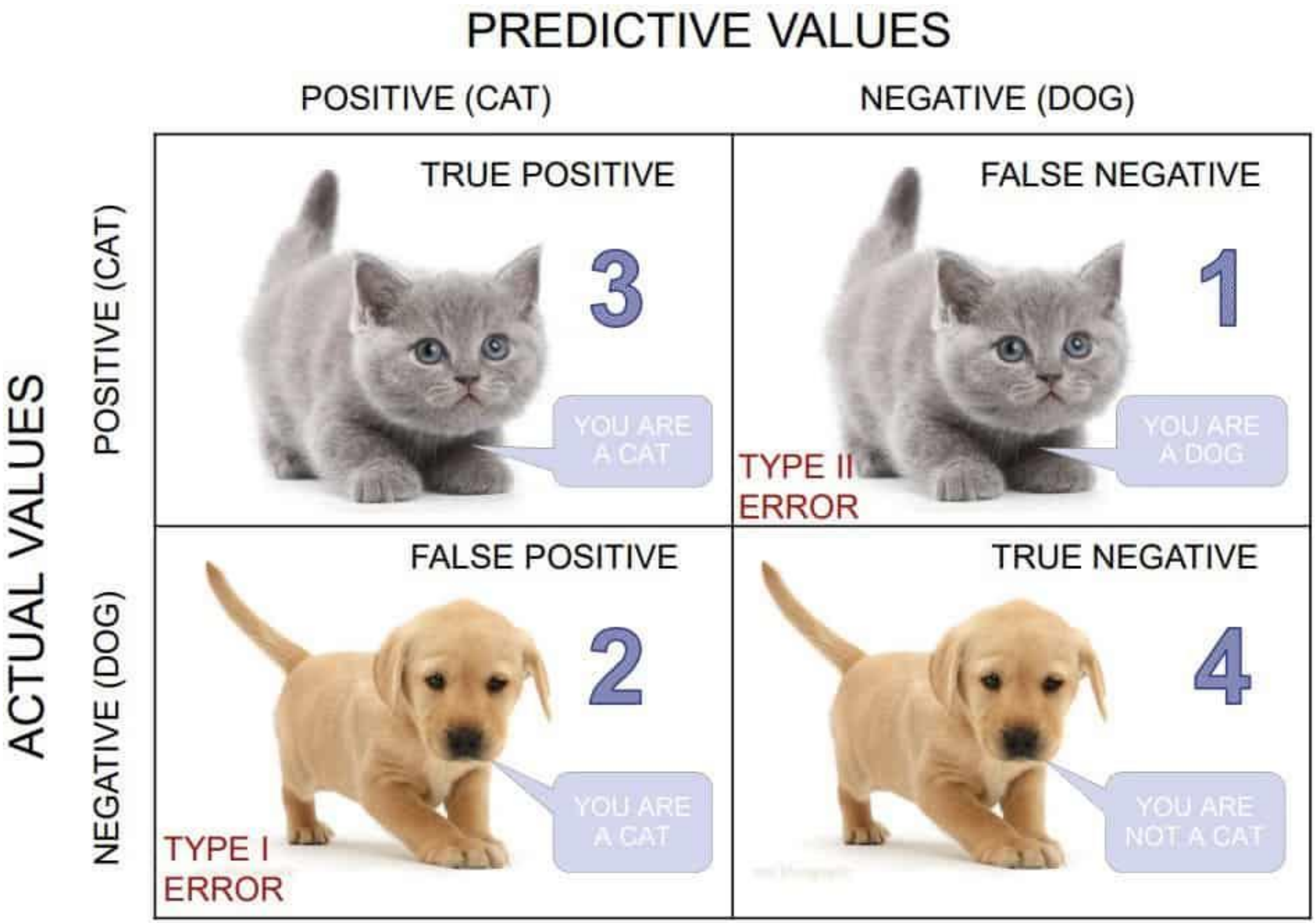
1. Doğruya doğru demek (True Positive – TP) DOĞRU
2. Yanlışa yanlış demek (True Negative – TN) DOĞRU
3. Doğruya yanlış demek (False Positive – FP) YANLIŞ
4. Yanlışa doğru demek(False Negative – FN) YANLIŞ

Gerçek Pozitif (DP): Bunlar gerçek değeri 1 ve tahmin edilen değerin de 1 olduğu durumdur.

Gerçek Negatif (DN): Bunlar gerçek değeri 0 ve tahmin edilen değerin de 0 olduğu durumdur.

Yanlış Pozitif (YP): Bunlar gerçek değeri 0 ancak tahmin edilen değeri 1 olan durumlardır.

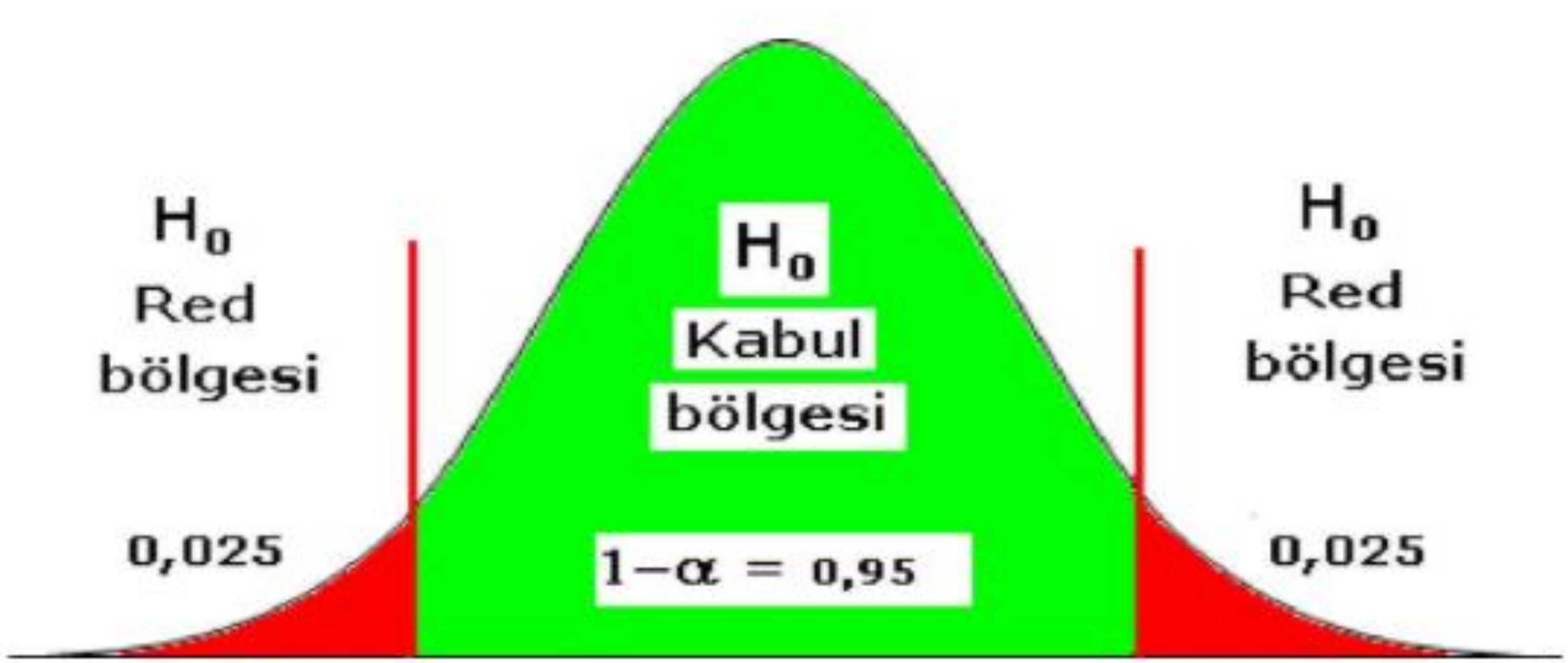
Yanlış Negatif (YN): Bunlar gerçek değeri 1 ancak tahmin edilen değerin 0 olduğu durumlardır.

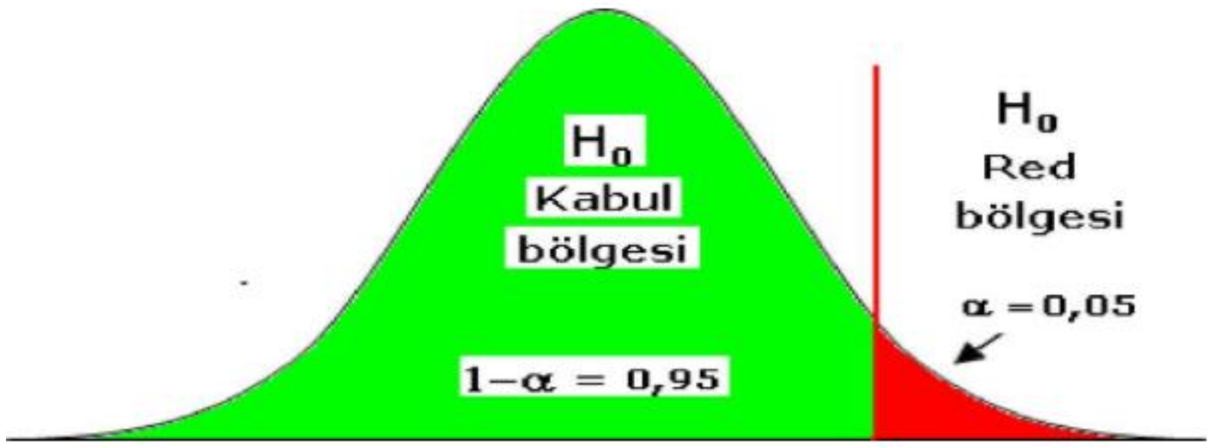

PREDICTIVE VALUES
POSITIVE (CAT)
NEGATIVE (DOG)
ACTUAL VALUES
POSITIVE (CAT)
NEGATIVE (DOG)
TRUE POSITIVE
3
YOU ARE A CAT
FALSE NEGATIVE
1
YOU ARE A DOG
TYPE II ERROR
FALSE POSITIVE
2
YOU ARE A CAT
TYPE I ERROR
TRUE NEGATIVE
4
YOU ARE NOT A CAT

Red Bölgesi (Rejection region) : Hipotezin reddedildiği alandır.

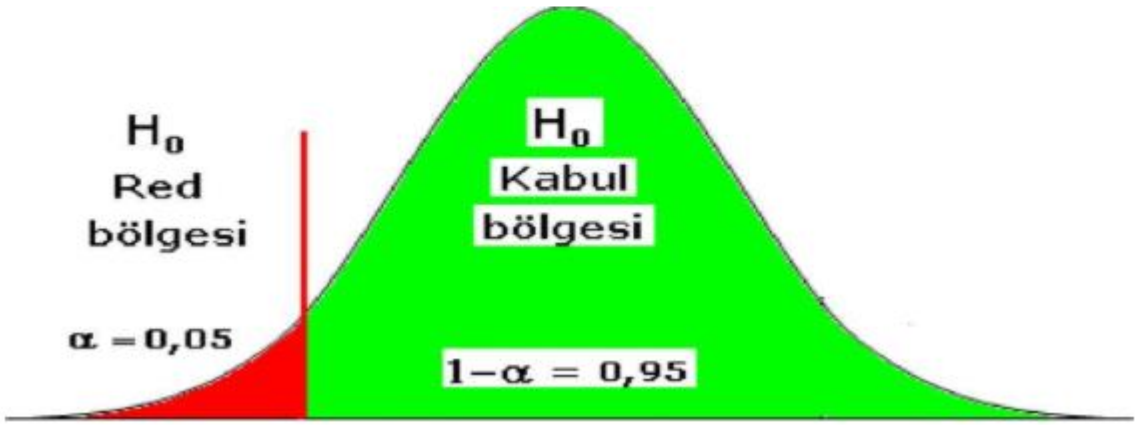
Kabul Bölgesi (Acceptence Region) : Hipotezin reddedilmediği alandır.

Örneğin α =0,05 düzeyinde yapılan iki yanlı bir testte kabul ve red bölgeleri aşağıdaki Normal Dağılım eğrisi üzerinde gösterilmiştir.

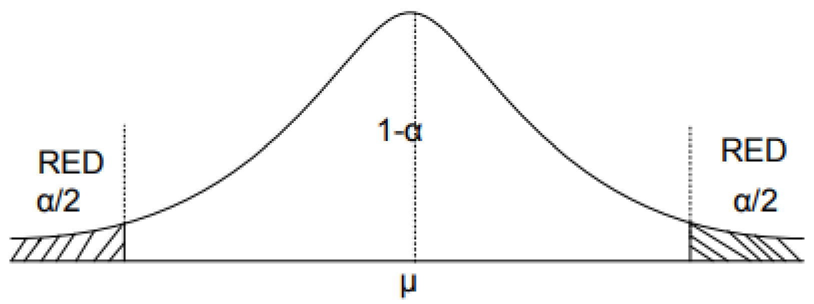


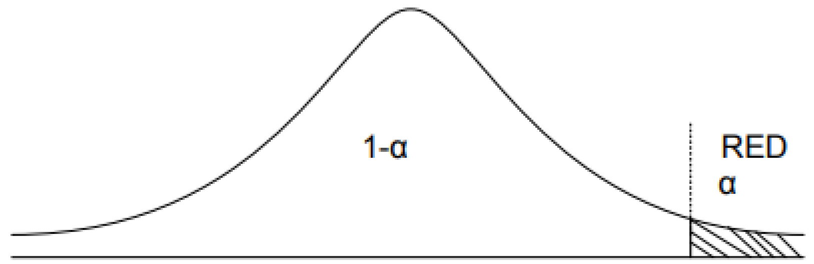


α =%5 düzeyinde yapılan tek yanlı bir testte kabul ve ret bölgeleri

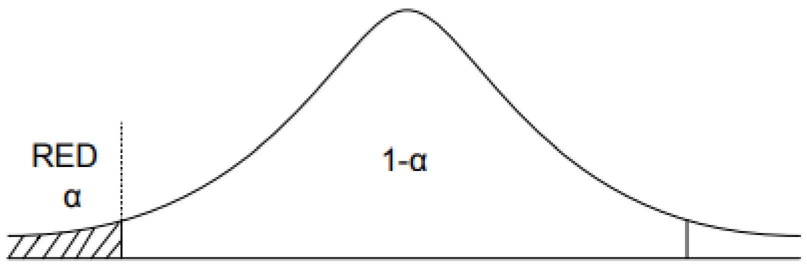


İki yanlı test (two sided test) : α/2 düzeyinde yapılan testtir.





Tek yanlı test (one sided test) : α düzeyinde yapılan testtir.



TEK ÖRNEKLEM HİPOTEZ TESTLERİ
(ONE SAMPLE HYPOTHESIS TESTING)

Tek örneklem hipotez testleri ,

- Kitle ortalaması (μ) için
- Kitle oranı (P) için
- Kitle varyansı (σ^2)için

oluşturulabilir.

KİTLE ORTALAMASI İÇİN HİPOTEZ TESTİ

Kitlenin dağılımının Normal dağılıma sahip olduğu varsayılır. Kitle ortalaması için hipotez testi aşamasında kitle varyansının $(\boldsymbol{\sigma^2})$ bilinip bilinmediğine ya da örneklem büyüklüğüne (n) bakılarak hangi testin kullanılacağına karar verilir.

i. Kitle varyansı biliniyorsa örneklem büyüklüğüne (n) bakılmaz, $Z = \dfrac{\bar{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}}$

ii. Kitle varyansı bilinmiyorsa ve n≥30 ise , $Z = \dfrac{\bar{X}-\mu_0}{\frac{S}{\sqrt{n}}}$

iii. Kitle varyansı bilinmiyorsa ve n< 30 ise, $t = \dfrac{\bar{X}-\mu_0}{\frac{S}{\sqrt{n}}}$ kullanılır.

Hipotez	$H_0: \mu=\mu_0$ $H_S: \mu\neq \mu_0$	$H_0: \mu=\mu_0$ $H_S: \mu> \mu_0$	$H_0: \mu=\mu_0$ $H_S: \mu< \mu_0$
Karar	$\lvert Z\rvert \geq Z_{\frac{\alpha}{2}}$ ise H_0 reddedilir	$Z \geq Z_{\alpha}$ ise H_0 reddedilir	$Z \leq -Z_{\alpha}$ ise ($\lvert Z\rvert \geq Z_{\alpha}$ ise) H_0 reddedilir
Test İstatistiği	$Z = \frac{\bar{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}}, \quad Z = \frac{\bar{X}-\mu_0}{\frac{S}{\sqrt{n}}}$		

Hipotez	$H_0: \mu=\mu_0$ $H_S: \mu \neq \mu_0$	$H_0: \mu=\mu_0$ $H_S: \mu > \mu_0$	$H_0: \mu=\mu_0$ $H_S: \mu < \mu_0$
Karar	$\lvert t \rvert \geq t_{(n-1),\frac{\alpha}{2}}$ ise H_0 reddedilir	$t \geq t_{(n-1),\,\alpha}$ ise H_0 reddedilir	$t \leq -t_{(n-1),\alpha}$ ise ($\lvert t \rvert \geq t_{(n-1),\alpha}$ ise) H_0 reddedilir
Test İstatistiği	$t = \dfrac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}}$		

p değeri

İstatistiksel paket programları, bir hipotez testi sonucunda gerçekleşen I. Tip Hata miktarını hesaplar ve bu değere p değeri denir. p değeri önceden belirlenmiş α değeri ile karşılaştırılarak karar verilir.

Eğer:

$p \leq \alpha$ ise H_0 reddedilir. Bunun anlamı, H_0 'ı reddetmekle gerçekleşen yanılgı öngörülenden küçüktür. Dolayısıyla H_0 reddedilebilir.

$p > \alpha$ ise H_0 reddedilmez. Bunun anlamı gerçekleşen yanılgı öngörülenden küçük olmadığı için H_0 reddedilemez.

Serbestlik Derecesi

Örneklemden hesaplanan bir istatistiğin kitlenin gerçek değerini tahmin etmek amacıyla yapılan hesaplamalarda ya da test istatistiğinin tablo değerlerini belirlemede serbestlik derecesine (degrees of freedom) ihtiyaç vardır. Serbestlik derecesi, t, F ve k,-kare dağılımlarında dağılımın şeklini belirler.

•Serbestlik derecesinin sayısı, bir istatistiğin hesaplamasında, değişiklik göstermekte serbest olan sayıdır. Bir değişkene ilişkin elde edilen değerlerin değişiklik gösterebilme serbestliği olarak ifade edilebilir.
Hipotez Testinin Aşamaları:

I. H_0 ve H_S Hipotezlerinin belirlenir
II. Test İstatistiğine karar verilir (t, Z, ki-kare, F gibi).
III. Anlamlılık düzeyine karar verilir (α=%1, %5,%10).
IV. Tablo değeri elde edilir
V. Karar verilir ve yorumlanır

Örneğin kolesterol düzeyi normal sınırlarda kitle ortalaması 200 standart sapması 34 ise bu kitleden çekilen bir örneklemin çekildiği kitlenin ortalamasının 200 olup olmadığını incelemek için kurulacak hipotezler:

Örneklemin çekildiği kitlenin ortalaması 200'den farklıdır.	$H_0: \mu = 200$ $H_S: \mu \neq 200$ İki yanlı test
Örneklemin çekildiği kitlenin ortalaması 200'den büyüktür.	$H_0: \mu = 200$ $H_S: \mu > 200$ Tek yanlı test
Örneklemin çekildiği kitlenin ortalaması 200'den küçüktür.	$H_0: \mu = 200$ $H_S: \mu < 200$ Tek yanlı test

Örnek 1: Ağrı kesici bir ilacın ortalama 60 dakikadan daha fazla bir sürede etkisini göstereceği iddia ediliyor. Rastgele seçilen bir örneklemde yer alan hastalardan 64'üne bu ilaç veriliyor ve ortalama ağrıyı 63 dakikada kestiği görülüyor ve standart sapması 12 bulunuyor. Bu iddianın doğruluğunu %5 anlamlılık düzeyinde test ediniz.

Çözüm: Kitle varyansı bilinmiyor ancak n>30 onun için z testi kullanılır.

H_0: μ = 60
H_s: μ > 60

$$Z = \frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}} = \frac{63-60}{\frac{12}{\sqrt{64}}} = 2$$

$$Z_{tablo} = Z_{\alpha=0,05} = 1,65$$

$Z_{hesap} > Z_{tablo}$ olduğundan H_0 hipotezi reddedilir.

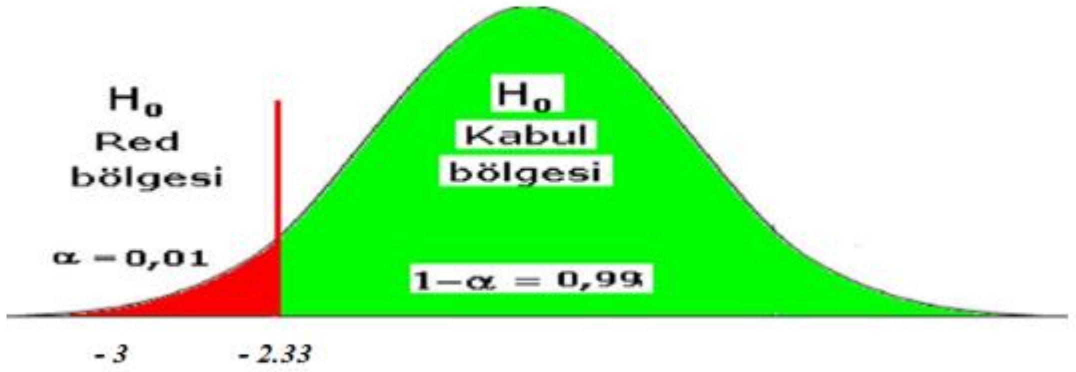
Yorum: Bu ağrı kesici ilacın ağrıyı kesme ortalamasının 60 dakikadan daha fazla olduğunu %5 anlamlılık düzeyinde söyleyebiliriz.

Örnek 2: Bir hastalığın en fazla 10 günde iyileştiği iddia edilmektedir. Bu hastalığa sahip 144 kişi üzerinde yapılan araştırmada ortalama 9 günde iyileşme olduğu ve standart sapmasının 4 olduğu tespit edilmiştir. %1 anlamlılık düzeyinde bu iddianın doğruluğunu test ediniz.

Çözüm: Kitle varyansı bilinmiyor ancak n>30 onun için Z testi kullanılır.

$H_0 : \mu = 10$
$H_s : \mu < 10$

$$Z = \frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}} = \frac{9-10}{\frac{4}{\sqrt{144}}} = -3$$



$$Z_{tablo} = -Z_{\alpha=0,01} = -2{,}33$$

$Z_{hesap} < Z_{tablo}$ olduğundan H_0 hipotezi reddedilir.

Yorum: Bu hastalıkta ortalama iyileşme süresinin10 gün'den daha az olduğunu %1 anlamlılık düzeyinde söyleyebiliriz.

Örnek 3: Bir çikolata firması 500 gramlık paketler halinde üretim yapmayı planlamaktadır. Üretimin planlandığı gibi gerçekleşip gerçekleşmediğini kontrol etmek için rastgele olarak seçilen 100 paketin ortalama ağırlığı 495 gram ve standart sapması 20 gram olarak bulunmuştur. Üretimin planlandığı gibi gerçekleşip gerçekleşmediğini α = 0.05 anlamlılık düzeyinde test ediniz.

Çözüm: Kitle varyansı bilinmiyor ancak n>30 onun için z testi kullanılır.

$H_0 : \mu = 500$
$H_s : \mu \neq 500$

$$Z = \frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}} = \frac{495-500}{\frac{20}{\sqrt{100}}} = -2,5$$

$Z_{tablo} = Z_{\alpha=0,05/2} = 1,96$

$|Z_{hesap}| \geq Z_{tablo}$ olduğundan H_0 hipotezi reddedilir.

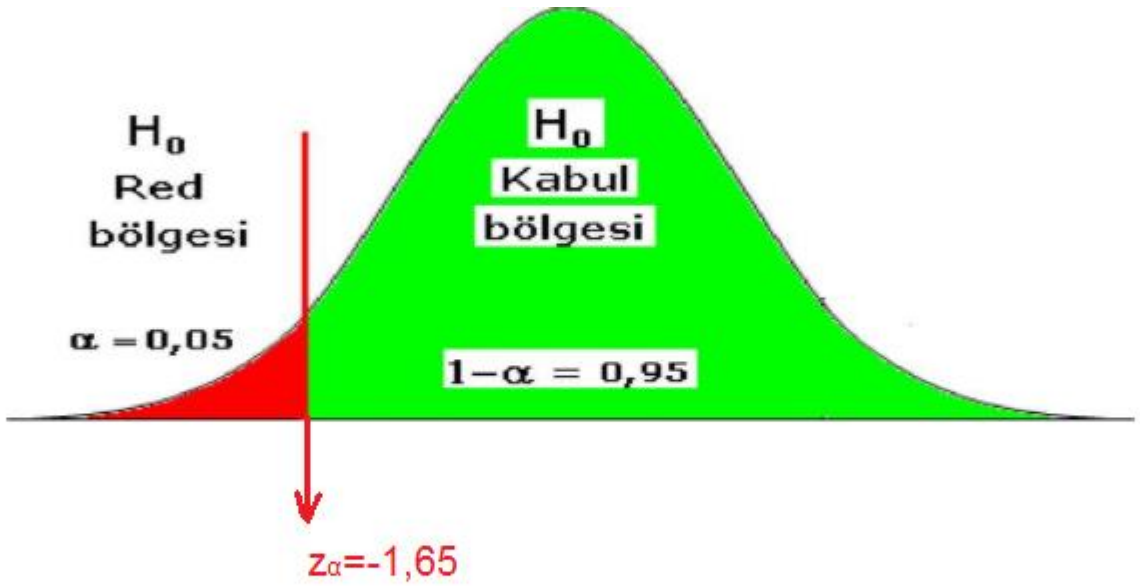
Yorum: Paket ortalamasının 500 gramdan farklı olduğunu %5 anlamlılık düzeyinde söyleyebiliriz. Üretim planlandığı gibi gerçekleşmemiştir.

Örnek 4. Bir firma ürettiği sabunlardaki PH değerinin 5,5’dan küçük olduğunu iddia etmektedir. Bu sabunlardan satın alacak bir market o günkü üretimden rastgele 36 sabun seçerek incelemiş, ortalama PH değerini 5,01 standart sapmasını 1,28 bulmuştur. Sabunlarının ortalama PH değerinin 5,5’dan daha düşük olup olmadığını %5 anlamlılık düzeyinde test ediniz.

Çözüm: Kitle varyansı bilinmiyor, n>30

H_0: $\mu=5,5$

H_S: $\mu< 5,5$



$$Z = \frac{\bar{X}-\mu_0}{\frac{S}{\sqrt{n}}} = \frac{5-5,5}{\frac{1,28}{\sqrt{36}}} = -2,301$$

$$Z_{tablo} = -Z_{\alpha=0,05} = -1,65$$

$Z_{hesap} < Z_{tablo}$ olduğundan H_0 reddedilir (p=0,027<0,05 H_0 Ret).

Yorum: Sabunların ortalama PH değerinin 5,5’dan daha düşük olduğu %5 anlamlılık düzeyinde söyleyebilir. Üreticinin iddiası doğrudur.

Örnek 5. 15 hastaya ait sistolik kan basınçları verilmiştir. Bu bireyler kan basınç değerleri sağlıklı toplumda istenen 120 mm/HG'ye eşit midir (α=0,05)?

Çözüm: σ^2 bilinmiyor ve n<30 olduğu için t testi kullanılacak.

Sistolik kan basıncı
117
125
118
140
104
120
127
113
112
128
119
126
104
130
134

H_0: $\mu = 120$
H_S: $\mu \neq 120$

Verilerden $\bar{X} = 121{,}13$ S=10,31 hesaplanır.

$$t = \frac{\bar{X}-\mu_0}{\frac{S}{\sqrt{n}}} = \frac{121{,}13-120}{\frac{10{,}31}{\sqrt{15}}} = 0{,}426$$

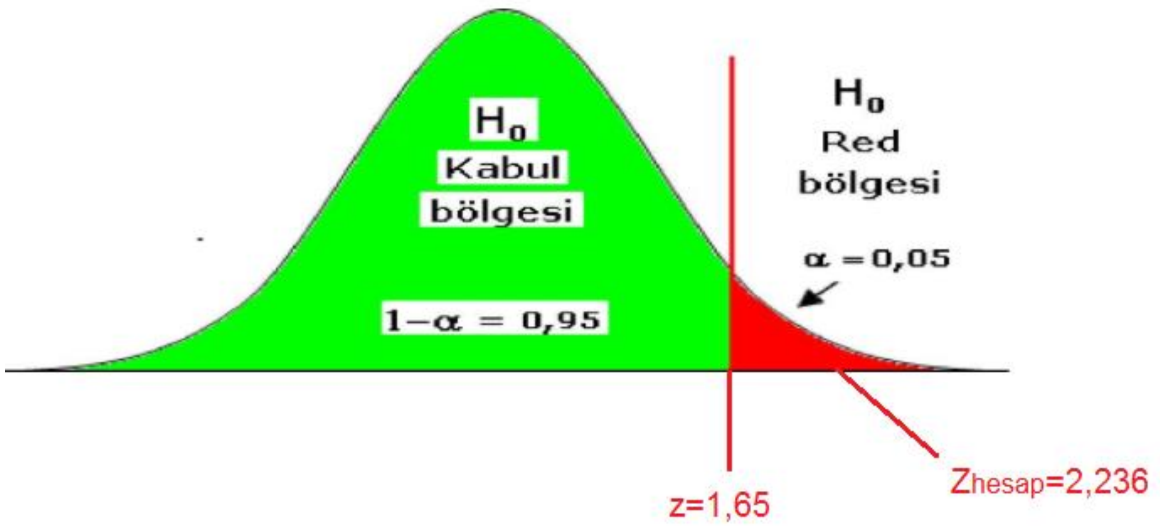
t dağılımı tablosundan $t_{0,05/2,\ 14}=t_{0.025,14}$=2,14 bulunur.

$t_{hesap}<t_{tablo}$ olduğundan H_0 reddedilmez.

Bireylerin sistolik kan basınçlarının sağlıklı bireylere ait toplum parametresi olan 120 mm/Hg' ye eşit olduğunu %5 anlamlılık düzeyinde söyleyebiliriz. Başka bir anlatımla bu bireyler sağlıklı topluma ait bireylerdir.

Örnek 6. Bir tür sigara paketindeki bir sigaradaki ortalama nikotin miktarının 0,7mg aştığı iddia edilmektedir. Sigaralardaki nikotin miktarları, 0,06 standart sapma ile normal dağılış gösterdiği bilinmektedir. Bunun için bu tür markadan rasgele bir paket alınıyor ve nikotin miktarları ölçülüyor. Nikotin miktarlarının ortalaması 0,73 olarak bulunuyor. Bu iddianın doğruluğunu 0,05 anlamlılık düzeyinde test ediniz. Bir pakette 20 sigara bulunmaktadır.

Çözüm: σ^2 biliyor ve örneklem büyüklüğüne bakılmaksızın z testi kullanılır.



H_0: μ=0,7mg

H_S: μ> 0,7mg

$$Z = \frac{\bar{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{0,73-0,7}{\frac{0,06}{\sqrt{20}}} = 2,236 \qquad Z_{tablo} = Z_{\alpha=0,05} = 1,65$$

$Z_{hesap} \geq Z_{tablo}$ olduğundan H_0 reddedilir.

Yorum: %5 anlamlılık düzeyinde bir sigaradaki ortalama nikotin miktarı 0,7mg'ı aştığını söyleyebiliriz.

Örnek 7. Otomobil güvenlik testlerin ile ilgili yapılan çalışmalarda, saatte 80 km hızla giden bir arabanın ortalama fren mesafesinin 25cm'den az olduğu iddia edilmektedir. Bu amaçla 100 sürüş testinde kat edilen ortalama mesafe 24,5cm ve standart sapma 4 olarak bulunmuştur. (α=0,05)?

Çözüm: σ^2 bilinmiyor ve n>30 olduğu için Z testi kullanılır.

H_0: μ=25cm

H_S: μ< 25cm

$$Z = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} = \frac{24,5 - 25}{\frac{4}{\sqrt{100}}} = -1,25 \qquad Z_{tablo} = Z_{\alpha=0,05} = 1,65$$

$|Z_{hesap}| < Z_{tablo}$ olduğundan H_0 hipotezi reddedilemez.

Yorum: %5 anlamlılık düzeyinde, 80 km/saat hızla giden bir aracın fren mesafesinin 25cm az olmadığını söyleyebiliriz.

Örnek 8. Ankara ilinde 28 Mart 2020 gününde ortalama 2mm yağış aldığı iddia ediliyor. 28 Mart 2020 gününde saat başı yapılan tahminleri sonucunda yağış miktarının ortalaması 1,8mm ve standart sapması 1mm bulunuyor (n=24). Yağış miktarının normal dağılıma uyduğu bilindiğine göre 28 Mart 2020 gününde ortalama yağış miktarının 2mm olup olmadığını %5 anlamlılık düzeyinde test ediniz. (α=0,05)?

Çözüm: σ^2 bilinmiyor ve n<30 olduğu için t testi kullanılır.

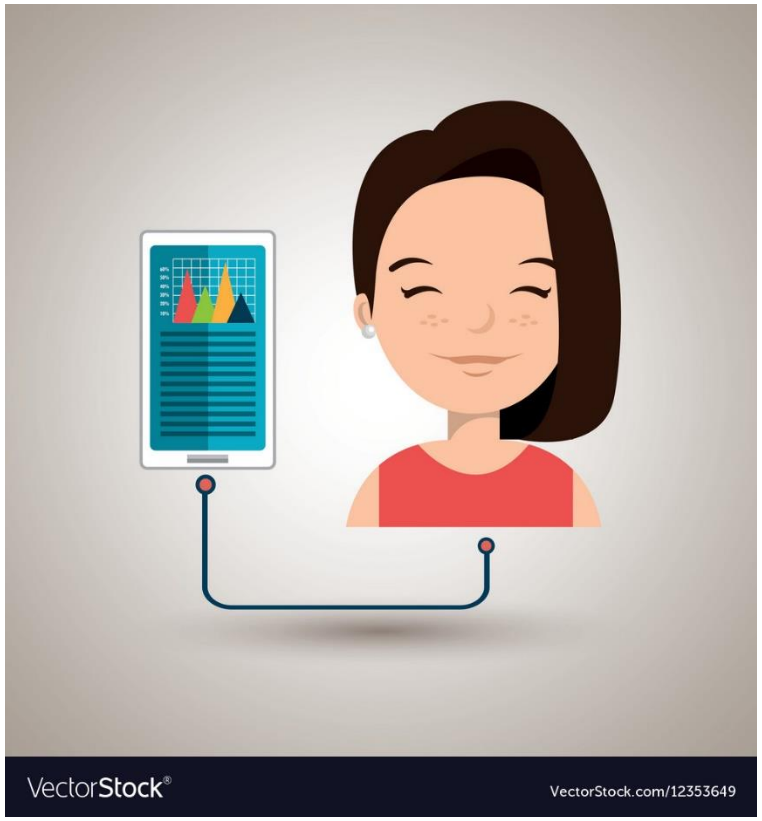
H_0: μ=2mm

H_S: μ≠ 2mm

$$t = \frac{\bar{X}-\mu_0}{\frac{S}{\sqrt{n}}} = \frac{1{,}8-2}{\frac{1}{\sqrt{24}}} = -0{,}979 \qquad t_{tablo} = t_{\frac{\alpha}{2}=0{,}025,23} = 2{,}064$$

$|t_{hesap}| < t_{tablo}$ olduğundan H_0 reddedilemez.

Yorum: %5 anlamlılık düzeyinde 28 Mart 2020 gününde ortalama 2mm yağış aldığı iddiası reddedilemez.



Bir sonraki derste kitle oranı ve kitle varyansı için hipotez testleri incelenecek.

KAYNAKLAR

1-) "İstatistiksel Yöntemlere Giriş", H.Demirhan, C.Hamurkaroğlu H.Ü.Yayınları, 2011.

2-) "Discovering Statistics Using SPSS for Windows : Advanced Techniques for the Beginner", Andy Field, Ref No: HA32.F54 2000.

3-) "SPSS for Windows : An Introduction to Use and Interpretation in Research", George A. Morgan, Orlando V. Griego, Gene W. Gloeckner. Ref No: HA 32M667 2001.

4-) "Modern Elementary Statistics", John.E.Freund, Prentice Hall, 2004.

5-) "Temel İstatistik Yöntemler", Serpil Cula, Zehra Muluk, Başkent Üniversitesi yayınları,2006.

https://docplayer.biz.tr/14975474-Hipotez-testleri-yrd-doc-dr-emre-atilgan.html

http://genderi.org/bolum-13-istatistikte-tahmin-ve-hipotez-testleri-istatistik.html