HACETTEPE
ÜNİVERSİTESİ
İSTATİSTİK BÖLÜMÜ

İST156 İSTATİSTİĞE GİRİŞ II

DERS 6- HİPOTEZ TESTLERİ-Bölüm 2

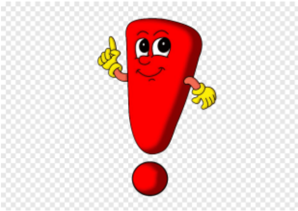
Ders sorumluları: Prof.Dr.Serpil AKTAŞ ALTUNAY (01 Şubesi)
Doç.Dr. Ayten YİĞİTER (02 Şubesi)

1. KİTLE ORANI (P) İÇİN HİPOTEZ TESTİ

Bir bölgede nadir bir hastalığa sahip olan kişilerin oranı, bir kamuoyu yoklamasında anayasa değişikliğine «evet» diyen kişilerin oranı, bir web sitesinden alışveriş yapan kişilerin oranı, bir ilacın yan etkisinin görülme oranı vb. araştırmalarda gözlem birimleri belli bir özelliğe sahip olup olmadıklarına göre iki gruba ayrılır ve kitle İkiterimli (Binom) kitlesi olarak adlandırılır. Kitledeki orana (P) ilişkin hipotezlerin (iddiaların) testinde Binom dağılımının normal dağılıma yakınsamasından yararlanılır. Bu yakınsama istatistikte *Merkezi Limit Teoremi* olarak bilinir.

Örneğin, hazır gıda tüketimine bağlı olarak İstanbul'da obezite oranının kadınlarda %45 olduğu iddia edilmektedir.

H_0: $P = 0,45$ hipotezine karşılık



Hs: $P \neq 0,45$,

Hs: $P > 0,45$,

Hs: $P < 0,45$ seçenek hipotezlerinden birisi kurulur.

Kitlenin dağılımının Binom (n, p) dağılımına sahip olduğu varsayılır. Binom dağılımının kitle ortalaması $\boldsymbol{np}$ ve kitle varyansı $(\boldsymbol{np}(\mathbf{1}-\boldsymbol{p}))$'dir. Örneklemden elde edilen oran $\frac{\boldsymbol{X}}{\boldsymbol{n}}$, kitle oranı p'nin nokta tahminidir.

Merkezi Limit Teoremi'nin bir sonucu olarak, örneklem büyüklüğü (n) yeterince büyük olduğu durumda yani ($n \to \infty$) için,

$$Z=\frac{X-np}{\sqrt{np(1-p)}}\sim N(0,1)$$

istatistiğinin dağılımı standart normal dağılıma sahiptir.

Hipotez	$H_0: P = P_0$ $H_S: P \neq P_0$	$H_0: P = P_0$ $H_S: P > P_0$	$H_0: P = P_0$ $H_S: P < P_0$
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	$Z = \dfrac{p - P_0}{\sqrt{\dfrac{P_0(1-P_0)}{n}}}$		

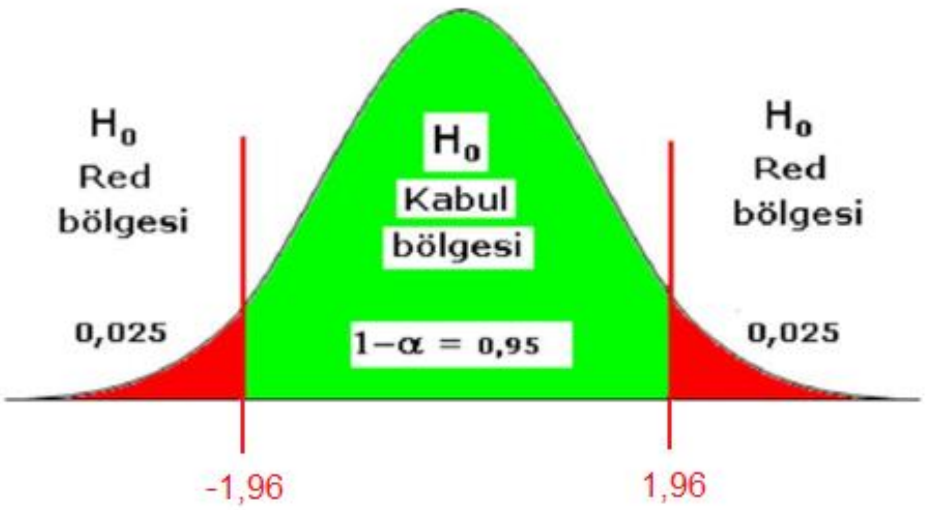
Örnek 1: Televizyonda yeni yayınlanmaya başlayan bir dizi için bu diziyi izleyenlerden rastgele seçilen 100 izleyiciye «diziyi beğendiniz mi?» sorusuna 70’nin cevabının «evet» olduğu görülmüştür. Bu bilgileri kullanarak, kitledeki gerçek beğeni oranının %60 olup olmadığını %5 anlamlılık düzeyinde araştırınız.

Çözüm:

$$Z = \frac{p-P_0}{\sqrt{\frac{P_0(1-P_0)}{n}}} = \frac{0,70-0,60}{\sqrt{\frac{0,60(1-0,60)}{100}}} = 2,0412$$

H_0: P = 0,60
H_s: P $\neq$ 0,60

örneklemdeki oran $\frac{X}{n} = p = \frac{70}{100} = 0,70$



$Z_{tablo} = Z_{\frac{\alpha}{2}=0,025}$=1,96

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

Yorum: Diziyi beğenenlerinin gerçek oranının 0,60’tan farklı olduğunu %5 anlamlılık düzeyinde söyleyebiliriz.

Örnek 2: Piyasaya yeni bir ürün süren bir firmanın çağrı merkezine yeni ürün ile ilgili olarak gelen şikayet oranının 0,02'den çok olduğu iddia ediliyor. Bu iddianın araştırılması için çağrı merkezine gelen çağrılar arasından 120 tanesi rastgele seçiliyor ve bunlardan 5'nin yeni ürünle ilgili olduğu görülüyor. Bu iddianın doğruluğunu %5 anlamlılık düzeyinde araştırınız.

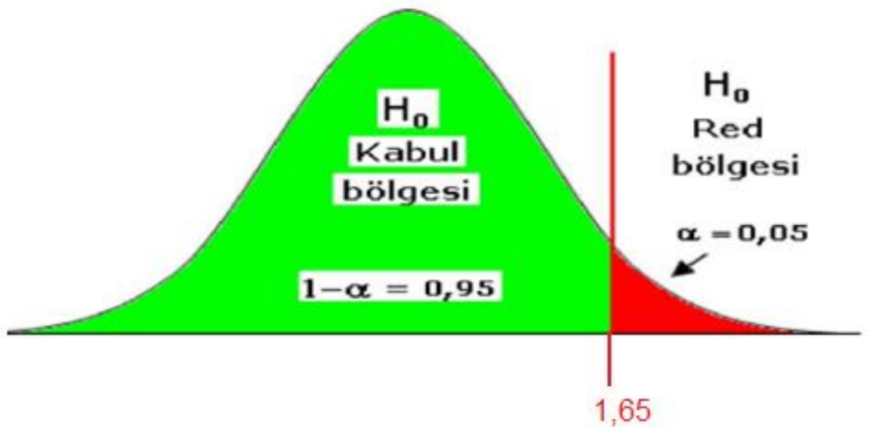
Çözüm:

H_0: P = 0,02

H_s: P > 0,02

$$Z = \frac{p-P_0}{\sqrt{\frac{P_0(1-P_0)}{n}}} = \frac{\frac{5}{120}-0,02}{\sqrt{\frac{0,02(1-0,02)}{120}}} = 1,695$$

örneklemdeki oran $\frac{X}{n} = \frac{5}{120} = 0,04167$



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

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

Yorum: Çağrı merkezine gelen şikayetlerin gerçek oranının 0,02'yi aştığını %5 anlamlılık düzeyinde, söyleyebiliriz.

Örnek 3: H. Ü. Öğrencilerinin Beytepe kampüsüne gelmektense, «uzaktan eğitimi tercih ederiz» görüşünde olan öğrencilerin oranının %50'den az olduğu iddia edilmektedir. Bu amaçla rastgele seçilen 400 öğrencinin görüşleri alınmış, 180'inin bu görüşte oldukları görülmüştür. İddianın doğruluğunu %4 anlamlılık düzeyinde araştırınız.

Çözüm:

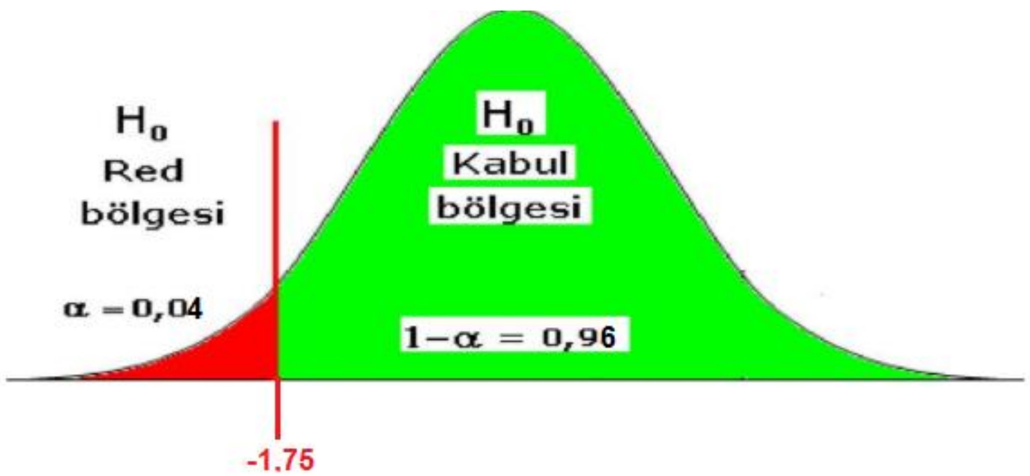
H_0: P = 0,50

H_s: P < 0,50

$$Z = \frac{p - P_0}{\sqrt{\frac{P_0(1-P_0)}{n}}} = \frac{0,45 - 0,50}{\sqrt{\frac{0,50(1-0,50)}{400}}} = -2$$

örneklemdeki oran $\frac{X}{n} = p = \frac{180}{400} = 0,45$

$Z_{tablo} = -Z_{\alpha=0,04} = -1,75$



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

Yorum: Öğrencilerin uzaktan eğitimi tercih etme oranının %50'den az olduğunu %4 anlamlılık düzeyinde söyleyebiliriz.

2. KİTLE VARYANSI (σ^2) İÇİN HİPOTEZ TESTİ

Ortalama gibi varyans da dağılım hakkında bilgi veren önemli bir niceliktir ve araştırmacılar ortalama etrafındaki değişkenliğini de değişip değişmediğini araştırmak isterler.

Normal dağılıma sahip bir kitleden rasgele alınan n birimlik bir örneklem olsun. Kitle varyansı **σ^2**'ye ilişkin hipotez testi ki-kare ($\chi^2_{(n-1)}$) dağılımına dayanır:

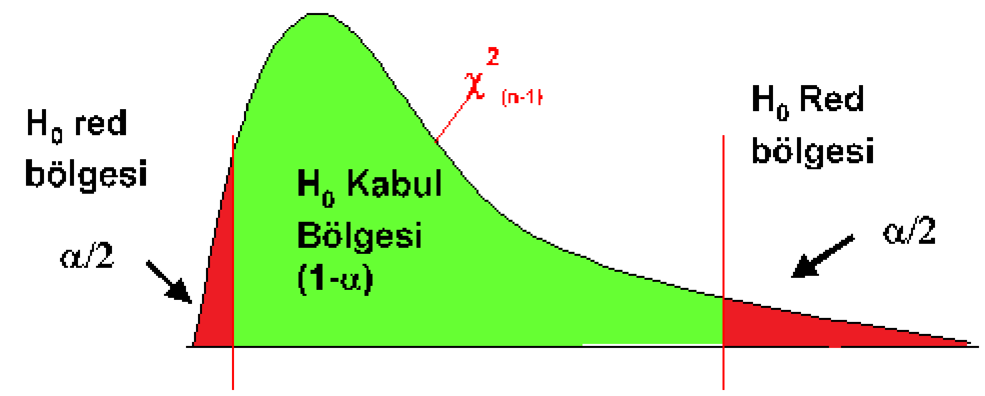
$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

Serbestlik Derecesi

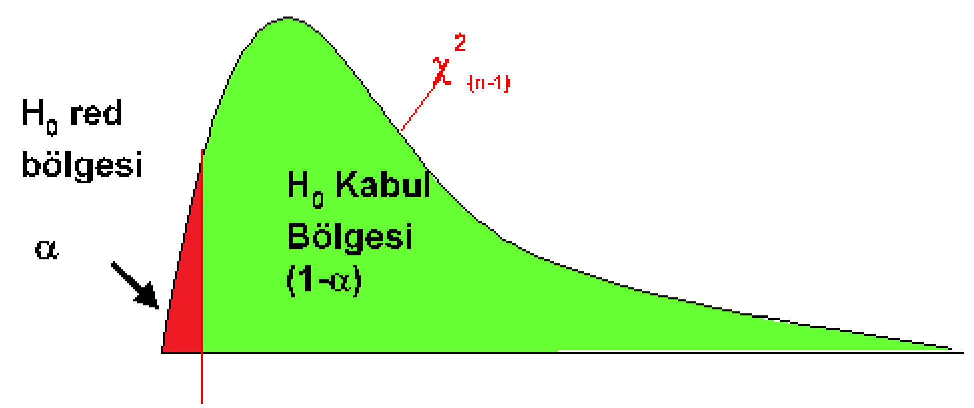
Burada $S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$ örneklemin varyansıdır.

Hipotez	H_0: $\sigma^2 = \sigma_0^2$ H_S: $\sigma^2 \neq \sigma_0^2$	H_0: $\sigma^2 = \sigma_0^2$ H_S: $\sigma^2 > \sigma_0^2$	H_0: $\sigma^2 = \sigma_0^2$ H_S: $\sigma^2 < \sigma_0^2$
Karar	$\chi^2 \geq \chi^2_{(n-1),\frac{\alpha}{2}}$ ya da $\chi^2 \geq \chi^2_{(n-1),1-\frac{\alpha}{2}}$ ise H_0 reddedilir	$\chi^2 \geq \chi^2_{(n-1),\alpha}$ise H_0 reddedilir	$\chi^2 \leq \chi^2_{(n-1),1-\alpha}$ise H_0 reddedilir
Test İstatistiği	$\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$		

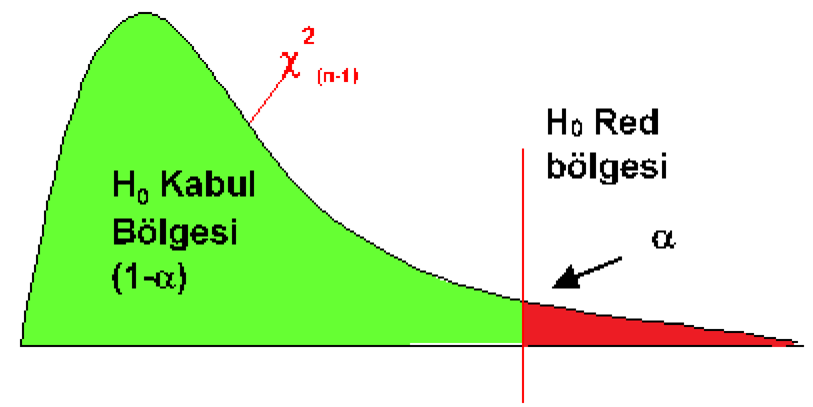
İki yanlı test (two sided test) : α/2 düzeyinde yapılan testtir.



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



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



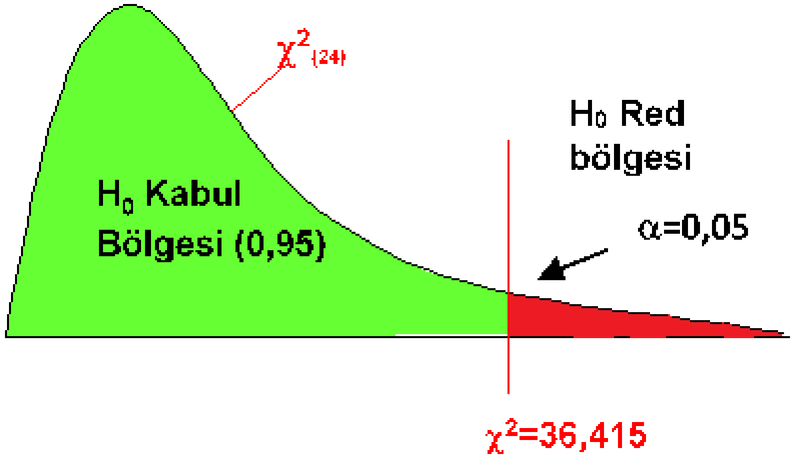
Örnek 4: Genetik olarak değiştirilmiş domates bitkisinin yapraklardaki klorofil miktarı (μmol/gr) üzerinde yapılan bir çalışmada bir seradan 25 adet domates fidesi rastgele olarak seçilmiştir. Her fidenin üzerinden de bir yaprak rasgele olarak seçilmiş ve klorofil miktarı ölçümü yapılmıştır. Yapılan 25 ölçümün standart sapması 1,765 (μmol/gr) olarak bulunmuştur. Genetik olarak değiştirilmiş domates bitkisinin yapraklarındaki klorofil miktarının varyansının 3,24'ten büyük olması iddiasını, klorofil miktarlarının normal dağılımlı olduğu varsayımı altında test ediniz (α=0,05).

Çözüm:

H_0: $\sigma^2 = 3{,}24$

H_S: $\sigma^2 > 3{,}24$

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} = \frac{(25-1)(1{,}765^2)}{3{,}24} = 23{,}075$$



$$\chi^2{}_{tablo} = \chi^2{}_{24,\alpha=0{,}05} = 36{,}415$$

$\chi^2{}_{hesap} < \chi^2{}_{tablo}$ olduğundan H_0 hipotezi reddedilmez.

Yorum: Klorofil miktarının varyansının 3,24 olduğunu %5 anlamlılık düzeyinde söyleyebiliriz.

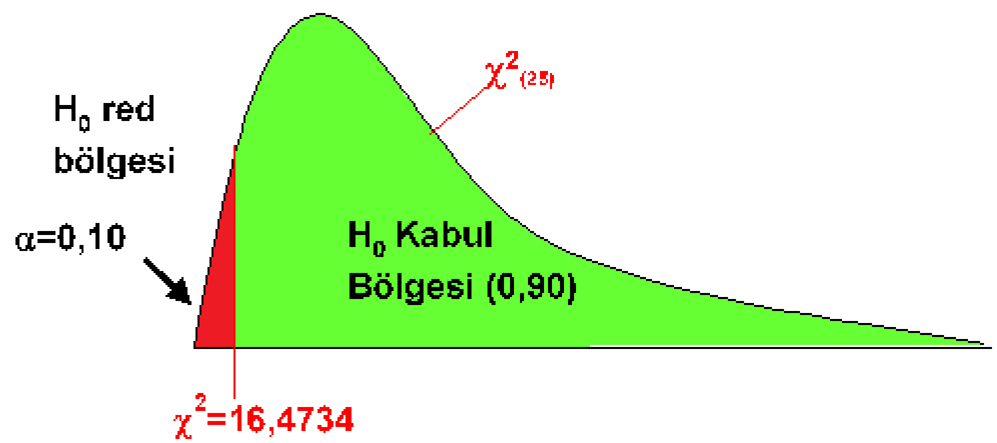
Örnek 5: Kimya laboratuvarında yapılan test için oda sıcaklığındaki değişimin $25C^0$'den az olması istenmektedir. Gün içinde 26 farklı zamanda sıcaklık ölçülmüş ve sıcaklıkların standart sapması $4C^0$ olarak bulunmuştur. Test yapılacak odanın istenen standartlara uygun olup olmadığını, sıcaklık ölçümlerinin dağılımının normal dağılımlı olduğu varsayımı altında test ediniz (α=0,10).

Çözüm:

$H_0: \sigma^2 = 25$

$H_S: \sigma^2 < 25$

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} = \frac{(26-1)(4^2)}{25} = 16$$



$${\chi^2}_{tablo} = {\chi^2}_{25,\alpha=0,90} = 16,4734$$

${\chi^2}_{hesap} < {\chi^2}_{tablo}$ olduğundan H_0 hipotezi reddedilir.

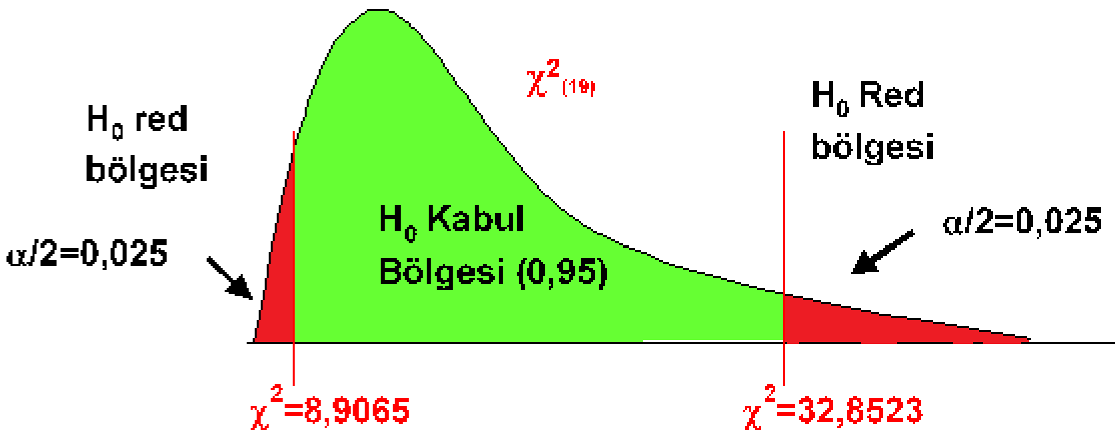
Yorum: Test yapılacak odanın istenen standartlara uygun olduğunu %10 anlamlılık düzeyinde söyleyebiliriz.

Örnek 6: Bir havayolu şirketinin uçuşlarının her zaman zamanında olduğunu ve ortalama 15 dakika gecikmeyle hak talebinde bulunulduğunu varsayalım. Şirket ortalama gecikme süresinde o kadar tutarlıdır ki, standart sapmanın 12 dakika olduğunu iddia etmektedir. İddianın tutarlılığından şüphe duyan bir yolcu, sonraki 20 uçuşunun gecikmelerini hesaplar. Bu 20 uçuş için ortalama gecikme 15 dakika, standart sapma 14 dakika olarak bulur. Havayolu şirketinin iddiasını doğru olup olmadığını test ediniz (α=0,05)

Çözüm:

H_0: $\sigma^2 = 144$

H_S: $\sigma^2 \neq 144$

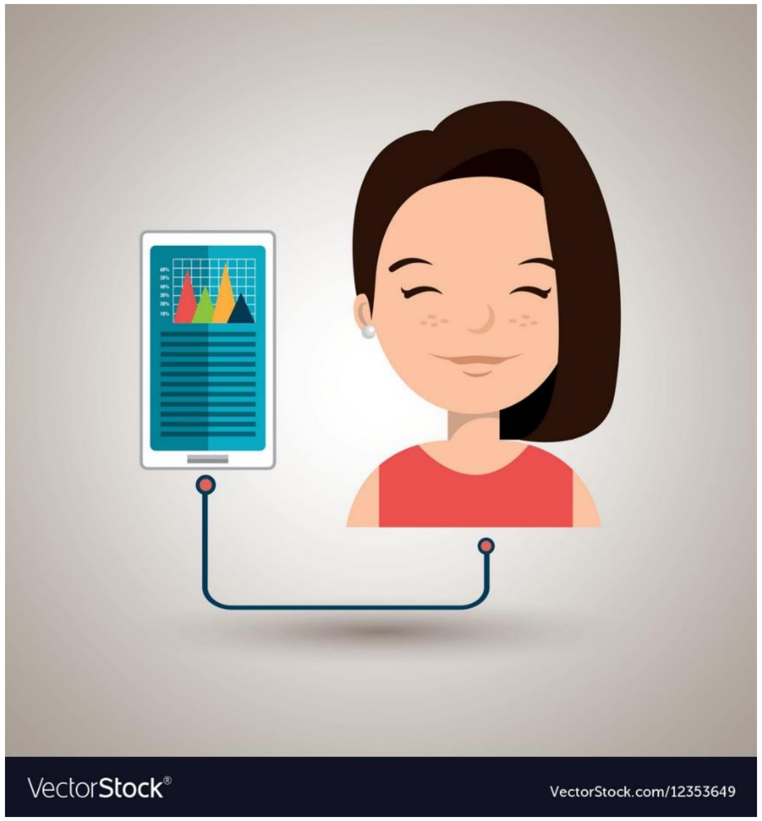


$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} = \frac{(20-1)(14^2)}{12^2} = 25{,}86$$

$\chi^2{}_{tablo} = \chi^2{}_{19,\alpha=0{,}975} = 8{,}9065$ ve $\chi^2{}_{19,\alpha=0{,}025}$=32,8523

$\chi^2{}_{hesap} > \chi^2{}_{19,\alpha=0{,}975}$ ve $\chi^2{}_{hesap} < \chi^2{}_{19,\alpha=0{,}025}$ olduğundan H_0 hipotezi reddedilemez.

Yorum: %5 anlamlılık düzeyinde şirketin iddiasının reddedecek yeterli kanıt bulunamamıştır.



Bir sonraki derste «Örneklem büyüklüğünün hesaplanması konusu» 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://opentextbc.ca/introbusinessstatopenstax/chapter/test-of-a-single-variance/