



Bölüm 6

Hipotez Testi: Tek Örnekli Testler



UZAKTAN EĞİTİM UYGULAMA VE ARAŞTIRMA MERKEZİ

telefon 0(312) 202 82 00 • eposta guzem@gazi.edu.tr • adres Gazi Üniversitesi Rektörlük Binası No:6/f

guzem.gazi.edu.tr • uzaktanegitim.gazi.edu.tr • lms.gazi.edu.tr

Öğrenme hedefleri

Bu bölümde aşağıdaki konular öğrenilecektir:

- Hipotez testinin temel prensipleri
- Hipotez testinin bir ortalama, oranın veya varyansın test edilmesinde nasıl kullanılacağı
- Her bir hipotez testi sürecinin varsayımları, nasıl değerlendirileceği ve ihlal edilmişse getireceği sonuçlar
- Hipotez testinde gizli tehlikelerin nasıl önleneyeceği
- Hipotez testinde gizli tehlikeler & ahlaki konular

Giriş



- Daha önceki bölümde örneklem verisini kullanarak bir popülasyon parametresi ile ilgili nasıl güven aralığı oluşturulacağı üzerinde durduk.
- Ancak bir çok mühendislik probleminde bir popülasyon parametresi ile ilgili çelişen iddialar veya ifadelerden hangisinin doğru olduğu ile ilgili karar vermemiz gerekebilir.
- Bu ifadeler **hipotezler** olarak bilinir ve karar verme prosedürü de **hipotez testi** olarak anılır.
- Hipotez testleri istatistiksel çıkarsamanın çok önemli bir kısmını oluşturmaktadır çünkü mühendislikteki bir çok karar verme problemleri, testler, deneyler hipotez testi problemleri olarak formüle edilebilir.
- Ayrıca konunun içinde göreceğimiz gibi güven aralığı ile hipotez testi arasında yakın bir ilişki vardır.

Hipotez nedir?

- Hipotez, bir popülasyon parametresi hakkında bir iddiadır (beyandır):

- Popülasyon ortalaması

Örnek: Bulunduğumuz şehirdeki ortalama aylık cep telefonu faturası miktarı $\mu = \$42$ 'dir

- Popülasyon oranı

Örnek: Bu şehirdeki cep telefonlu yetişkinlerin oranı $\pi = 0.68$ 'dir.



Sıfır hipotezi, H_0

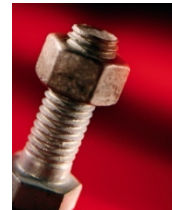
- Test edilecek iddia veya beyanı ifade eder

Örnek: Üretilmiş bir somunun ortalama çapı 30mm'dir
($H_0 : \mu = 30$)

- Her zaman popülasyon parametresi ile ilgilidir, örnek istatistiği ile ilgili değildir

$$H_0 : \mu = 30$$

$$H_0 : \bar{X} = 30$$

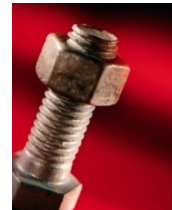


Sıfır hipotezi, H_0

- Sıfır hipotezinin doğru olduğu varsayımı ile başlanır
- Mevcut duruma atıf verir.
- “=”, veya “ $\leq$ ”, veya “ $\geq$ ” işaretlerini içerirler, bu derste her zaman “=” işareti kullanılacaktır
- Reddedilebilir veya reddedilemez

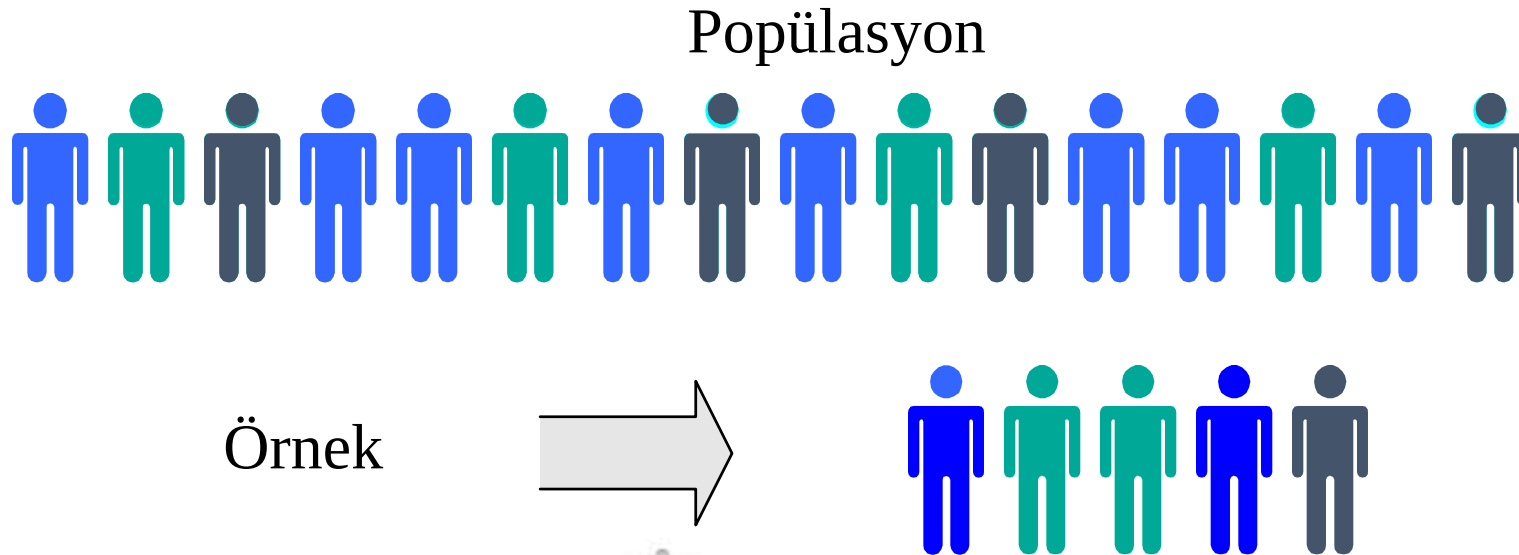
Karşıt hipotez, H_1

- Sıfır hipotezinin tersidir
 - yani, üretilmiş somunların ortalama çapı 30mm değildir. ($H_1: \mu \neq 30$)
- Mevcut durumu sorgular (karşı koyar)
- “ $\neq$ ”, veya “ $\leq$ ”, veya “ $\geq$ ” işaretlerini içerebilir
- İspatlanabilir de , ispatlanamayabilir de
- Genellikle araştırmacının ispatlamaya çalıştığı hipotezdir.



Hipotez testi süreci

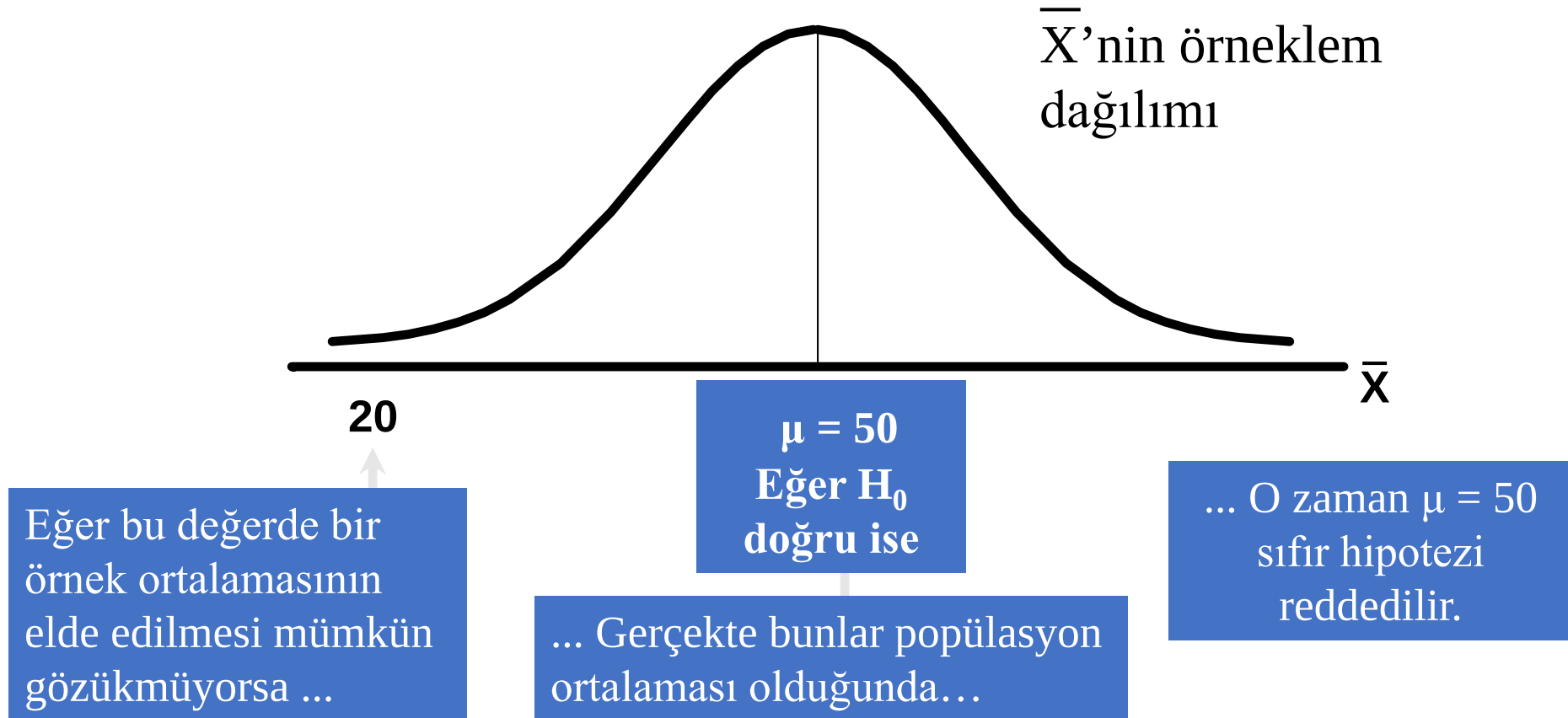
- İddia: Popülasyonun ortalama yaşı 50'dir.
 - $H_0: \mu = 50$, $H_1: \mu \neq 50$
- Popülasyonu örnekle ve örnek ortalamasını bul.



Hipotez testi süreci

- Örnek yaş ortalamasının $\bar{X} = 20$ olduğunu varsayalım.
- Bu, iddia edilen popülasyon ortalama yaşı olan 50'den oldukça düşüktür.
- Eğer sıfır hipotezi doğru ise, bu şekilde farklı bir ortalamanın elde edilme olasılığı oldukça düşük olacaktır, bu şekilde sıfır hipotezi reddedilir.
- Başka bir deyişle, eğer popülasyon ortalaması 50 idiyse, örnek ortalamasının 20 olarak elde edilmesi mümkün değildir, buradan popülasyon ortalamasının 50 olmaması gerektiği sonucuna varılır.

Hipotez testi süreci



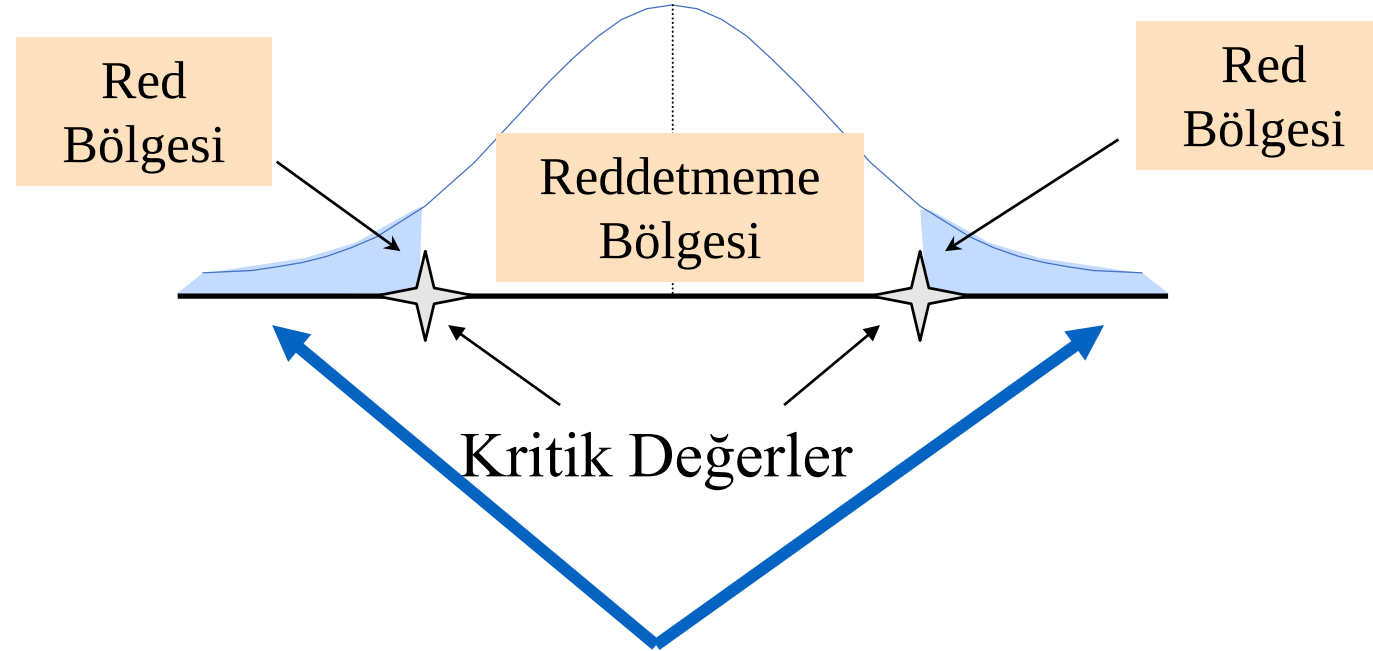
Test istatistiđi ve kritik deđerler



- Eđer örnek ortalaması belirtilen popölasyon ortalamasına çok yakınsa, sıfır hipotezi reddedilmez.
- Eđer örnek ortalaması belirtilen popölasyon ortalamasından çok uzaksa, sıfır hipotezi reddedilir.
- H_0 'ı reddetmek için ne kadarlık bir uzaklık “yeterince uzak” olarak kabul edilir?
- Bir test istatistiđinin kritik deđerini karar vermek için tam anlamıyla “son noktayı koyar” ve ne kadarlık bir uzaklıđın yeterince uzak olduđu sorusuna cevap verir.

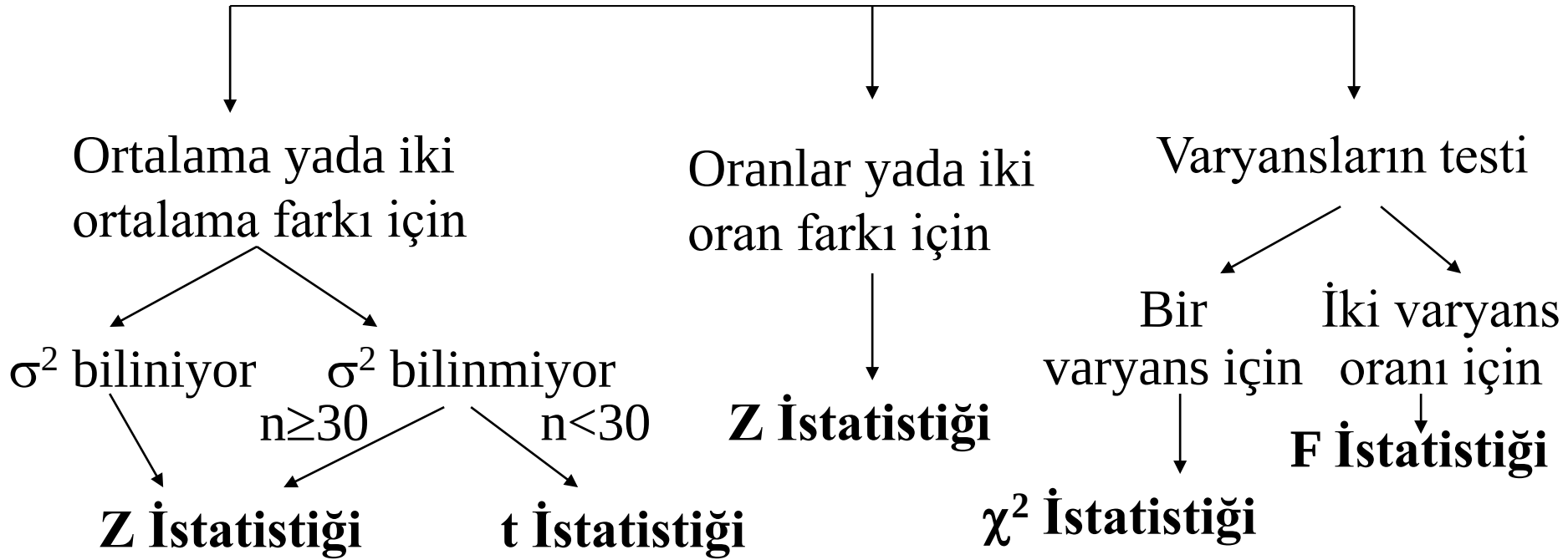
Test istatistiđi ve kritik deđer

Test İstatistiđinin Örnekleme Dağılımı



Örnekleme Dağılımının Ortalamasından “çok çok uzakta”

Hipotez Testinde Test İstatistiğinin Belirlenmesi



Hipotez testi ile karar vermede olası hatalar

- **Tip I Hata**

- Doğru bir sıfır hipotezinin reddi
- Ciddi bir hata türü olarak değerlendirilir
- Tip I hatanın olasılığı α 'dır
 - ✓ Testin önem derecesi olarak adlandırılır
 - ✓ Araştırmacı tarafından önceden belirlenir

- **Tip II Hata**

- Yanlış bir sıfır hipotezinin reddedilmemesindeki hata
- Tip II hatanın olasılığı β 'dır.

Hipotez testi ile karar vermede olası hatalar

Olası Hipotez Testi Çıktıları		
	Gerçek Durum	
Karar	H_0 Doğru	H_0 Yanlış
H_0 'ı kabul et	Hata yok Olasılık $1 - \alpha$	Tip II Hata Olasılık β
H_0 'ı reddet	Tip I Hata Olasılık α	Hata yok Gücü $1 - \beta$

Hipotez testi ile karar vermede olası hatalar

- **Güven katsayısı** $(1-\alpha)$, doğru olduğunda H_0 'ı reddetmeme olasılığıdır.
- Bir hipotez testinin **güven seviyesi** $(1-\alpha)*\%100$ 'dür.
- **İstatistiksel testin gücü** $(1-\beta)$, yanlış olduğunda H_0 'ı reddetmenin olasılığıdır.

Tip I & II hata ilişkileri

- Tip I ve Tip II hataları aynı anda oluşamazlar
 - Tip I hata sadece H_0 **doğru** olduğunda oluşur
 - Tip II hata sadece H_0 **yanlış** olduğunda oluşur

Tip I hata olasılığı (α) , o zaman

Tip II hata olasılığı (β) 

Tip II hatayı etkileyen faktörler

- Her şey eşit olduğunda,

- Hipotez parametresi ile gerçek değeri arasındaki fark $\downarrow$ β $\uparrow$

- α $\downarrow$ ise β $\uparrow$

- σ $\uparrow$ ise β $\uparrow$

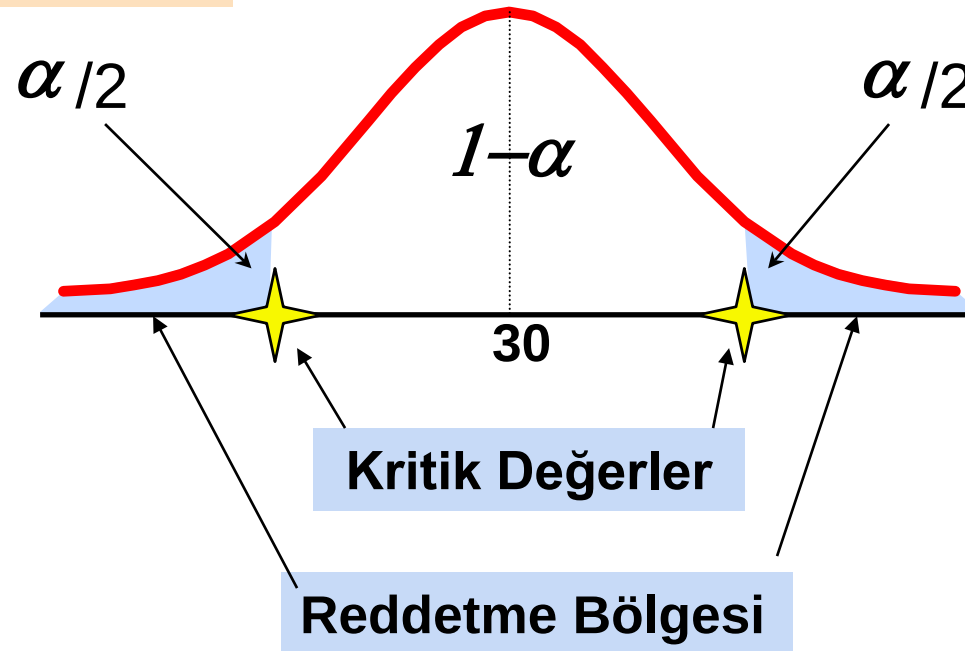
- n $\downarrow$ ise β $\uparrow$

Önem seviyesi ve reddetme bölgesi

$$H_0: \mu = 30$$

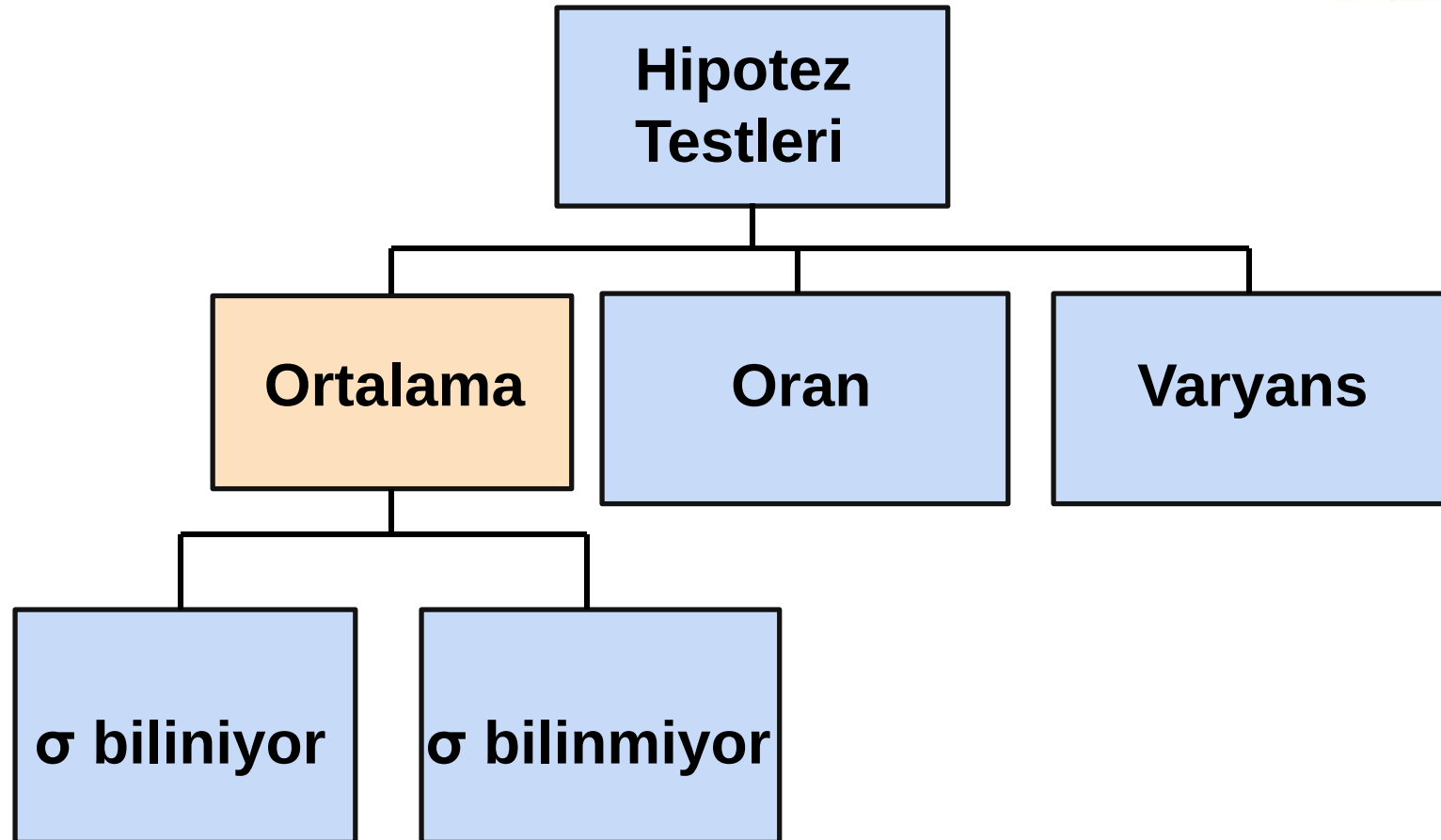
$$H_1: \mu \neq 30$$

Önem Seviyesi= α

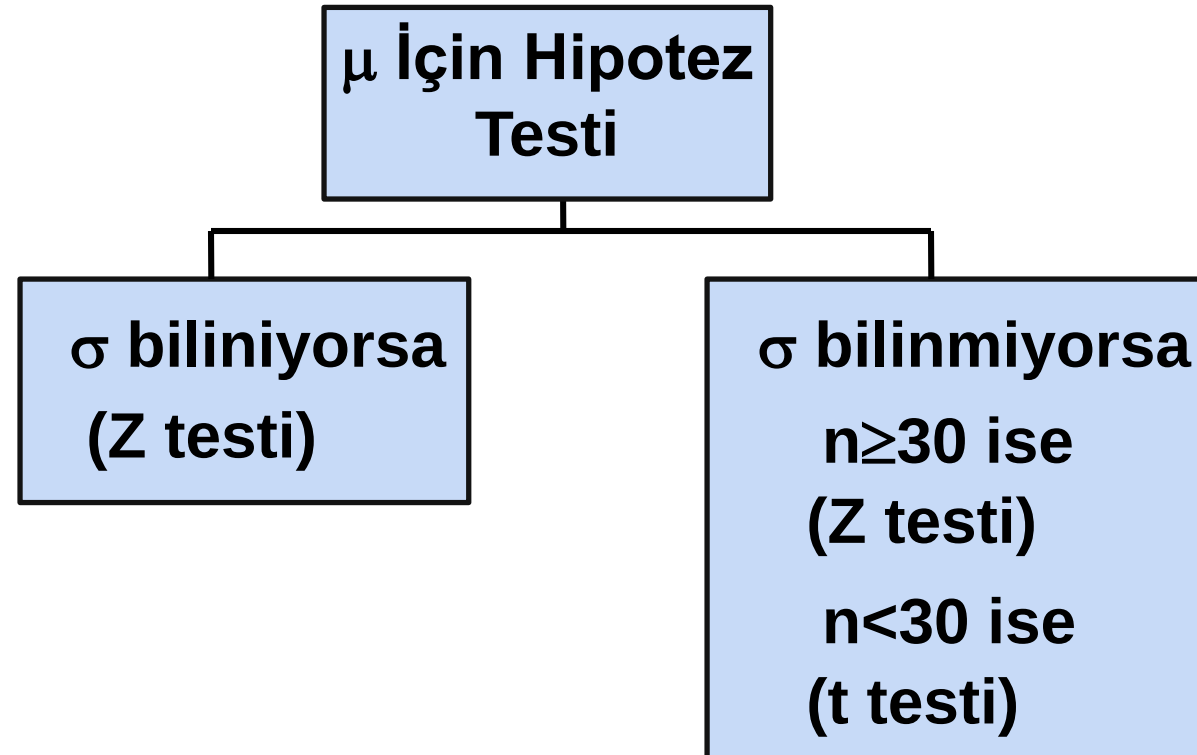


Her iki kuyruk bölgesinde de reddetme bölgesi olduğundan bu **iki kuyruklu bir test** tir.

Hipotez testleri

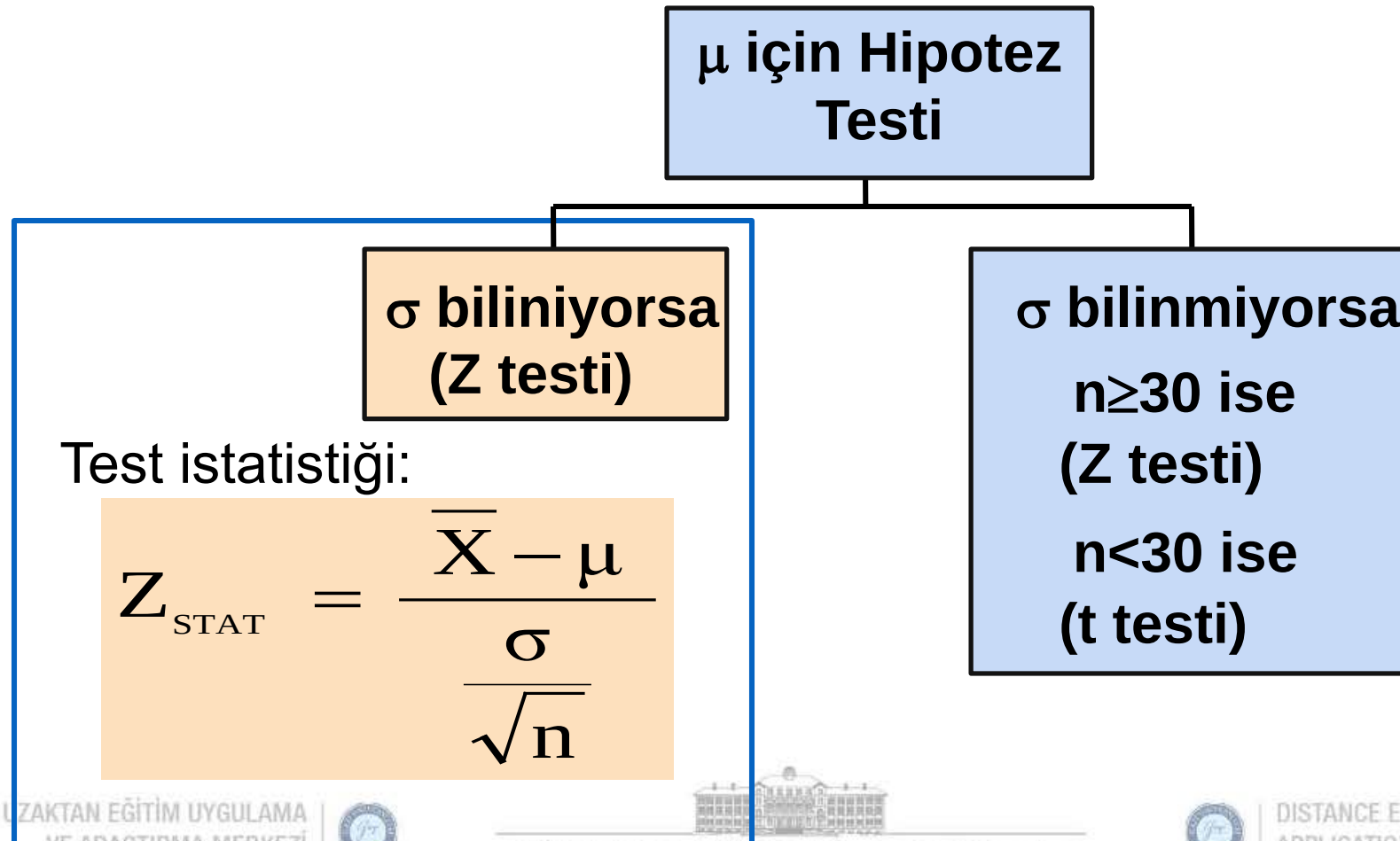


Ortalama için hipotez testi



Ortalama için Z hipotez testi (σ bilindiğinde)

- Örnek istatistiğini($\bar{X}$) test istatistiğine Z_{STAT} dönüştürelim



Hipotez testine kritik deęer yaklařımı

- σ bilindięinde ortalama iin iki kuyruklu test iin:
- rnek istatistięini ($\bar{X}$) test istatistięine (Z_{STAT}) dnřtr
- Tablodan veya bilgisayardan α nem seviyesine gre kritik Z deęerlerini belirle
- **Karar Kuralı:** Eęer test istatistięi reddetme blgesine dřyorsa, H_0 'ı reddet; deęilse H_0 'ı reddetme

Ortalama ile ilgili hipotez testleri

Varyans Bilindiğinde Ortalama için Hipotezlerin Test Edilmesi (Z-Testleri)

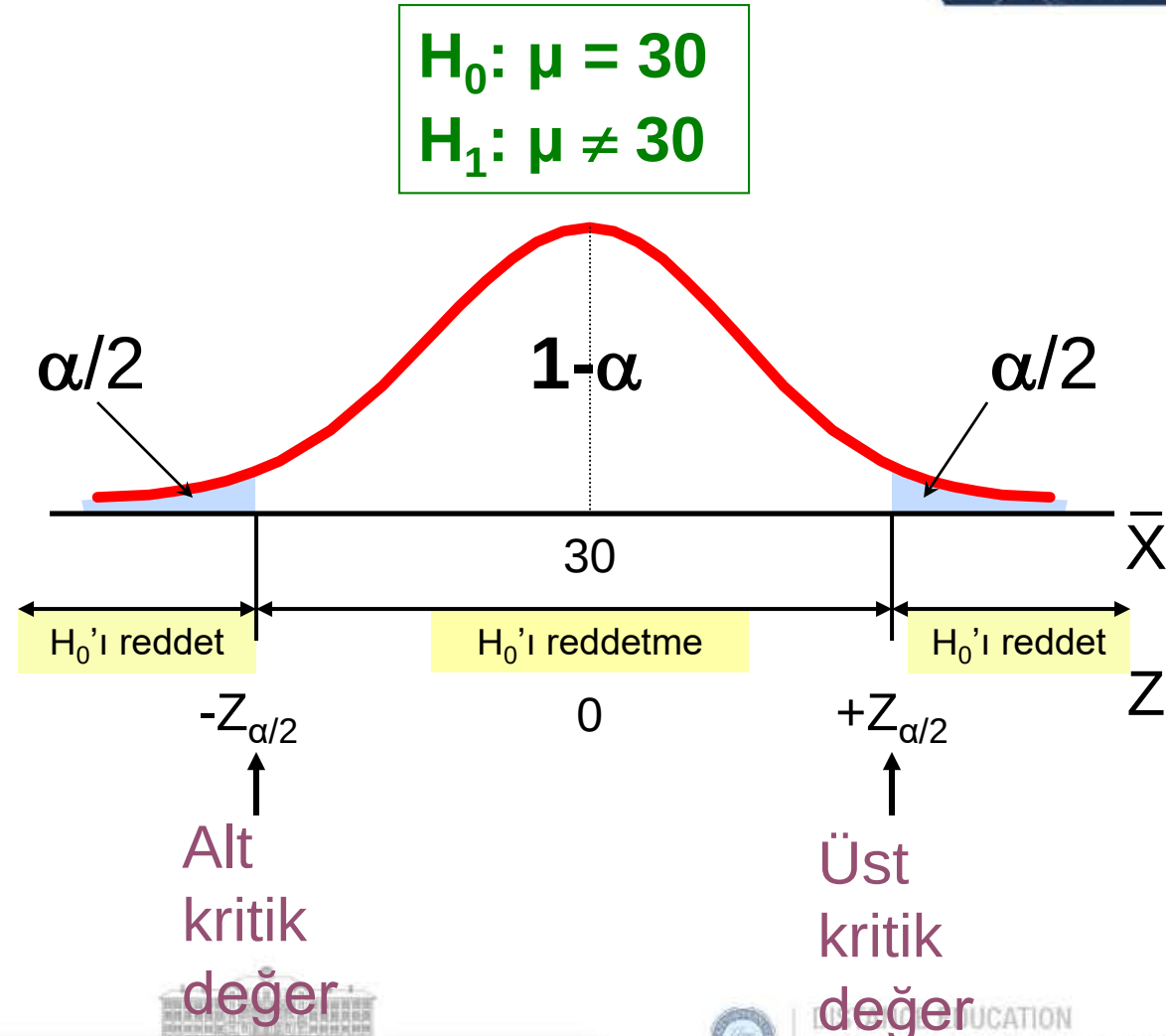
Sıfır Hipotezi: $H_0: \mu = \mu_0$

Test İstatistiği: $Z_0 = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$

Alternatif Hipotezler	P-Değeri	Sabit Seviye Testleri İçin Reddetme Kriteri
$H_1: \mu \neq \mu_0$	$ z_0 $ 'ın üstündeki ve $- z_0 $ 'ın altındaki olasılık. $P = 2[1 - \Phi(z_0)]$	$z_0 > z_{\alpha/2}$ veya $z_0 < -z_{\alpha/2}$
$H_1: \mu > \mu_0$	z_0 'ın üstündeki olasılık $P = 1 - \Phi(z_0)$	$z_0 > z_{\alpha}$
$H_1: \mu < \mu_0$	z_0 'ın altındaki olasılık $P = \Phi(z_0)$	$z_0 < -z_{\alpha}$

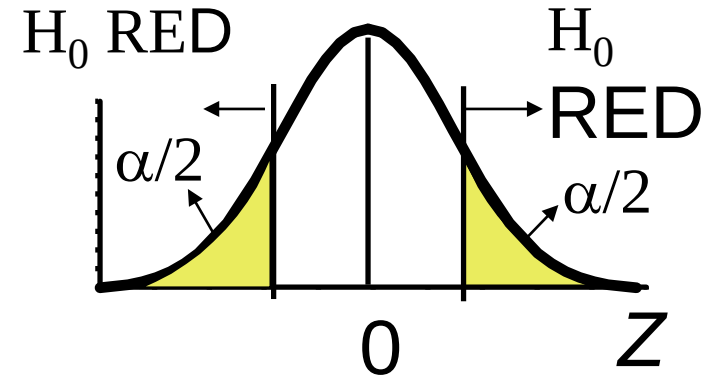
İki kuyruklu testler

- Reddetme bölgesini tanımlayan iki kesme değeri (kritik değerler) vardır

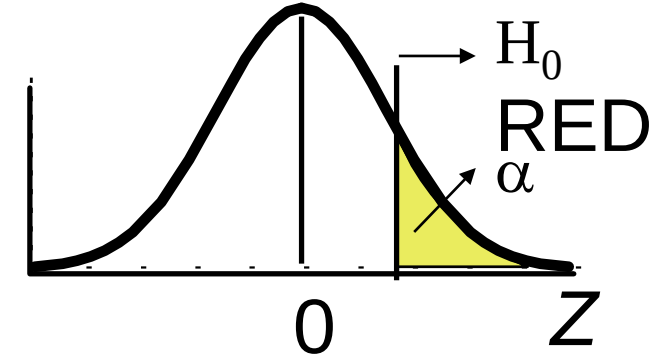


Hipotez Çiftleri:

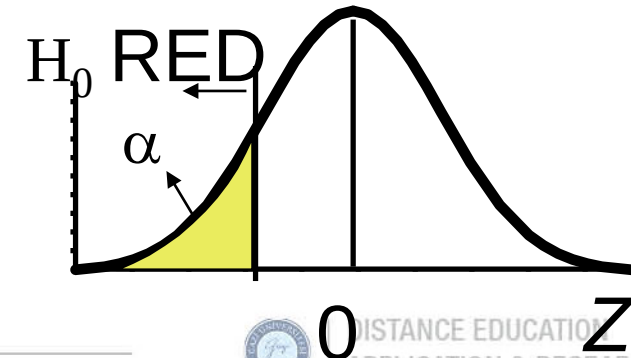
$$\left. \begin{array}{l} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{array} \right\} \text{ÇİFT TARAFLI TEST}$$



$$\left. \begin{array}{l} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{array} \right\} \text{TEK TARAFLI TEST} \\ \text{(Sağ taraf testi)}$$



$$\left. \begin{array}{l} H_0 : \mu = \mu_0 \\ H_1 : \mu < \mu_0 \end{array} \right\} \text{TEK TARAFLI TEST} \\ \text{(Sol taraf testi)}$$



Hipotez testinde 6 adım

1. Sıfır hipotezini , H_0 ve karşıt hipotezi , H_1 ifade et
2. Önem derecesi, α' yı ve örnek boyutunu , n' i seç
3. Uygun test istatistiği ve örnekleme dağılımını belirle
4. Reddetme ve reddetmeme bölgelerini ayıran kritik değerleri belirle

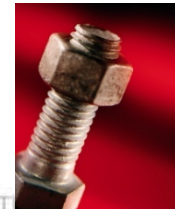
Hipotez testinde 6 adım

5. Veri topla ve test istatistiğinin değerini hesapla
6. İstatistiksel kararı ver ve yönetsel sonucu belirt. Eğer test istatistiği reddetmeme bölgesine düşüyorsa, sıfır hipotezi H_0 'ı reddetme. Eğer test istatistiği reddetme bölgesine düşüyorsa, sıfır hipotezini reddet. Yönetsel sonucu, problem bağlamında ifade et.

Hipotez testi örneği

**Üretilmiş bir somunun gerçek ortalama çapının 30mm olduğu iddiasını test edelim.
($\sigma = 0.8$ olduğunu varsayalım)**

1. Gerekli uygun sıfır hipotezi ve karşıt hipotezini belirleyelim
 - $H_0: \mu = 30$ $H_1: \mu \neq 30$ (İki kuyruklu bir testtir)
2. Gerekli önem seviyesini ve örnek boyutunu tanımlayalım
 - Bu test için $\alpha = 0.05$ ve $n = 100$ olarak seçildiğini varsayalım

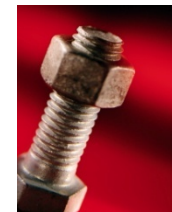


Hipotez testi örneği

3. Uygun tekniği belirle
 - σ biliniyor varsayımı mevcut, dolayısıyla Z testi uygulanır.
4. Kritik değerleri belirle
 - $\alpha = 0.05$ için kritik Z değerleri ± 1.96 'dir.
5. Veriyi topla ve test istatistiğini hesapla
 - Örnek sonuçları $n = 100$, $\bar{X} = 29.84$ 'dir.
($\sigma = 0.8$ bilindiği varsayımı mevcut)

Buradan test istatistiği:

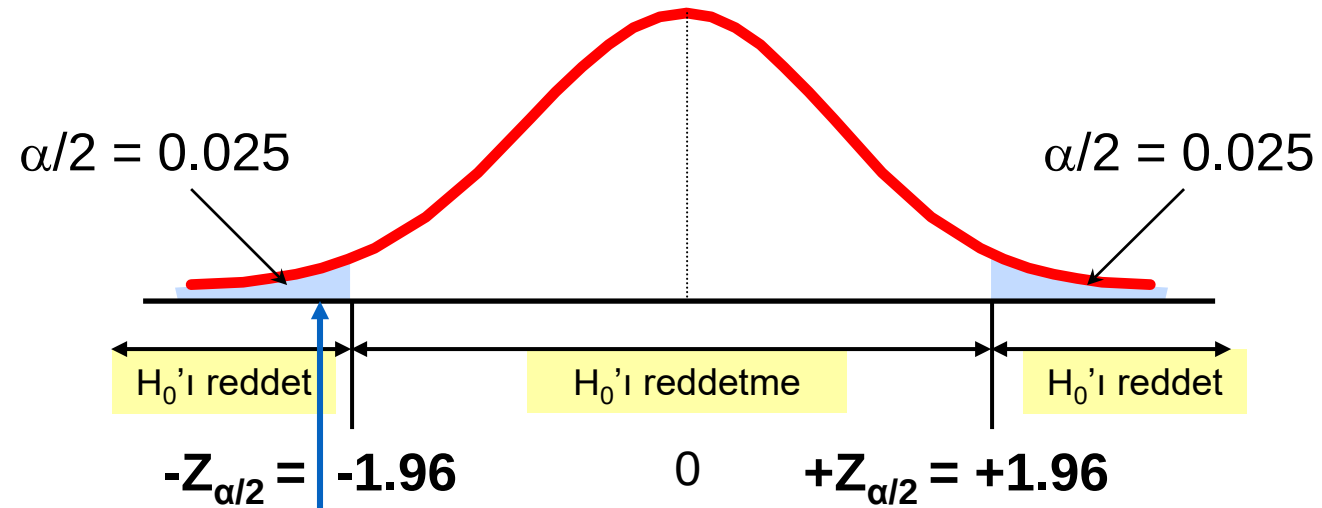
$$Z_{\text{STAT}} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{29.84 - 30}{\frac{0.8}{\sqrt{100}}} = \frac{-0.16}{0.08} = -2.0$$



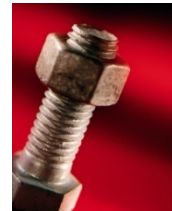
Hipotez testi örneği

- 6. Test istatistiği red bölgesinde mi?

Eğer $Z_{STAT} < -1.96$ veya $Z_{STAT} > 1.96$ ise H_0 'ı reddet; diğer durumda H_0 'ı reddetme

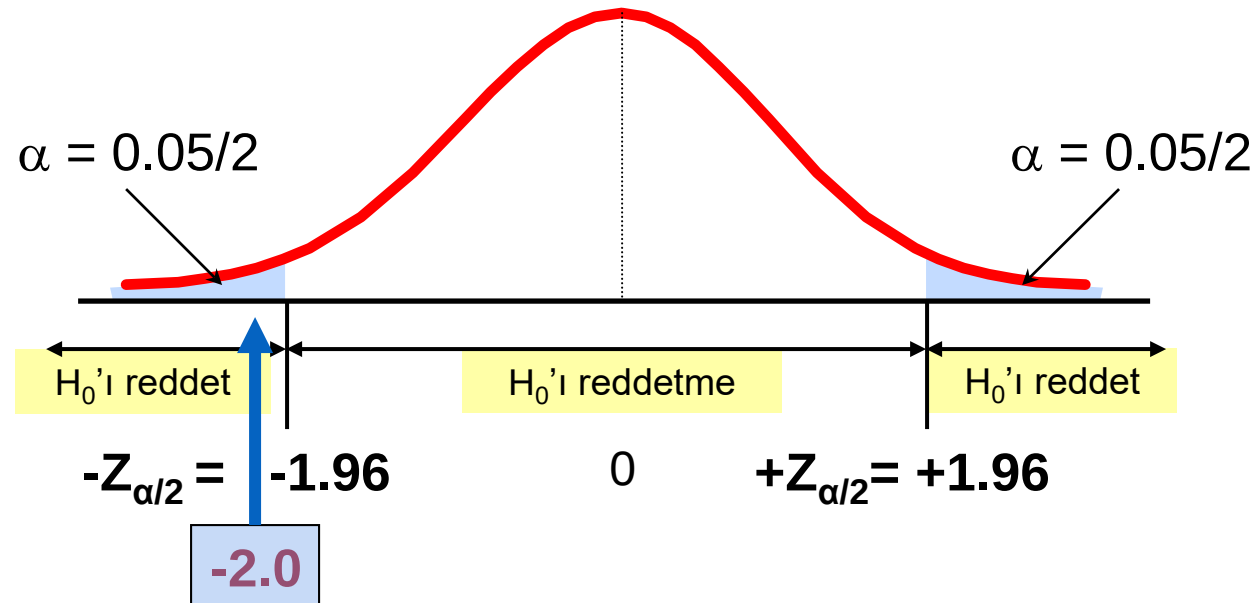


Burada, $Z_{STAT} = -2.0 < -1.96$, olduğundan test istatistiği red bölgesindedir

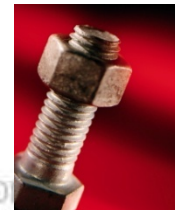


Hipotez testi örneği

6 (devamı). Bir karara varılır ve sonuç hakkında çıkarım yapılır.



$Z_{STAT} = -2.0 < -1.96$ olduğundan, sıfır hipotezi reddedilir ve üretilmiş somunların ortalama çapının 30mm eşit olmadığı konusunda yeterli delilin olduğu sonucuna varılır.



Test için p-değeri yaklaşımı

- p-değeri: H_0 doğru olarak verildiğinde elde edilen bir test istatistiğinin gözlemlenen örnek değerinden büyük veya ona eşit olması olasılığıdır.
 - p-değeri gözlemlenen önem derecesi olarak da adlandırılır
 - H_0 'ın reddedilebileceği en küçük α değeridir.

Test için p-değeri yaklaşımı: p-değerinin yorumlanması

- p-değerini, α ile kıyaslayalım

- Eğer $p\text{-değeri} < \alpha$, ise H_0 'ı reddet

- Eğer $p\text{-değeri} \geq \alpha$, ise H_0 'ı reddetme

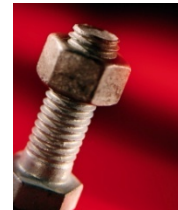
Hipotez testine 5 adımlı p-değeri yaklaşımı

1. Sıfır hipotezi, H_0 ve karşıt hipotezi, H_1 belirlenir.
2. Önem seviyesi, α ve örnek boyutu, n seçilir.
3. Uygun test istatistiği ve örnekleme dağılımı belirlenir
4. Veri toplanır ve test istatistiğinin değeri ve p-değeri hesaplanır
5. İstatistiksel karar verilir ve yönetsel çıkarım belirlenir. Eğer $p\text{-değeri} < \alpha$ ise H_0 reddedilir, diğer durumlarda H_0 reddedilmez. Problem bağlamında yönetsel çıkarım oluşturulur.

p-değeri hipotez testi örneği

Üretilmiş somunların ortalama gerçek çapının 30 mm olduğu iddiasını test edelim ($\sigma = 0.8$ olduğu varsayılıyor)

1. Uygun sıfır ve karşıt hipotezlerini belirleyelim
 - $H_0: \mu = 30$ $H_1: \mu \neq 30$ (Çift-kuyruklu bir testtir)
2. İstenen önem derecesini ve örnek boyutunu tanımlayalım
 - Bu test için $\alpha = 0.05$ olduğunu ve $n = 100$ olarak seçildiğini varsayalım



p-değeri hipotez testi örneği

3. Uygun tekniği belirleyelim

- σ 'nın bilindiği varsayılmaktadır, dolayısıyla bu bir Z testi olacaktır.

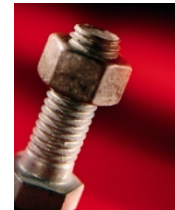
4. Veriyi toplayalım, test istatistiğini ve p-değerini hesaplayalım

- Örnek sonuçlarının

$n = 100$, $\bar{X} = 29.84$ olduğunu varsayalım ($\sigma = 0.8$ olarak bilindiği varsayımı mevcut)

Test istatistiği şu şekildedir:

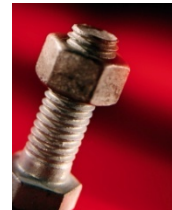
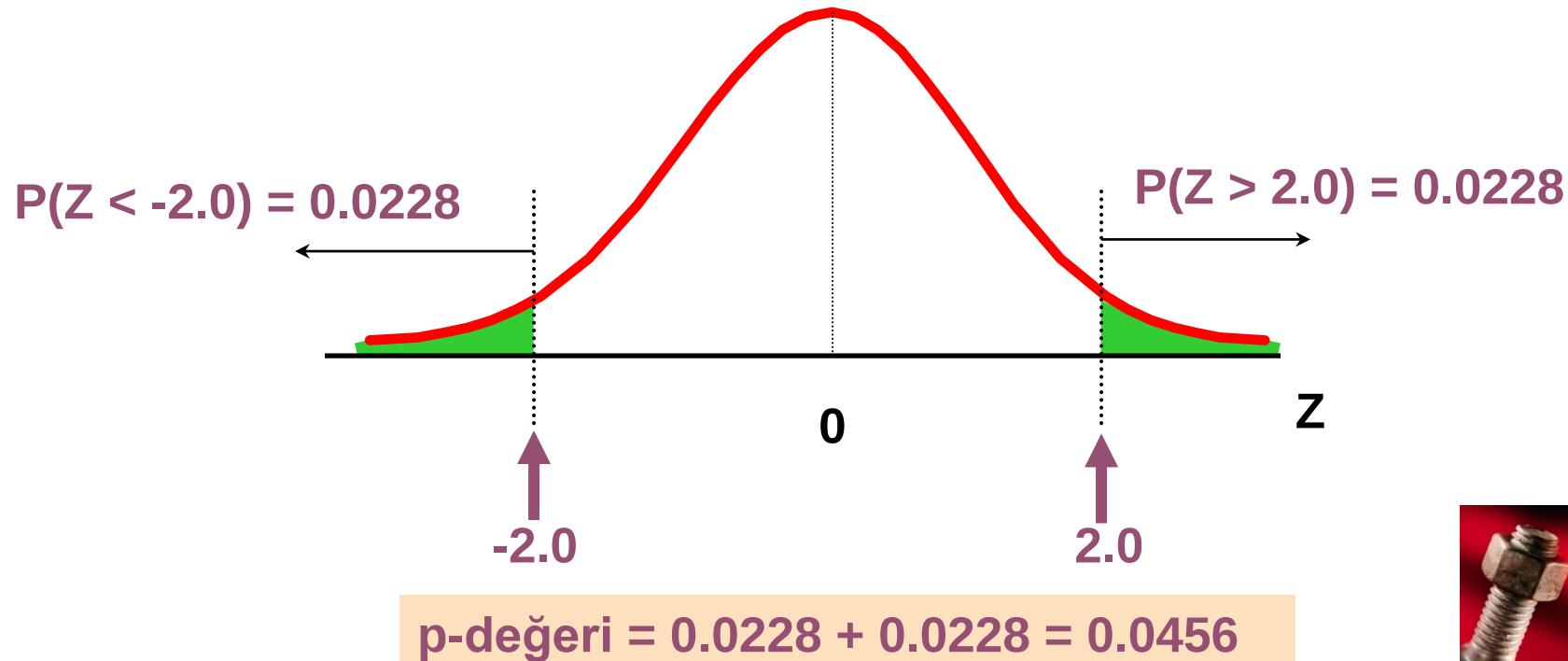
$$Z_{\text{STAT}} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{29.84 - 30}{\frac{0.8}{\sqrt{100}}} = \frac{-0.16}{0.08} = -2.0$$



p-değeri hipotez testi örneği: p-değerinin hesaplanması

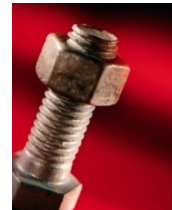
4. (devamı) p-değerini hesaplayalım.

- o H_0 doğru olduğunda Z_{STAT} değerini -2 olarak (veya ortalama değeri (0)'dan her iki yönde de ne kadar uzak bir değer) elde etmek ne kadar mümkündür ?



p-değeri hipotez testi örneği

- 5. $p\text{-değeri} < \alpha$?
 - $p\text{-value} = 0.0456 < \alpha = 0.05$ olduğundan H_0 'ı reddet
- 5. (devamı) Durum bağlamında yönetimsel çıkarımı yapalım.
 - Üretilmiş somunların ortalama çapının 30mm eşit olmadığı konusunda yeterli delilin olduğu sonucuna varılır.



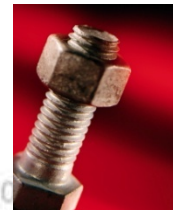
İki kuyruklu testlerin arasındaki ilişki ve güven aralıkları

- $\bar{X} = 29.84$, $\sigma = 0.8$ ve $n = 100$, için %95 güven aralığı şu şekildedir:

$$29.84 - (1.96) \frac{0.8}{\sqrt{100}} \quad \text{ile} \quad 29.84 + (1.96) \frac{0.8}{\sqrt{100}}$$

$$29.6832 \leq \mu \leq 29.9968$$

- Hipotezde geçen ortalama (30) aralık içerisinde yer almadığından, $\alpha = 0.05$ için sıfır hipotezi reddedilir



Örnek: Katı yakıtın yanma hızı

- Hava mürettebatı kaçış sistemi katı bir yakıt ile enerjisini sağlamaktadır.
- Bu katı yakıtın yanma hızı önemli bir ürün karakteristiğidir. Spesifikasyonlar yanma hızının 50 cm/sn olması gerektiğini belirtmektedir.
- Yanma hızına ait standart sapmanın $\sigma = 2 \text{ cm/sn}$ olduğu bilinmektedir. Deneyi gerçekleştiren kişi tip I hata olasılığını veya önem düzeyini $\alpha = 0,05$ olarak belirlemiş ve $n = 25$ gözlemden oluşan bir rassal örneklem almıştır.
- Yanma hızına ait örneklem ortalamasını $\bar{x} = 51,3 \text{ cm/sn}$ olarak hesaplamıştır. Elde edilecek sonuçlar nelerdir?

Örnek: Katı yakıtın yanma hızı

İlgilenilen parametre popülasyon ortalaması μ 'dür. Sıfır hipotezi $H_0: \mu = 50 \text{ cm/sn}$ olarak kurulur. Alternatif hipotez $H_1: \mu \neq 50 \text{ cm/sn}$ şeklinde geliştirilir. Test istatistiğinin değeri

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

şeklinde olacaktır. Eğer P -değeri 0,05'ten küçükse sıfır hipotezi H_0 reddedilecektir. Sabit önem seviyesi testini kullanmak için kritik bölgenin sınırları $z_{0,025} = 1,96$ ve $-z_{0,025} = -1,96$ şeklinde olacaktır.

$\bar{x} = 51,3$ ve $\sigma = 2$ olduğundan

$$z_0 = \frac{51,3 - 50}{2/\sqrt{25}} = 3,25$$

Sonuç olarak $z_0 = 3,25$ olduğundan $P - \text{değeri} = 2[1 - \Phi(3,25)] = 0,0012$ çıkmaktadır.

Sıfır hipotezi $H_0: \mu = 50$ 'yi $\alpha = 0,05$ önem seviyesinde reddedeceğiz.

Örnek

- Spor malzemeleri üreten bir firma, ürettiği olta iplerinin dayanma mukavemeti ortalamasının **15 kg/mm²**, standart sapmasının **0.5 kg/mm²** olduğunu açıklamıştır.
- Firmanın bu iddiasını test etmek isteyen bir tüketici örgütü firmanın üretiminden rastgele **50** olta ipi almış ve ortalamasını **14.8 kg/mm²** olarak belirlemiştir.
- %1 hata seviyesinde olta ipleri mukavemetinin **15 kg/mm²** ye eşit olup olmadığını test ediniz.

Çözüm

Tüketici örgütünün aldığı örneklerden elde ettiği 14.8 kg/mm² değeri ile firmanın beyan ettiği 15 kg/mm² değeri arasındaki 0.2 kg/mm² lik farklılığın tesadüfi mi olduğu, yoksa önemli bir farklılık mı olduğu test edilecektir.

i. $H_0: \mu = 15 \text{ kg/mm}^2$

$H_1: \mu \neq 15 \text{ kg/mm}^2$

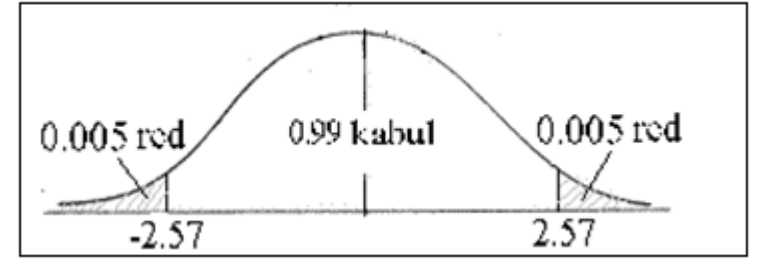
ii. $\alpha = 0.01 \quad \alpha/2 = 0.01/2 = 0.005 \quad TÇY \rightarrow Z_{\alpha/2} = Z_{0.005} = \pm 2.57$

iii. $Z_h = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} = \frac{14.8 - 15}{\sqrt{0.5^2/50}} = -2.828$

iv. $|Z_h| = |-2.828| > |Z_{0.005}| = |\pm 2.57| \rightarrow H_0 \text{ reddedilir.}$

$\alpha = 0.01$ hata {veya $(1-\alpha) = 0.99$ güven} seviyesinde olta ipleri mukavemeti ortalamasının 15 kg/mm² 'den farklı olduğu söylenebilir.

Başka bir ifadeyle, tüketici örgütünün gözlediği farklılık tesadüfi (şans eseri) değil, önemli bir farktır.



Örnek

- Bir kutu mısır gevreğinin ağırlığının **368 gr'dan fazla olduğu iddia edilmektedir**. Ayrıca $\sigma = 15$ gram olduğunu belirtmiştir. $n = 25$ kutuluk bir örnek alınmış ve $\bar{X} = 372.5$ gr. olarak bulunmuştur. 0.05 seviyesinde test ediniz.

Çözüm

$H_0: \mu = 368$

$H_1: \mu > 368$

$\alpha = 0.05$

$n = 25$

• Kritik değer:

Test İstatistiği:

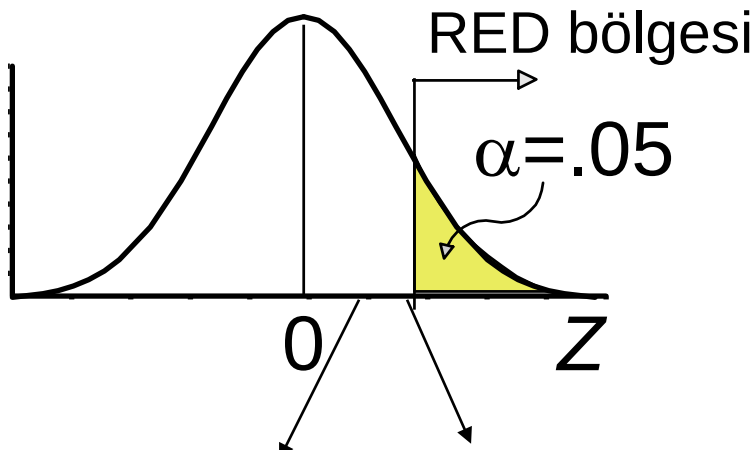
$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{372.5 - 368}{\frac{15}{\sqrt{25}}} = +1.5$$

Karar:

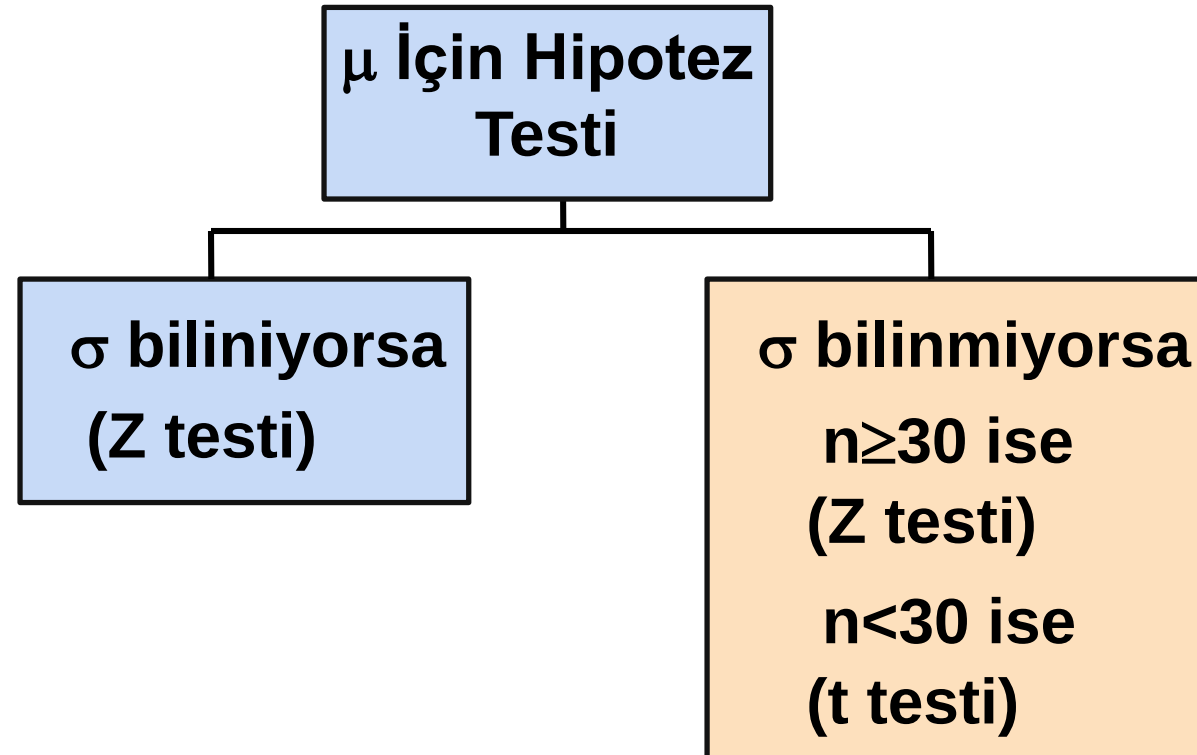
$\alpha = .05$ için H_0 hipotezi reddedilemez.

Yorum:

Ortalamanın 368 gr.dan fazla olduğuna dair yeterli kanıt yoktur.



Ortalama için hipotez testi



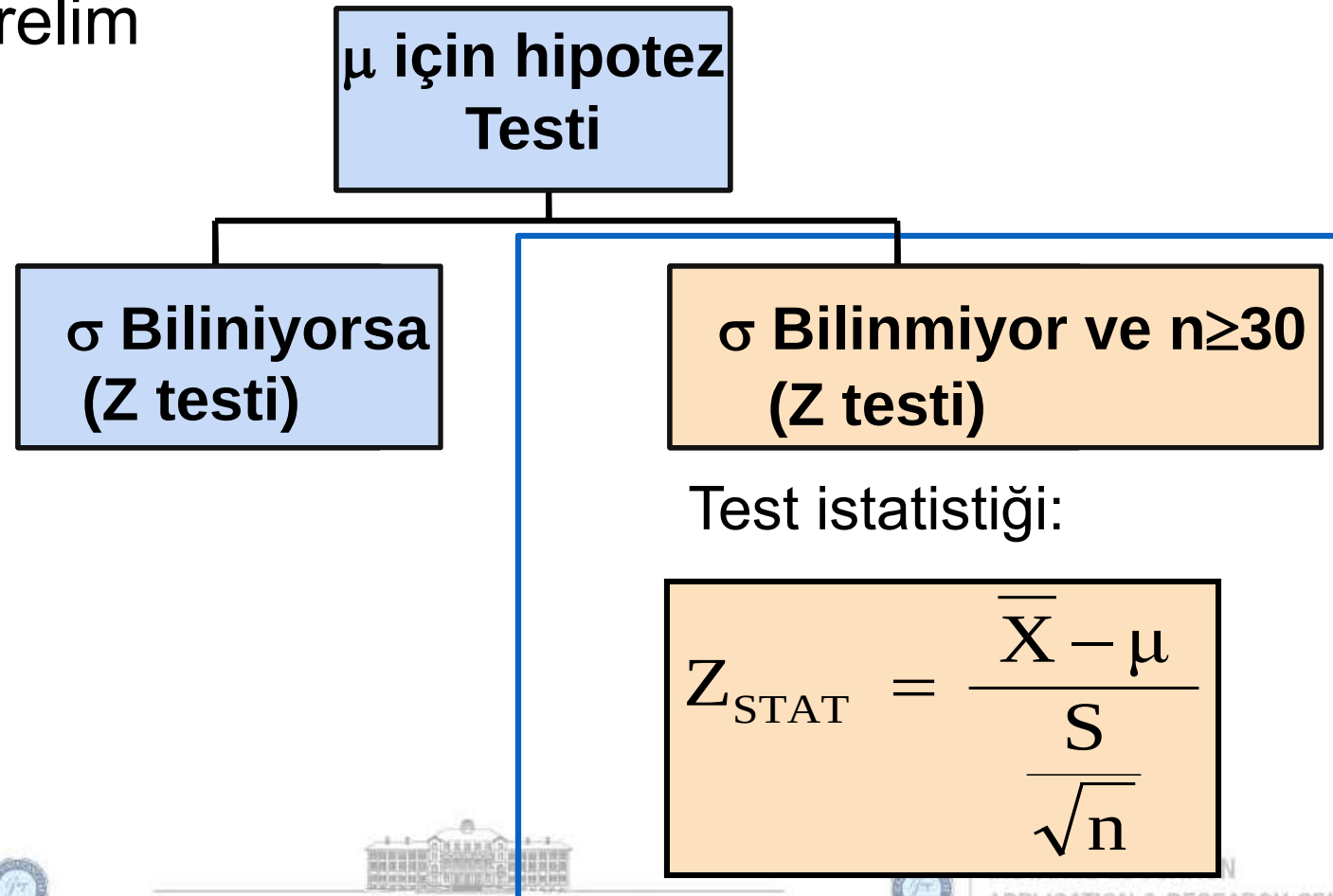
Hipotez testi:

σ bilinmiyor ve $n \geq 30$

- Popülasyonun normal dağıldığı ve σ^2 'nin bilindiği varsayımıyla $H_0: \mu = \mu_0$ sıfır hipotezi için testler geliştirdik.
- Ancak çoğu durumda σ^2 değeri bilinmez. Ayrıca popülasyonun normal dağılımla iyi bir şekilde modelleyebileceğimizden emin olamayız.
- Bu durumlarda eğer n yeterince büyükse söz gelimi $n \geq 30$ ise örneklem standart sapması s 'in popülasyon standart sapması σ yerine kullanılması test prosedürü üzerinde çok az etkiye sahiptir.
- Böylelikle varyansı bilinen bir normal dağılımın ortalaması için bir test yaparken popülasyonun dağılım formundan bağımsız şekilde geçerli olan **bilinmeyen σ^2 için büyük-örneklem test prosedürünü** kullanabiliriz.

Ortalama için Z testi (σ bilinmiyor ve $n \geq 30$)

- Örnek istatistiğini ($\bar{X}$) **test istatistiğine** t_{STAT} 'a dönüştürelim



Örnek

- Bir araştırmaya göre 10 yaşından küçük çocukların günde ortalama 6.6 saat televizyon izledikleri iddia edilmektedir. Bu iddianın doğruluğunu araştırmak için 10 yaşından küçük 100 çocuğun ailesiyle birlikte yapılan çalışma neticesinde çocukların günde izledikleri televizyon süresi ortalaması 6.1 saat ve standart sapması 2.5 saat olarak bulunmuştur. 0.01 hata payı ile araştırmanın doğruluğunu test ediniz.

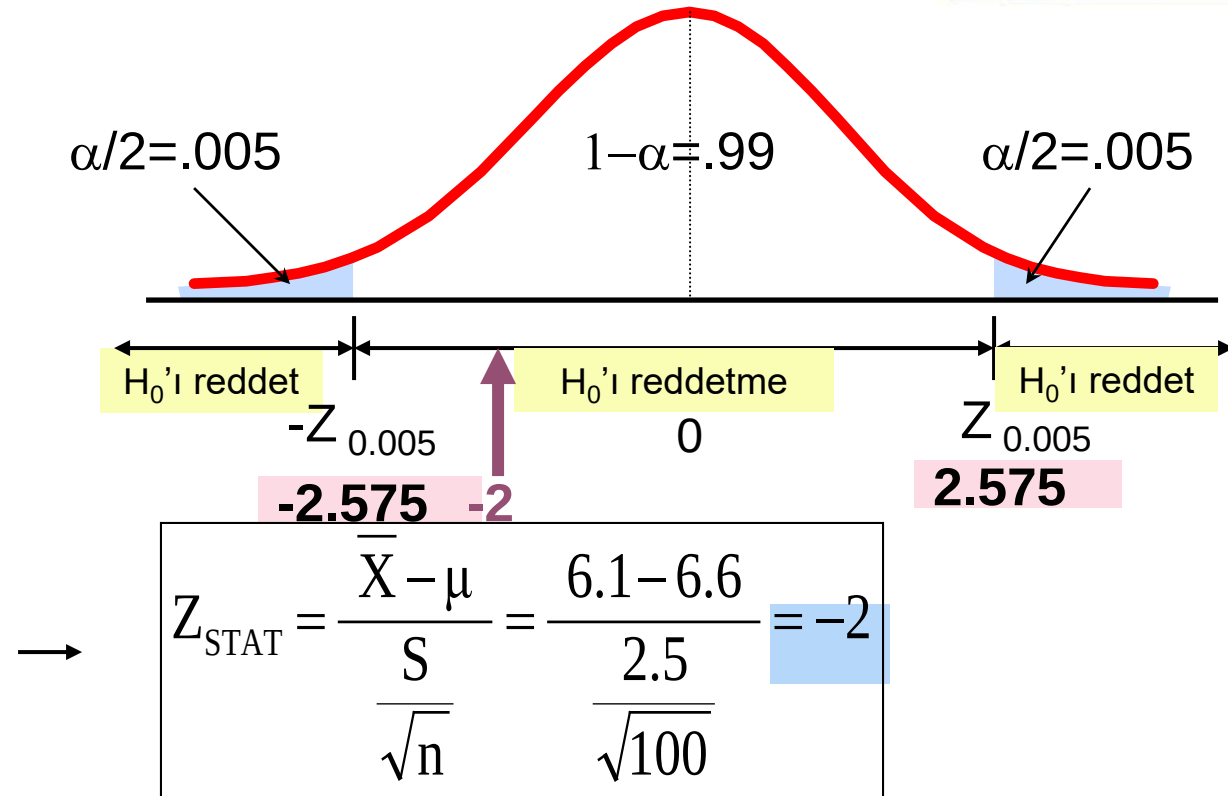
Örnek

$$H_0: \mu = 6.6$$

$$H_1: \mu \neq 6.6$$

- $\alpha = 0.01$
- $n = 100$
- σ bilinmiyor ve $n \geq 30$, o zaman **Z** istatistiği kullanılır
- Kritik Değer:

$$\pm Z_{0.005} = \pm 2.575$$



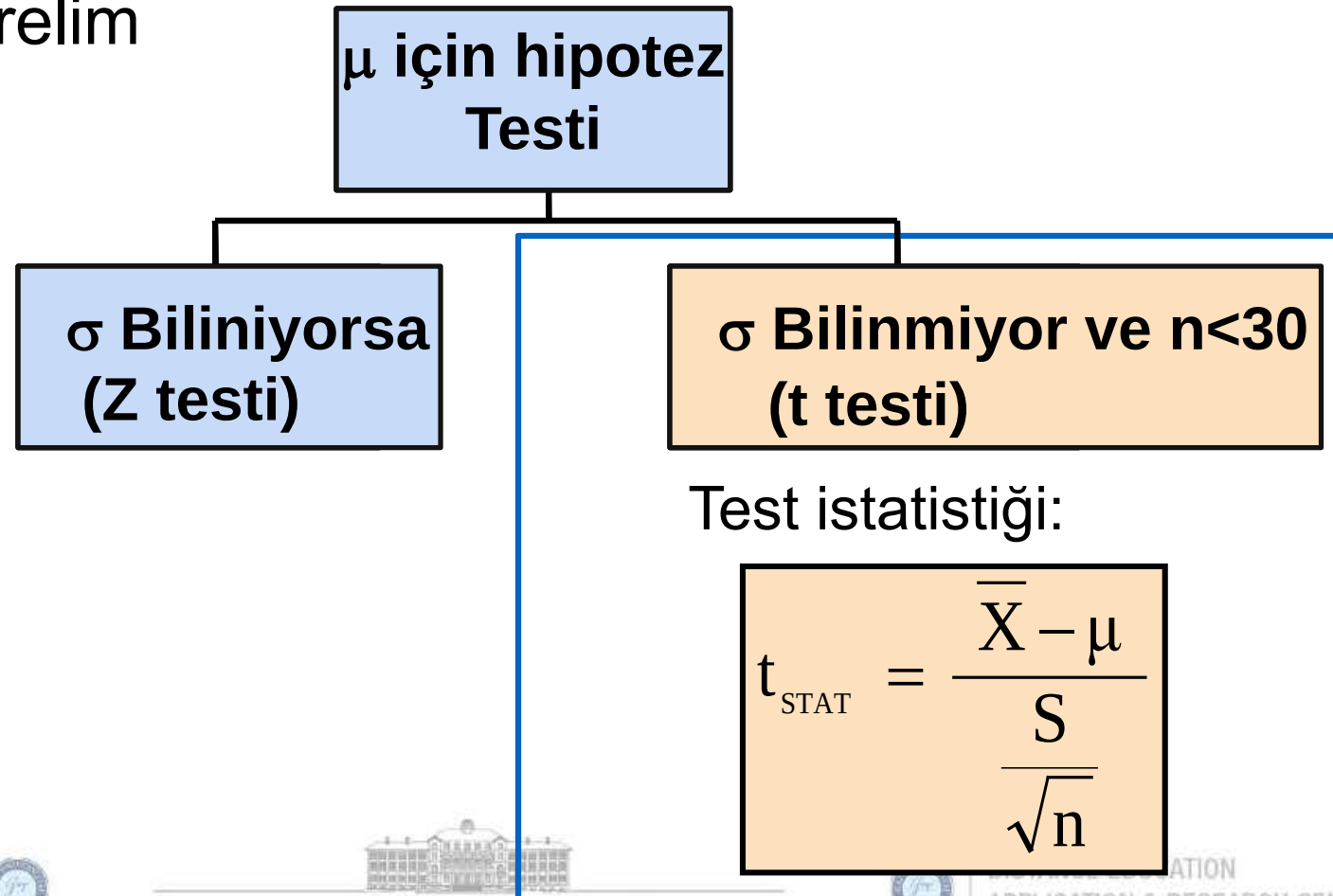
H_0 'ı reddetme: gerçek ortalama TV izleme süresinin 6.6 saatten farklı olduğuna dair yeterli delil yoktur

Hipotez testi: σ bilinmiyor ve $n < 30$

- Eğer popülasyon standart sapması bilinmiyorsa, bunun yerine örnek standart sapması S kullanılmalıdır.
- Bu değişimden dolayı, ortalama hakkındaki sıfır hipotezinin test edilmesi için eğer örneklem sayısı $n < 30$ ise Z dağılımı yerine t dağılımı kullanılmalıdır.
- t dağılımını kullanırken örnekleme yaptığınız popülasyonun normal dağılıma uyduğunu varsaymanız gerekmektedir.
- Tüm diğer adımlar, kavramlar ve çıkarımlar aynıdır.

Ortalama için t-testi (σ bilinmiyor ve $n < 30$)

- Örnek istatistiğini ($\bar{X}$) **test istatistiğine** t_{STAT} 'a dönüştürelim



t-testi



Varyans Bilinmediğinde Normal Dağılımın Ortalaması için Hipotez Testleri

Sıfır Hipotezi: $H_0: \mu = \mu_0$

Test İstatistiği: $T_0 = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$

Alternatif Hipotez

P-Değeri

Sabit Seviye Testler için Ret Bölgesi

$H_1: \mu \neq \mu_0$

$|t_0|$ üzerindeki olasılık ve
 $-|t_0|$ altındaki olasılık

$t_0 > t_{\alpha/2, n-1}$ veya $t_0 < -t_{\alpha/2, n-1}$

$H_1: \mu > \mu_0$

t_0 üzerindeki olasılık

$t_0 > t_{\alpha, n-1}$

$H_1: \mu < \mu_0$

t_0 altındaki olasılık

$t_0 < -t_{\alpha, n-1}$

Örnek: İki kuyruklu test (σ bilinmiyor ve $n < 30$)

New York'taki bir otel odasının gecelik ortalama fiyatının \$168 olduğu söylenmektedir. Bunun gerçekliğini belirlemek için, rassal olarak 25 otel alınmış ve \$172.50'lık bir $\bar{X}$ ve \$15.40'lık bir S ile sonuçlanmıştır. Gerekli olan hipotezleri $\alpha = 0.05$ önem seviyesinde test edin.

(Popülasyonun normal dağılıma uyduğunu varsayın)



$$H_0: \mu = 168$$

$$H_1: \mu \neq 168$$

Örnek çözümü: İki kuyruklu t-testi

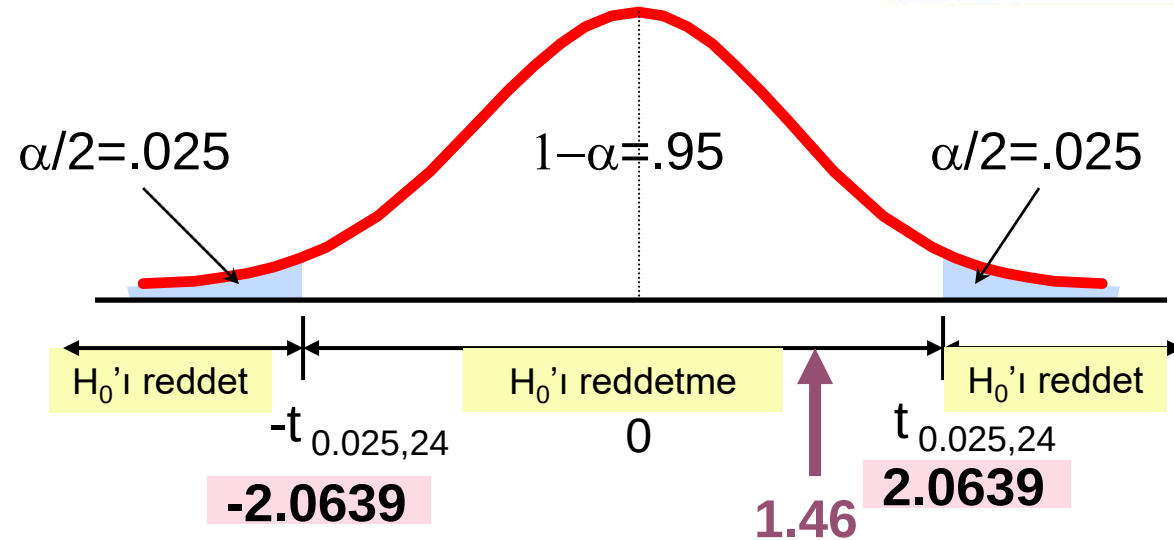
$$H_0: \mu = 168$$

$$H_1: \mu \neq 168$$

- $\alpha = 0.05$
- $n = 25, sd = 25-1=24$
- σ bilinmiyor, o zaman **t** istatistiği kullanılır

- Kritik Değer:

$$\pm t_{0.025,24} = \pm 2.0639$$



$$t_{STAT} = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} = \frac{172.50 - 168}{\frac{15.40}{\sqrt{25}}} = 1.46$$

H_0 'ı reddetme: gerçek ortalama maliyetin \$168'den farklı olduğuna dair yeterli delil yoktur

Excel'den elde edilen p-değerini kullanarak iki kuyruklu t-testi örneği

- Bu bir t testi olduğundan hesaplama yardımı olmadan p-değerini hesaplayamayız.
- Aşağıdaki Excel çıktısı bunu yapmaktadır:

Ortalama Hipotezi için t Testi			
Veri			
Sıfır Hipotezi $\mu=$	\$	168,00	
Önem Derecesi		0,05	
Örnek Boyutu		25	
Örnek Ortalaması	\$	172,50	
Örnek Standart Sapması	\$	15,40	
Ara İşlemler			
Ortalamanın Standart Hatası	\$	3,08	=B8/SQRT(B6)
Serbestlik Derecesi		24	=B6-1
t test istatistiği		1,46	= (B7-B4)/B11
İki Kuyruklu Test			
Alt Kritik Değer		-2,0639	=-TINV(B5,B12)
Üst Kritik Değer		2,0639	=TINV(B5,B12)
p-değeri		0,157	=TDIST(ABS(B13),B12,2)
Sıfır hipotezini Reddetme			=IF(B18<B5, "Sıfır Hipotezini Reddet", "Sıfır Hipotezini Reddetme")

p-değeri > α
Öyleyse H_0 'ı
reddetme

İki kuyruklu testlerin güven aralığı ile bağlantısı

- $\bar{X} = 172.5$, $S = 15.40$ ve $n = 25$, μ için %95 güven aralığı:

$$172.5 - (2.0639) 15.4/\sqrt{25} \text{ den } 172.5 + (2.0639) 15.4/\sqrt{25} \text{ 'e}$$

$$166.14 \leq \mu \leq 178.86$$

- Aralık hipotezi yapılan ortalamayı (168), içerdiğinden $\alpha = 0.05$ için sıfır hipotezi reddedilmez.

Örnek: Golf sopası tasarımı

- Hafif ve dayanıklı malzemelerdeki artış golf sopalarının tasarımında ve üretiminde bir devrime neden olmuştur.
- İnce yüzlü ve baş kısmının içi boş olan sopalar özellikle yeni başlayan oyuncular için daha uzun mesafeli vuruşlara neden olmaktadır.
- Belirli bir golf sopası üreticisinden 15 tane sopa rassal olarak seçilmiş ve belirli bir tasarım özelliği test edilmek istenmektedir.
- Burada test ölçüm ortalamasının 0,82 değerini aştığını iddiasını destekleyecek bir bulgu ($\alpha = 0,05$) olup olmadığını belirlemektir. Gözlem değerleri aşağıdaki gibidir:

0.8411	0.8191	0.8182	0.8125	0.8750
0.8580	0.8532	0.8483	0.8276	0.7983
0.8042	0.8730	0.8282	0.8359	0.8660

Örneklem ortalaması örneklem standart sapması $\bar{x} = 0,83725$ ve $s = 0,02456$ şeklindedir. Verilerin normal dağıldığı olasılık grafiğinden (probability plot) görülmektedir. Deneyin amacı popülasyon ortalamasının 0,82 değerini aştığını göstermek olduğundan tek-uçlu alternatif hipotez uygundur.

İlgilenilen parametre popülasyon ortalaması μ .

Sıfır Hipotezi: $H_0: \mu = 0,82$

Alternatif Hipotez: $H_1: \mu > 0,82$

Test İstatistiği

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

Eğer P -Değeri 0,05 değerinden küçükse H_0 hipotezini reddet.

Hesaplamalar: $\bar{x} = 0,83725$, $s = 0,02456$, $\mu_0 = 0,82$ ve $n = 15$ olduğundan

$$t_0 = \frac{0,83725 - 0,82}{0,02456/\sqrt{15}} = 2,72$$

Serbestlik derecesi 14 olan t dağılımı için t -tablosuna baktığımızda $t_0 = 2,72$ değeri 2,64 ($\alpha = 0,01$) ve 2,977 ($\alpha = 0,005$) arasına düşmektedir. Bu tek uçlu bir kuyruk olduğu için biliyoruz ki P -değeri bu iki değer arasındadır yani $0,005 < P < 0,01$. Böylelikle $P < 0,05$ olduğundan H_0 hipotezini reddedeceğiz ve popülasyon ortalamasının 0,82 değerini aştığı sonucuna ulaşacağız.

Örnek

- Bir bölümde okuyan öğrencilerin not ortalamasının 70 olduğu bölüm yöneticileri tarafından iddia edilmektedir. Bu iddianın doğru olup olmadığını test etmek üzere Bölüm öğrencileri arasından rasgele seçilen 26 öğrencinin notları tespit edilmiş ve ortalaması 67, varyansı 9 olarak hesaplanmıştır.
- %99 güven (%1 hata) seviyesinde Bölüm yöneticilerinin iddiasının doğru olduğu söylenebilir mi?
- *(Başka bir ifadeyle, sözü edilen bölümdeki öğrencilerin not ortalamalarının 70 den az olduğu söylenebilir mi?)*

Çözüm

Anakütle varyansı bilinmediğinden ve örnek hacmi 30'dan küçük olduğundan dolayı küçük örneklerin analizi için geliştirilen t dağılımı ile hipotez testi yapılacaktır.

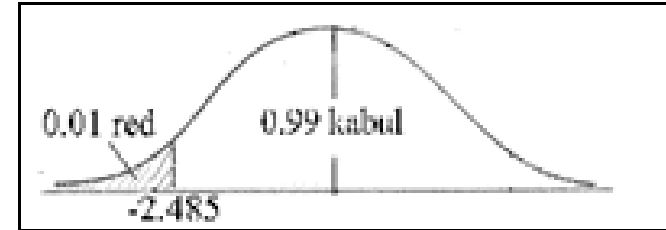
i. $H_0: \mu=70$ puan

$H_1: \mu<70$ puan

ii. $1-\alpha=0.99 \rightarrow \alpha=0.01$ $SD=n-1=26-1=25$ $TTY \rightarrow t_{0.01; 25}=-2.485$

iii. $t_h = \frac{\bar{X} - \mu}{S_{\bar{X}}} = \frac{67 - 70}{\sqrt{\frac{9}{26}}} = -5.099$

iv. $|t_h| = |-5.099| > |t_{0.01; 25}| = |-2.485| \rightarrow H_0$ reddedilir.



v. $\alpha=0.01$ hata {veya $(1-\alpha)=0.99$ güven} seviyesinde sözü edilen bölümdeki öğrencilerin notları ortalamasının 70 puandan küçük olduğu söylenebilir. Başka bir ifadeyle, gözlenen farklılık tesadüfi (*şans eseri*) değil, önemli bir farktır.

Örnek

Örnek



- Bir konserve fabrikasının imal ettiği konservelerin üzerinde brüt 455 gr yazmaktadır. Bu konservelerin brüt ağırlıkları ile ilgili bir karar vermek üzere rasgele seçilen 17 kutunun ortalama ağırlığı 450 gr ve standart sapması 13 gr bulunmuştur. Brüt ağırlığın 455 gr olmadığını 0.05 önem seviyesinde söyleyebilir misiniz?

Çözüm

Çözüm



$$H_0 : \mu = 455$$

$$H_1 : \mu \neq 455$$

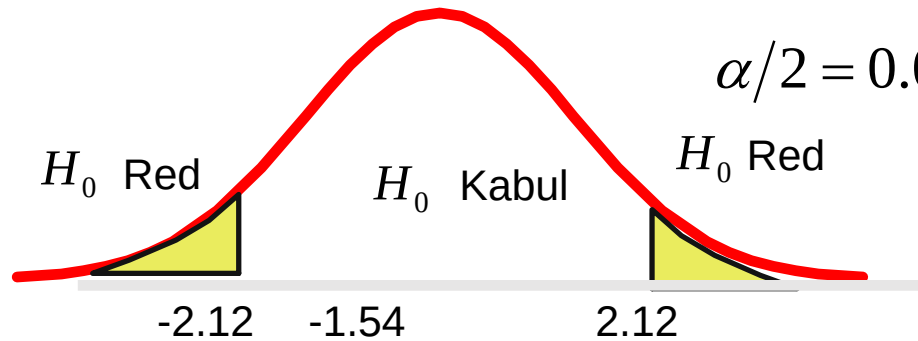
$$n = 17$$

$$\bar{X} = 450gr.$$

$$s = 13gr.$$

$$s.d. = n - 1 = 16$$

$$t_{tab} = \pm 2.12$$



$$t_h = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}} = \frac{450 - 455}{\frac{13}{\sqrt{17}}} \cong 1.54$$

Tek kuyruklu testler

- Çoğu durumda, karşıt hipotezler belirli bir yöne odaklanırlar

$$H_0: \mu = 3$$

$$H_1: \mu < 3$$



3 ortalamasının altındaki alt kuyruğa odaklandığından bu bir **alt**-kuyruk testidir

$$H_0: \mu = 3$$

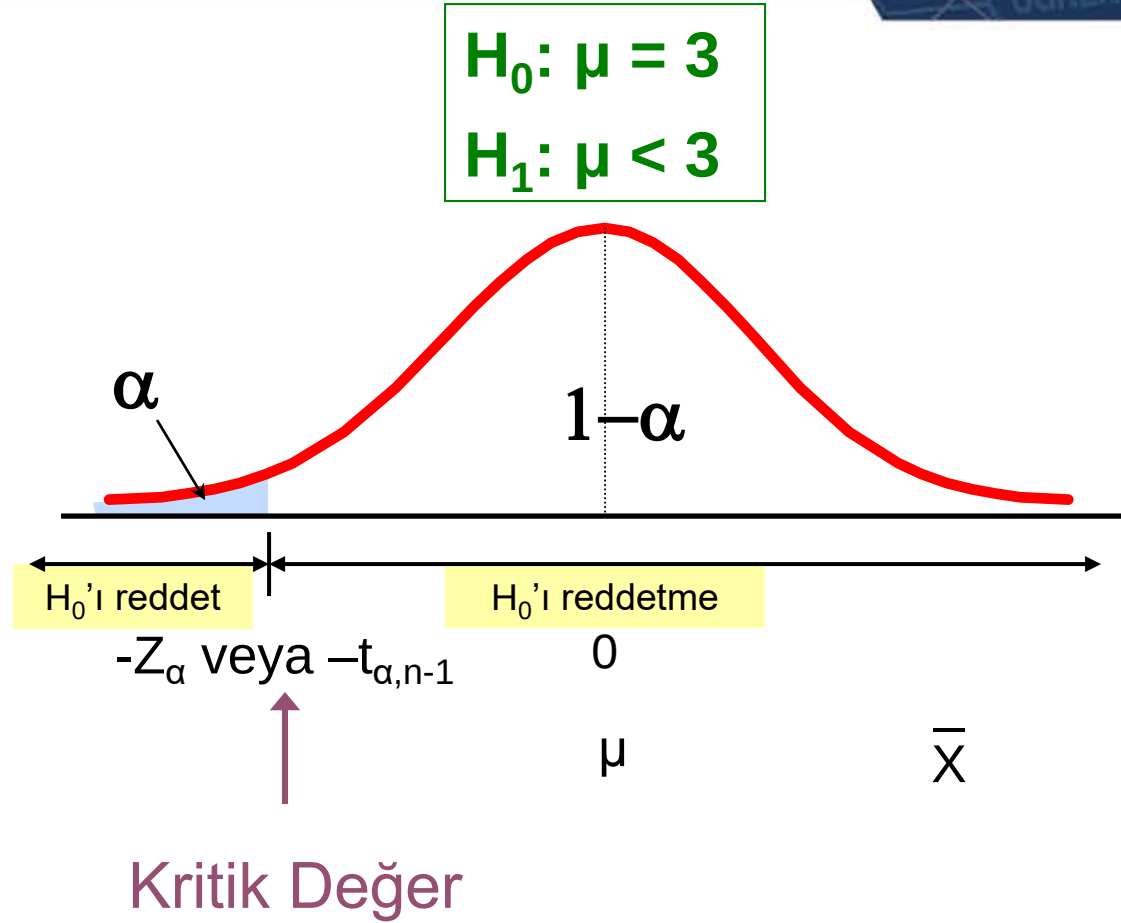
$$H_1: \mu > 3$$



3 ortalamasının üstündeki üst kuyruğa odaklandığından bu bir **üst**-kuyruk testidir

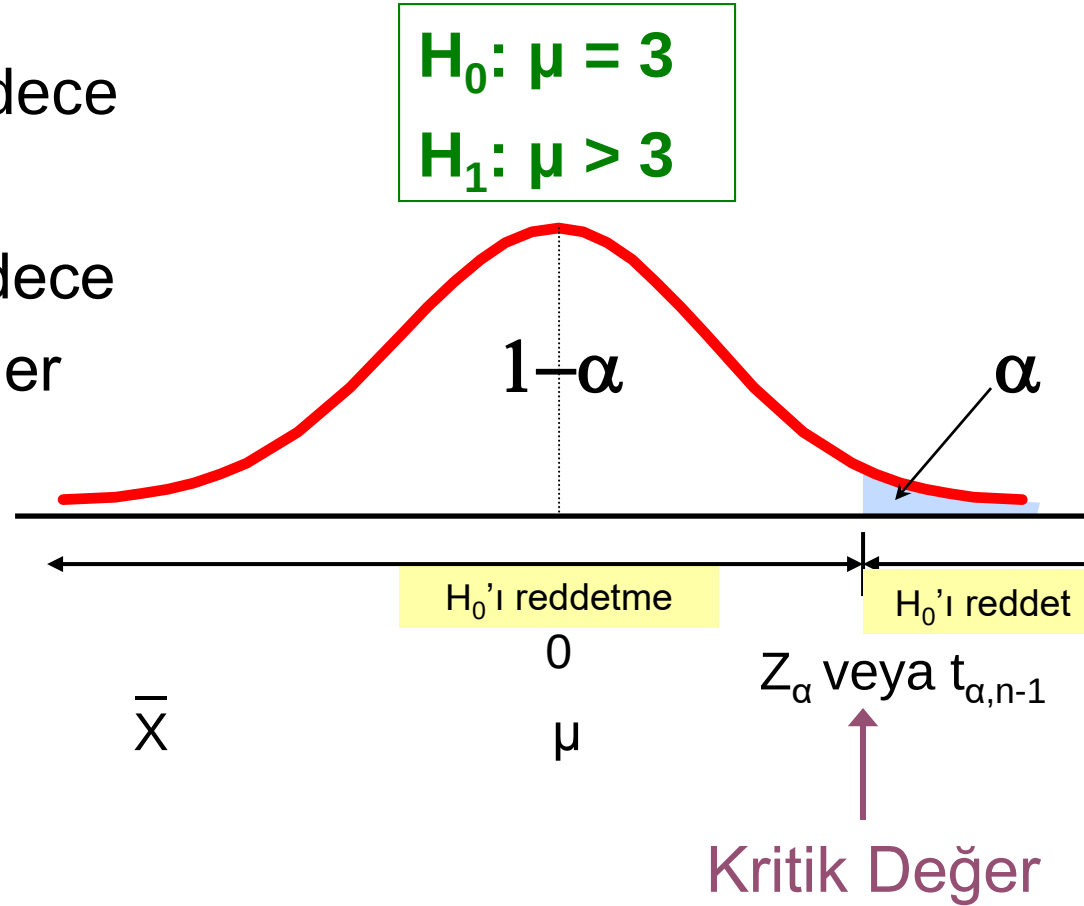
Alt-kuyruk testleri

- Red bölgesi sadece tek bir kuyrukta olduğundan sadece tek bir kritik değer mevcuttur.



Üst-kuyruk testleri

- Red bölgesi sadece tek bir kuyrukta olduğundan sadece tek bir kritik değer mevcuttur.



Örnek: Ortalama için üst-kuyruk t-testi (σ bilinmiyor ve $n < 30$)

Bir telefon endüstrisi yöneticisi müşterilerin aylık telefon faturalarının yükseldiğini düşünmektedir, ve şu anda aylık ortalama \$52'dır. Şirket bu iddiayı test etmek istemektedir. (Normal dağılan bir popülasyon olduğunu varsayın)



Hipotez testi oluşturalım:

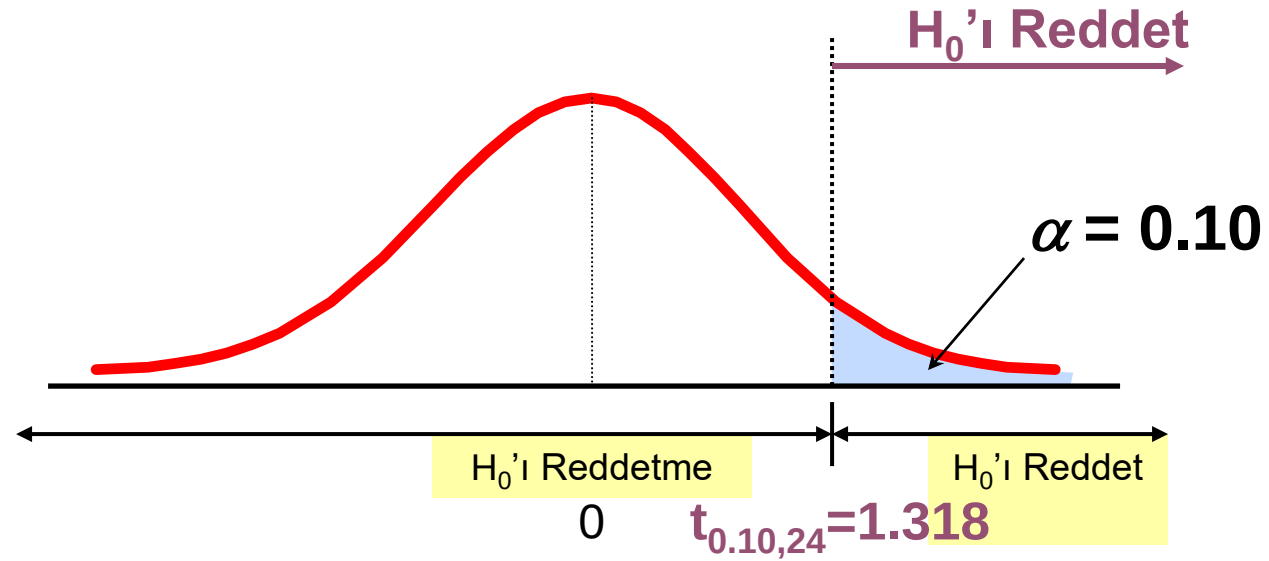
$H_0: \mu = 52$ ortalama aylık \$52'a eşittir

$H_1: \mu > 52$ ortalama aylık \$52'ın üzerinde**dir**
(yani, yöneticinin iddiasını destekleyen yeterli kanıt mevcuttur)

Örnek

- Bu test için $\alpha = 0.10$ seçildiğini ve $n = 25$ olduğunu varsayalım.

Red bölgesini bulalım:



Eğer $t_{STAT} > 1.318$ ise H_0 'ı reddet

Örnek

Örneği elde et ve test istatistiğini hesapla

Şu sonuçları veren bir örneğin alındığını varsayalım: $n = 25$, $\bar{X} = 53.1$,
ve $S = 10$

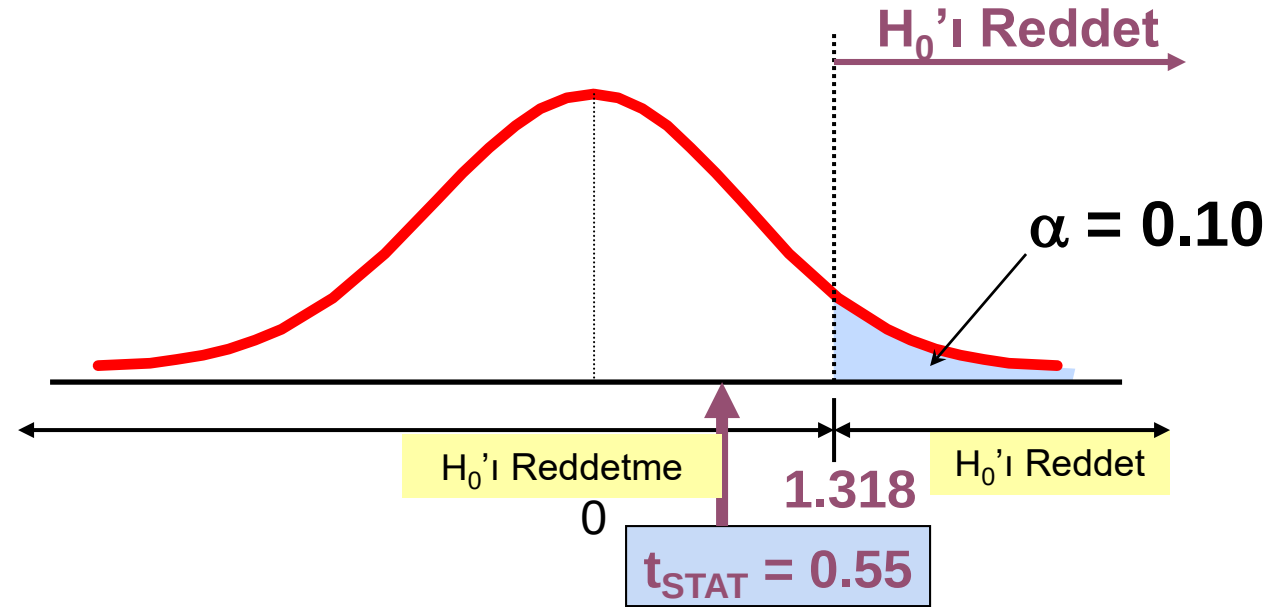
o Test istatistiği şöyledir:



$$t_{\text{STAT}} = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} = \frac{53.1 - 52}{\frac{10}{\sqrt{25}}} = 0.55$$

Örnek

Bir karara varılır ve sonuçlardan bir çıkarım yapılır:

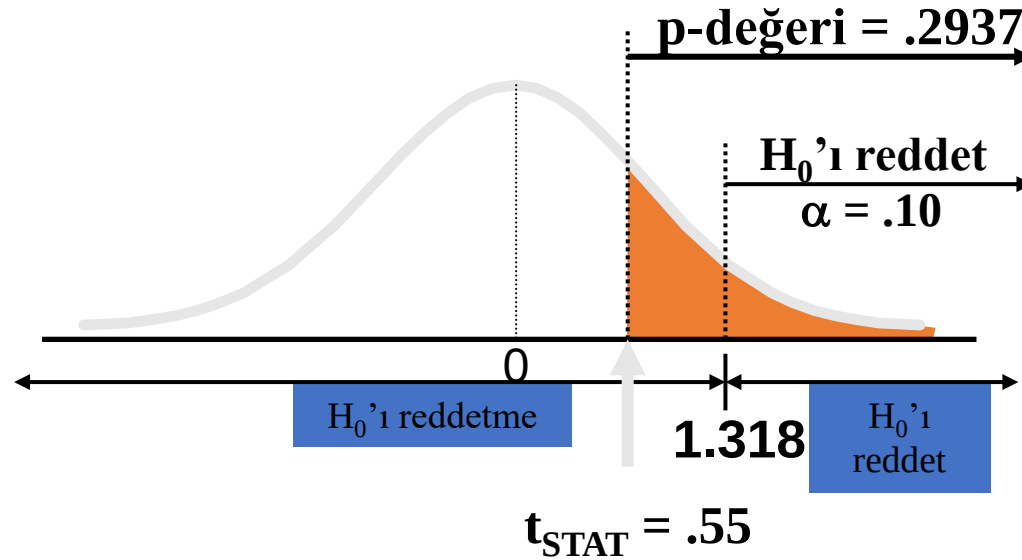


$t_{STAT} = 0.55 \leq 1.318$ olduğundan H_0 'ı reddetme

ortalama fatura tutarının \$52'in üzerinde olduğuna dair yeterli delil yoktur

Örnek: Test için p-değerinin kullanılması

- p-değerini hesaplayalım ve α ile kıyaslayalım (aşağıdaki p-değeri gelecek sayfada gösterilen excel sayfası ile hesaplanmıştır)



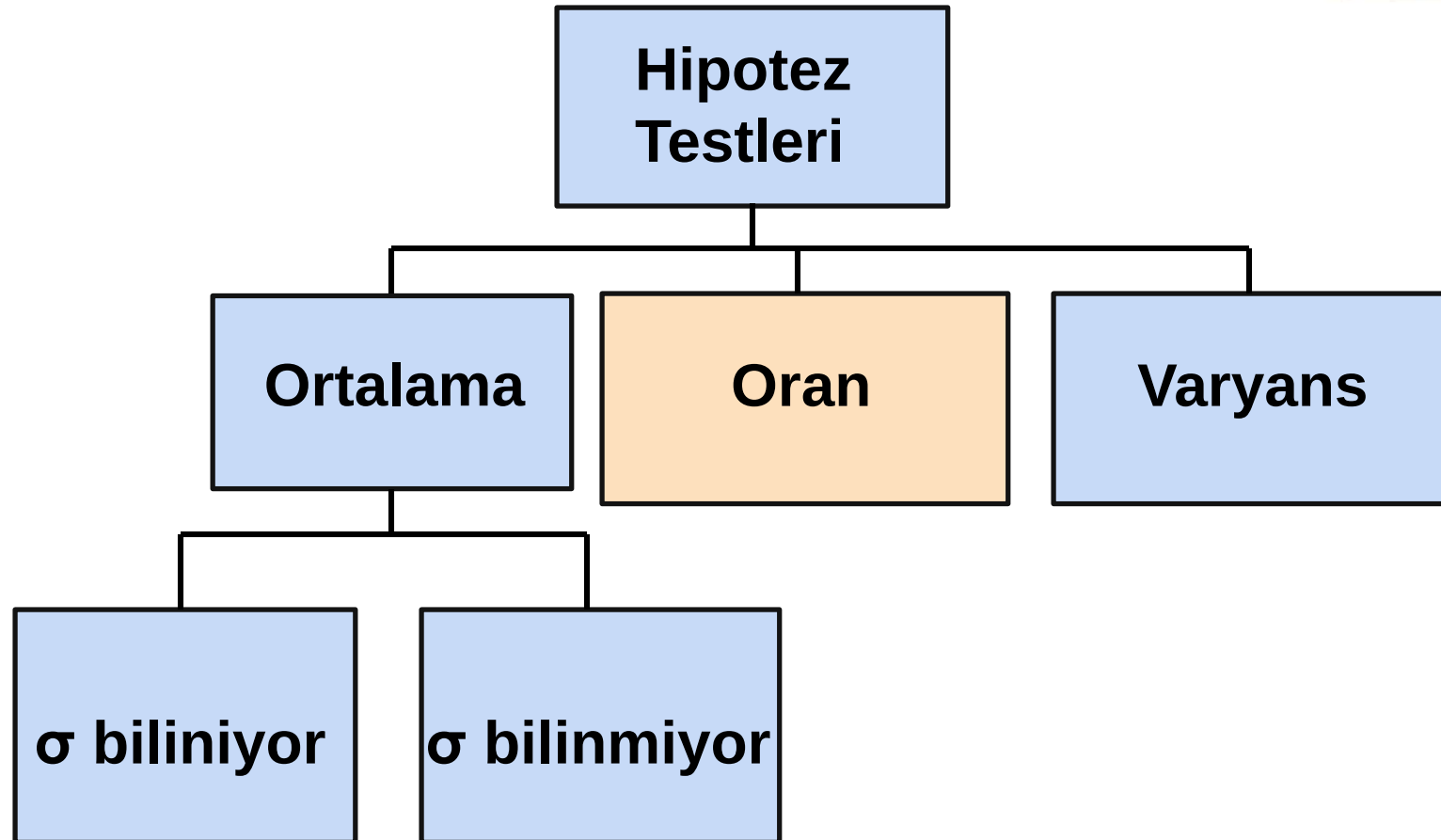
p-değeri = .2937 > $\alpha = .10$ olduğundan H₀'ı kabul et

Üst kuyruk t-testi için p-değerini hesaplayan excel sayfası

	A	B
1	t Test for the Hypothesis of the Mean	
2		
3	Data	
4	Null Hypothesis $\mu=$	184.2
5	Level of Significance	0.05
6	Sample Size	25
7	Sample Mean	170.8
8	Sample Standard Deviation	21.3
9		
10	Intermediate Calculations	
11	Standard Error of the Mean	4.2600 =B8/SQRT(B6)
12	Degrees of Freedom	24 =B6 - 1
13	t Test Statistic	-3.1455 =(B7 - B4)/B11
14		
15	Lower-Tail Test	
16	Lower Critical Value	-1.7109 =-T.INV.2T(2 * B5, B12)
17	p-Value	0.0022 =IF(B13 < 0, E11, E12)
18	Reject the null hypothesis	=IF(B17 < B5,"Reject the null hypothesis", "Do not reject the null hypothesis")

	D	E
10	One-Tail Calculations	
11	T.DIST.RT value	0.0022 =T.DIST.RT(ABS(B13), B12)
12	1-T.DIST.RT value	0.9978 =1 - E11

Hipotez testleri



Oranlar için hipotez testleri

- Kategorik değişkenleri içerir
- İki olası sonucu mevcuttur
 - İlgili duyulan özelliklere sahiptir
 - İlgili duyulan özelliklere sahip değildir
- İlgili duyulan kategorideki popülasyonun bir kesiri veya orantısı π ile gösterilir

Oran

- İlgili duyulan kategorideki örnek oran p ile gösterilir

- $$p = \frac{X}{n} = \frac{\text{örnek içerisinde ilgili duyulan kategorideki eleman sayısı}}{\text{örnek boyutu}}$$

- $n\pi$ ve $n(1-\pi)$ 'nin her ikisi de en az 5 olduğunda, p ortalaması ve standart sapması aşağıdaki değerler olan normal dağılıma yakınsar

-

$$\mu_p = \pi$$

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

Oranlar için hipotez testleri

- p 'nin örnekleme dağılımı yaklaşık olarak normal dağılıma uyar, buradan test istatistiği Z_{STAT} değeri:

$$Z_{STAT} = \frac{p - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}}$$

p için Hipotez Testi

$n\pi \geq 5$
ve
 $n(1-\pi) \geq 5$

$n\pi < 5$
veya
 $n(1-\pi) < 5$

Bu bölümde ele alınmayacaktır

İlgi duyulan kategorideki sayı cinsinden orantı için Z testi

- Son slayttakine eş biçimde, fakat ilgi duyulan kategorideki sayı cinsinden Oran X:

$$Z_{\text{STAT}} = \frac{X - n\pi}{\sqrt{n\pi(1-\pi)}}$$

X için Hipotez Testleri

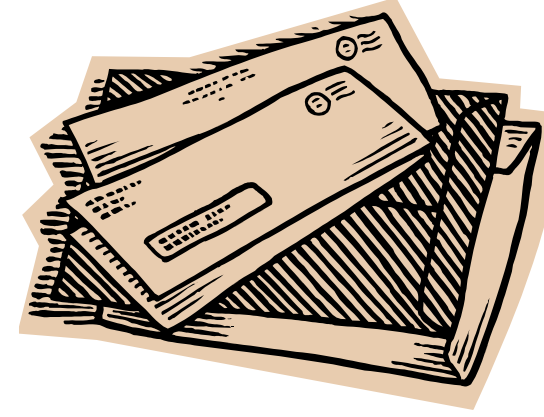
$X \geq 5$
ve
 $n-X \geq 5$

$X < 5$
veya
 $n-X < 5$

Bu bölümde ele alınmayacaktır

Örnek: Oran için Z testi

Bir pazarlama şirketi gönderdiği mektupların %8'ine dönüş yapıldığını iddia etmektedir. Bu iddianın test edilmesi için, 25 dönüşü sahip 500 kişilik bir örnek incelenmiştir. $\alpha = 0.05$ önem seviyesinde bunu test edelim.



Kontrol:

$$n\pi = (500)(.08) = 40$$

$$n(1-\pi) = (500)(.92) = 460$$



Oran için Z testi: Çözüm

$$H_0: \pi = 0.08$$

$$H_1: \pi \neq 0.08$$

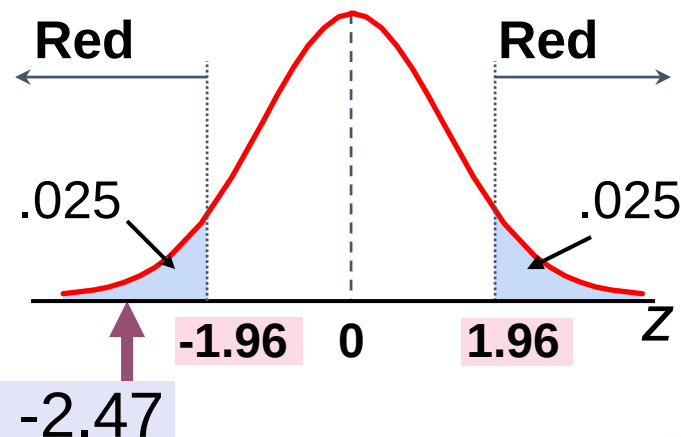
$$\alpha = 0.05$$

$$n = 500, p = 0.05$$

Test İstatistiği:

$$Z_{STAT} = \frac{p - \pi}{\sqrt{\frac{\pi(1 - \pi)}{n}}} = \frac{.05 - .08}{\sqrt{\frac{.08(1 - .08)}{500}}} = -2.47$$

Kritik Değerler: ± 1.96



Karar:

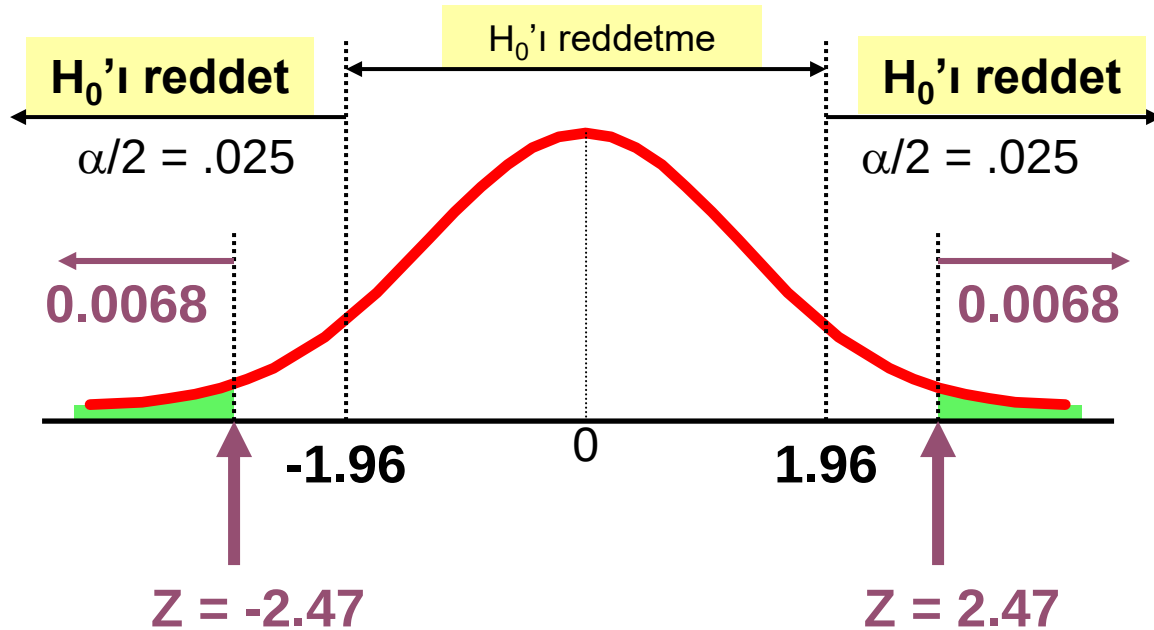
$\alpha = 0.05$ için H_0 'ı reddet

Çıkarım:

%8'lik dönüş oranının olduğu iddiasını reddetmek için yeterli delil mevcuttur.

p-değeri çözüm

p-değerini hesaplayalım ve α ile karşılaştıralım
(iki kuyruklu bir test için p-değeri her zaman iki-kuyrukludur)



p-değeri = 0.0136:

$$P(Z \leq -2.47) + P(Z \geq 2.47) \\ = 2(0.0068) = 0.0136$$

**p-değeri = 0.0136 < α = 0.05 olduğundan
 H_0 'ı reddet**

Örnek

- Bir tedavi yöntemi 40 hasta üzerinde uygulanmış ve bunlardan 30'u iyileşmiştir. Bu uygulama sonucuna göre %5 hata seviyesinde tedavi yönteminin etkinliğinin 0.80 den önemli derecede düşük olduğu söylenebilir mi?

Çözüm

i. $H_0: p=0.80$

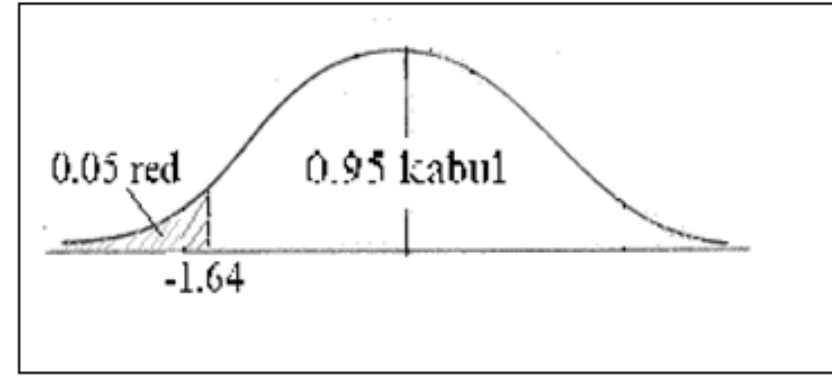
$H_1: p<0.80$

ii. $\alpha=0.05 \quad TTY \rightarrow Z_\alpha = Z_{0.05} = -1.64$

iii. $Z_h = \frac{\hat{p} - p}{\sigma_p} = \frac{\frac{30}{40} - 0.80}{\sqrt{\frac{0.80 * 0.20}{40}}} = -0.794$

iv. $|Z_h| = |-0.794| < |Z_{0.05}| = |-1.64| \rightarrow H_0$ reddedilemez.

v. $\alpha=0.05$ hata {veya $(1-\alpha)=0.95$ güven} seviyesinde tedavi yönteminin etkinliğinin 0.80'den düşük olmadığı söylenebilir Başka bir ifadeyle, tedavi uygulaması sonucunda ulaşılan düzey (0.75) ile ifade edilen düzey (0.80) arasındaki farklılığın önemli olmadığı, tesadüfi (*şans eseri*) olduğu kabul edilebilir.



Örnek

- Bir süpermarketler zinciri sahibi müşterilerinin %95'ten fazlasının süpermarketlerdeki fiyatlardan memnun olduğunu söylemektedir. Tesadüfi olarak seçilen 200 müşteriden 184'ü fiyatlardan memnun olduğunu bildirmektedir.%1 önem düzeyinde, süpermarketteki fiyatlardan memnun olanların oranının %95'e eşit olmadığını söyleyebilir miyiz?

Çözüm

$$H_0 : P = P_0$$

$$\alpha = 0.01$$

$$p = 184 / 200 = 0.92$$

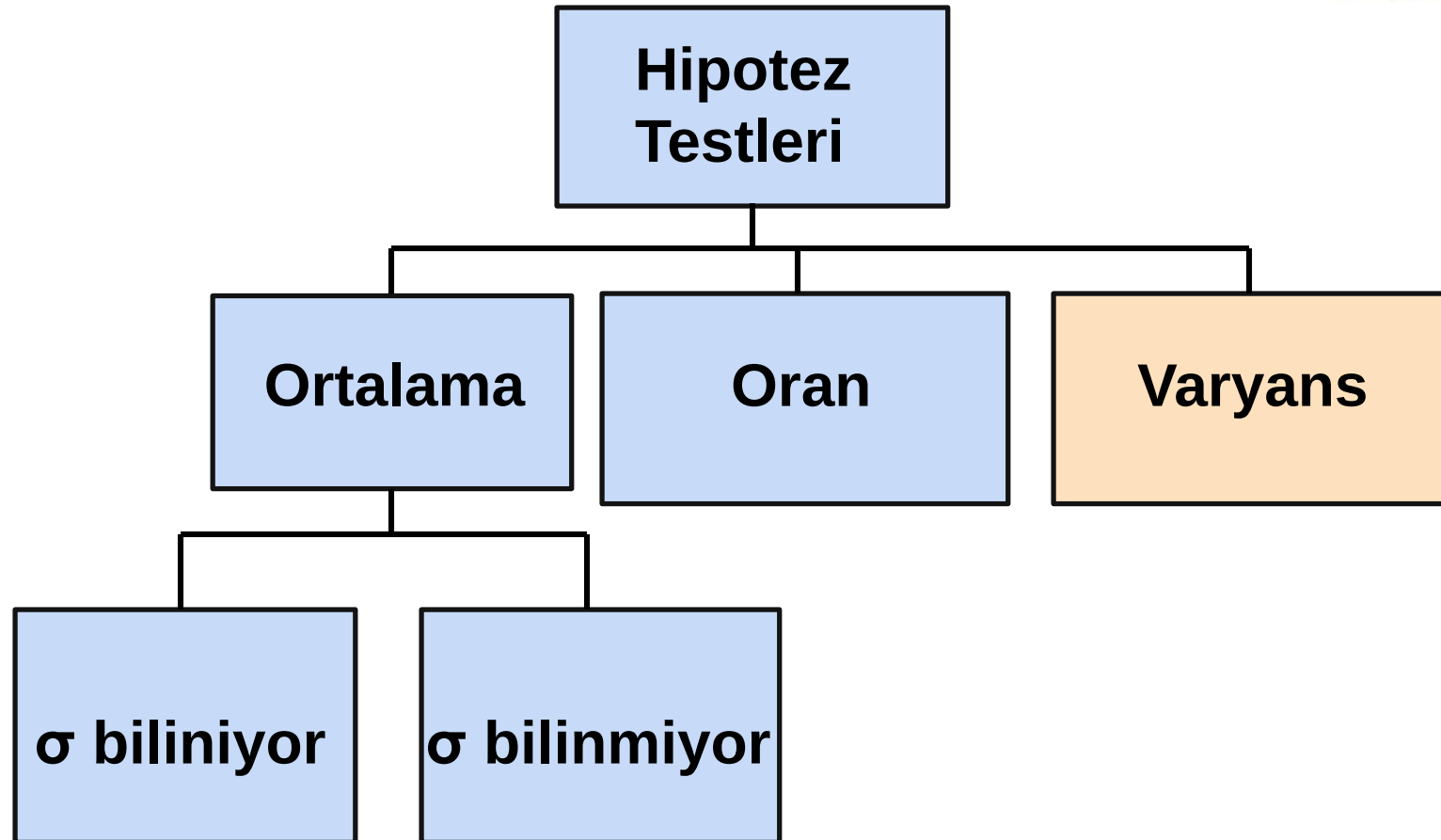
$$H_1 : P \neq P_0$$

$$Z_{tab} = \pm 2.58$$

$$Z = \frac{p - P}{\sigma_p} = \frac{p - P}{\sqrt{\frac{P(1-P)}{n}}} = \frac{0.92 - 0.95}{\sqrt{\frac{0.95(1-0.95)}{200}}} = -1.95$$

H_0 Kabul

Hipotez testleri



Normal dağılımın varyansı ve standart sapması ile ilgili testler

Bazen popülasyon varyansı veya standart sapması ile ilgili hipotez testlerine de ihtiyaç olmaktadır. Popülasyon normal dağılım ile modellendiğinde bu bölümde anlatılan testler uygulanır.

Varyans için Hipotez Testleri

Varsayalım ki normal dağılıma uyan popülasyonun varyansı σ^2 'nin önceden belirlenen σ_0^2 gibi bir değere eşit olup olmadığını veya benzer şekilde popülasyon standart sapması σ 'nın önceden belirlenmiş bir σ_0 değerine eşit olup olmadığını test etmek istiyoruz. $X_1, X_2, \dots, X_n$ bu popülasyondan alınmış n adet gözlemden oluşan rassal örneklem olsun.

$$H_0: \sigma^2 = \sigma_0^2$$

$$H_1: \sigma^2 \neq \sigma_0^2$$

Testini yapmak için aşağıdaki test istatistiğini kullanırız.

$$X_0^2 = \frac{(n-1)S^2}{\sigma_0^2}$$

Varyans için hipotez testleri

Eğer sıfır hipotezi $H_0: \sigma^2 = \sigma_0^2$ doğruysa yukarıdaki denklemde tanımlanan χ_0^2 test istatistiği $n - 1$ serbestlik derecesine sahip ki-kare dağılımına uyar. Test prosedürü için bu referans dağılımıdır.

Sabit önem seviyesi testi yapmak için ilgilenilen popülasyondan bir rassal örneklem alınır, χ_0^2 test istatistiğinin değeri χ_0^2 değeri hesaplanır ve eğer

$$\chi_0^2 > \chi_{\alpha/2, n-1}^2 \text{ ya da eğer } \chi_0^2 < \chi_{1-\alpha/2, n-1}^2$$

ise sıfır hipotezi $H_0: \sigma^2 = \sigma_0^2$ reddedilir. Burada $\chi_{\alpha/2, n-1}^2$ ve $\chi_{1-\alpha/2, n-1}^2$ değerleri $n - 1$ serbestlik derecesine sahip ki-kare dağılımında sırasıyla sağ ve sol kuyrukta kalan alanları $100 \alpha/2$ yapan değerlerdir. Şekil (a) da kritik bölge görülmektedir.

Varyans için hipotez testleri

Tek-uçlu alternatif hipotezler içinde aynı istatistik kullanılır. Tek-uçlu hipotezler için alternatif hipotez aşağıdaki gibi kurulduysa

$$H_0: \sigma^2 = \sigma_0^2$$

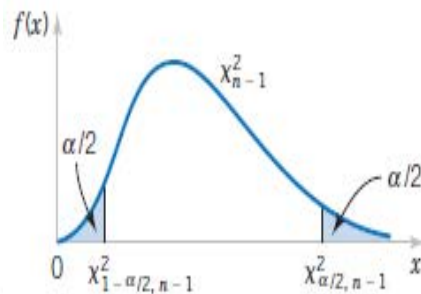
$$H_1: \sigma^2 > \sigma_0^2$$

$\chi_0^2 > \chi_{\alpha, n-1}^2$ iken sıfır hipotezi reddedilir ancak alternatif hipotez aşağıdaki gibi kurulduysa

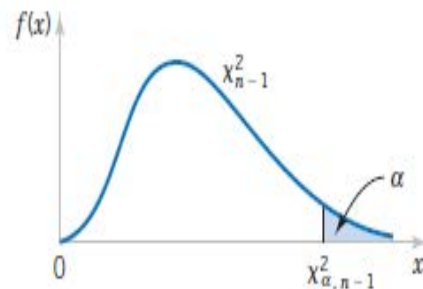
$$H_0: \sigma^2 = \sigma_0^2$$

$$H_1: \sigma^2 < \sigma_0^2$$

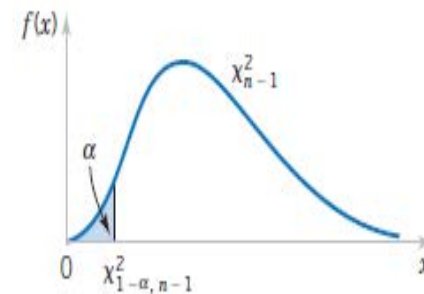
$\chi_0^2 > \chi_{1-\alpha, n-1}^2$ iken sıfır hipotezi reddedilir. Şekil (b) ve (c)'de tek-uçlu kritik bölgeler görülmektedir.



(a)



(b)



(c)

Varyans için hipotez testleri



Normal Dağılımın Varyansı için Testler

Sıfır Hipotezi: $H_0: \sigma^2 = \sigma_0^2$

Test İstatistiği: $\chi_0^2 = \frac{(n-1)S^2}{\sigma_0^2}$

Alternatif Hipotez Ret Kriteri .

$H_1: \sigma^2 \neq \sigma_0^2$ $\chi_0^2 > \chi_{\alpha/2, n-1}^2$ ya da $\chi_0^2 < \chi_{1-\alpha/2, n-1}^2$

$H_1: \sigma^2 > \sigma_0^2$ $\chi_0^2 > \chi_{\alpha, n-1}^2$

$H_1: \sigma^2 < \sigma_0^2$ $\chi_0^2 < \chi_{1-\alpha, n-1}^2$

Örnek: Otomatik dolum

- Şişelere sıvı deterjan dolumu otomatik bir dolum makinesi tarafından yapılmaktadır. 20 şişeden oluşan bir rassal örneklem için dolum hacminin varyansı $S^2 = 0,0153 \text{ lt}^2$ gelmiştir.
- Eğer dolum hacminin varyansı $0,01 \text{ lt}^2$ 'yi aşmışsa şişelerin kabul edilemez bir oranı az ya da çok doldurulmuştur.
- Örneklem verisinde üreticinin şişelerin az ya da çok doldurduğuna ilişkin bir bulgu var mıdır?
- Dolum hacimlerinin yaklaşık olarak normal dağıldığını kabul edeceğiz ve $\alpha = 0,05$ olarak alacağız.

Örnek

Burada ilgilenilen parametre popülasyon varyansı σ^2 'dir. Sıfır hipotezi $H_0: \sigma^2 = 0,01$ şeklinde ve alternatif hipotez $H_1: \sigma^2 > 0,01$ şeklinde olacaktır. Test istatistiği

$$\chi_0^2 = \frac{(n - 1)s^2}{\sigma_0^2} = \frac{19(0,0153)}{0,01} = 29,07$$

$\alpha = 0,05$ için $\chi_{0,05,19}^2 = 30,14$ bulunur. $\chi_0^2 = 29,07 < \chi_{0,05,19}^2 = 30,14$ olduğundan şişelerin dolum hacminin varyansının $0,01 \text{ lt}^2$ 'yi aştığına dair yeterli bulgu yoktur.

P-Değeri yaklaşımını da kullanabiliriz. Ki-kare tablosu incelendiğinde $\chi_{0,10,19}^2 = 27,20$ ve $\chi_{0,05,19}^2 = 29,07$ bulunur. $27,20 < 29,07 < 30,14$ olduğundan yukarıdaki problem için P-Değeri 0,05 ile 0,10 arasındadır.

Örnek

- Bir çimento fabrikasında üretilen çimentodan yapılan betonların dayanımına ilişkin standart sapmanın 1.24 kg/mm^2 olduğu iddia edilmektedir. Bu iddiayı test etmek üzere fabrikanın üretiminden rasgele alınan 25 birimlik örneğin ortalama dayanma miktarı ölçülmüştür. Daha sonra ölçülen verilerin ortalaması 25 kg/mm^2 ve varyansı da 2.4 olarak hesaplanmıştır. Bu verilere göre fabrika yöneticilerinin iddiasının doğru olduğu söylenebilir mi? ($\alpha=0.01$)

Çözüm

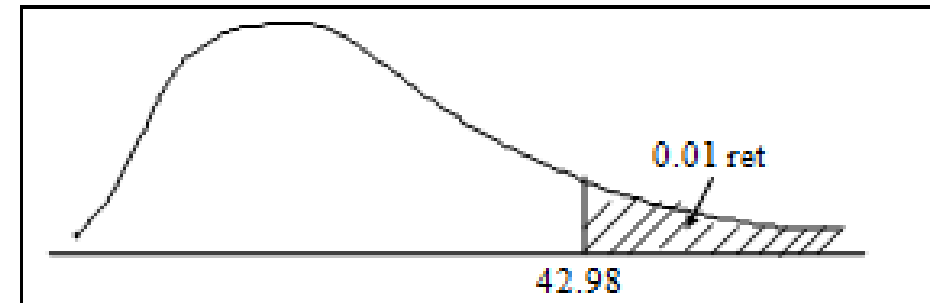
Varyansın büyük olması istenmeyen bir durumdur.

i. $H_0: \sigma^2 = (1.24)^2$

$H_1: \sigma^2 > (1.24)^2$

ii. $\alpha = 0.01$ $SD = n - 1 = 25 - 1 = 24$ ve $\chi^2_{0.01(24)} = 42.98$

iii. $\chi^2_h = \frac{(25-1)2.4}{(1.24)^2} = 37.44$



iv. $\chi^2_h = 37.44 < \chi^2_{0.01(24)} = 42.978 \rightarrow H_0$ reddedilemez.

v. $\alpha = 0.01$ hata {veya $(1-\alpha) = 0.99$ güven} seviyesinde fabrika yöneticilerinin iddiasının doğru (yani $\sigma = 1.24$) olduğu söylenebilir.

Uygunluk (uyum) testi

Birçok araştırmada gözlenen frekansların H_0 hipotezinde öne sürülen teorik frekanslara uyup uymadığı belirlenmek (test edilmek) istenir.

Örneğin:

- ✓ bir atölyede üretilen parçalarının boylarının parametreleri belirli olan bir normal dağılıma uyup uymadığı,
- ✓ bir kavşağa gelen araçların *gelişlerarası süresinin* parametresi belirli olan üstel dağılıma uyup uymadığı,...., vb konular test edilebilir.

Gözlenen frekansların, teorik olarak beklenen frekanslara uyup uymadığı ile ilgili olarak aşağıda verilen hipotez kurulabilir:

Hipotezler	Tablo değeri	Test istatistiği
H_0 : Gözlemler dağılımına uygundur H_1 : Gözlemler dağılımına uygun değildir	$\alpha = \alpha_0$ ve $\chi^2_{\alpha_0(k-1)} = a_0$ k : grup sayısı	$\chi^2_h = \frac{\sum_{i=1}^k (m_i - b_i)^2}{b_i}$

m_i : gözlenen değer

b_i : beklenen değer

Örnek

Bir zar 120 defa atılmış ve gelen değerler aşağıda verilmiştir.

X_i	1	2	3	4	5	6	Toplam	
m_i	17	18	24	26	21	14	120	→ Verilen
b_i	20	20	20	20	20	20	120	→ Hesaplanan

Bu veriler esas alınarak sözü edilen zarın hilesiz olduğu söylenebilir mi? ($\alpha=0.05$)

Çözüm

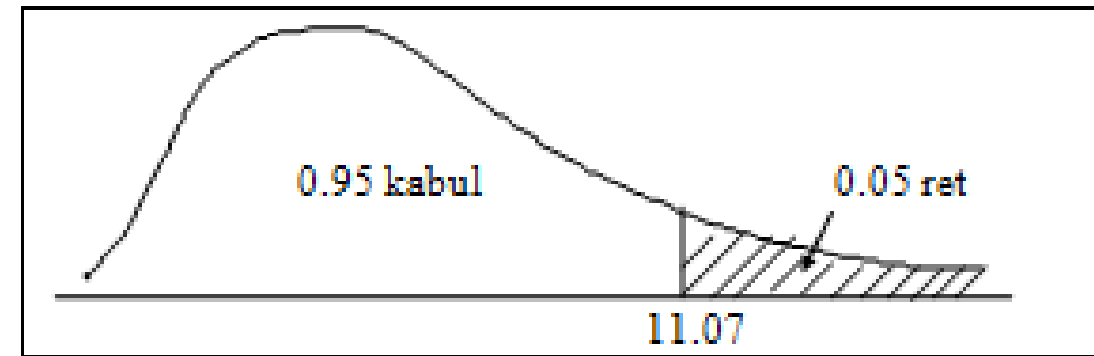
i. $H_0: p=1/6$ veya $b_i=20$

$H_1: p \neq 1/6$ veya $b_i \neq 20$

ii. $\alpha=0.05$ $SD=k-1=6-1=5 \rightarrow \chi^2_{0.05;(5)} = \mathbf{11.07}$

iii. $\chi^2_h = \frac{(17-20)^2}{20} + \frac{(18-20)^2}{20} + \dots + \frac{(14-20)^2}{20} = \mathbf{5.1}$

iv. $\chi^2_h = \mathbf{5.1} < \chi^2_{0.05;(5)} = \mathbf{11.07} \rightarrow H_0$ reddedilemez.



v. $\alpha=0.05$ hata {veya $(1-\alpha)=0.95$ güven} seviyesinde zarın hilesiz olduğu söylenebilir. Başka bir ifadeyle, 120 defa atılan zarın farklı değerlerde gelmiş olması onu hileli olarak nitelendirmek için yeterli kanıtı sağlamamaktadır.

Planlama aşamasında ele alınacak sorular

- Anketin, çalışmanın veya deneyin amacı nedir?
- Bu amacı bir sıfır hipotezine ve bir karşıt hipoteze nasıl çevirebilirsiniz?
- Hipotez testi tek veya iki kuyruklu mudur?
- Rassal bir örnek seçilebilir mi?
- Hangi tip veriler toplanabilir? Sayısal? Kategorik?
- Hangi önem derecesi kullanılmalıdır?
- İstenilen örnek büyüklüğü arzu edilen güce ulaşmak için yeterince büyük müdür?
- Hangi istatistiksel test süreci kullanılmalıdır?
- Planlanan hipotez testinin sonucundan hangi sonuçlar ve çıkarımlara ulaşılabilir?

Bu soruları dikkate almamak, önyargıya veya eksik sonuçlara neden olabilir!

İstatistiksel öneme karşı pratik önem

- İstatistiksel olarak önemli sonuçlar (sıfır hipotezini reddetme) her zaman pratik öneme sahip olmayabilir
 - Bu durum örnek boyutu çok büyük olduğunda daha çok olasıdır.
- Pratik öneme sahip sonuçlar istatistiksel olarak önemsiz olabilirler (sıfır hipotezinin reddedilememesi durumu)
 - Bu durum örnek boyutu görece olarak küçük olduğunda daha olasıdır

Bulguların raporlanması & etik konular

- İyi ve kötü sonuçların her ikisini de belgelemeli ve raporlamalıdır
- Sadece istatistiksel olarak önemli sonuçları raporlamamalıdır
- Raporlar kötü araştırma metodolojisi ile etik olmayan davranış arasında ayırım yapmalıdır
- Etik konular şu alanlarda ortaya çıkmaktadır:
 - o İnsan deneklerin kullanımı
 - o Veri toplama metodu
 - o Kullanılan testin tipi
 - o Kullanılacak önem seviyesi
 - o Verilerin temizlenmesi ve atılması
 - o İlgili bulguları bildirmedeki başarısızlık

Bölüm özeti

Bu bölümde şu konular işlenmiştir;

- Hipotez Testi Metodolojisi
- Ortalama için Z testi uygulanması (σ bilindiğinde)
- Hipotez Testi için kritik değer ve p–değeri yaklaşımları
- Tek kuyruklu ve iki kuyruklu testlerin uygulanması

Bölüm özeti

- Ortalama için t testi uygulanması (σ bilinmediğinde)
- Oran için Z testi uygulanması
- Varyans için Ki-kare testi
- Uygunluk (uyum) testi
- İstatistiksel ve pratik önem
- Güçlükler ve ahlaki konular



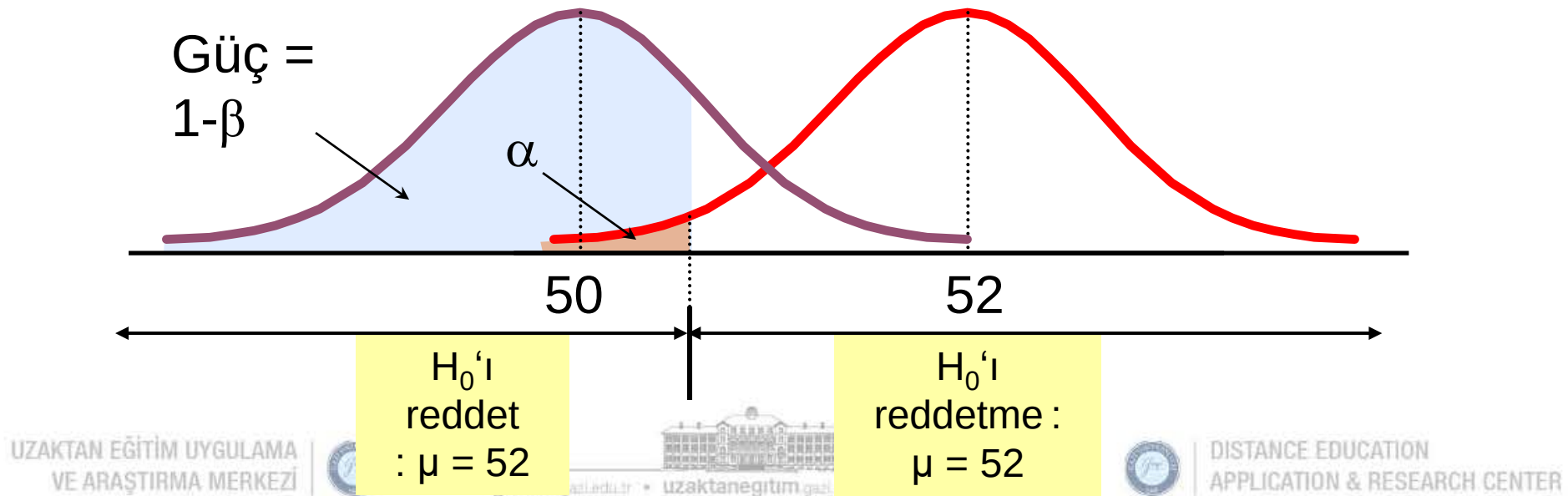
Bir Testin Gücü



Bir testin gücü

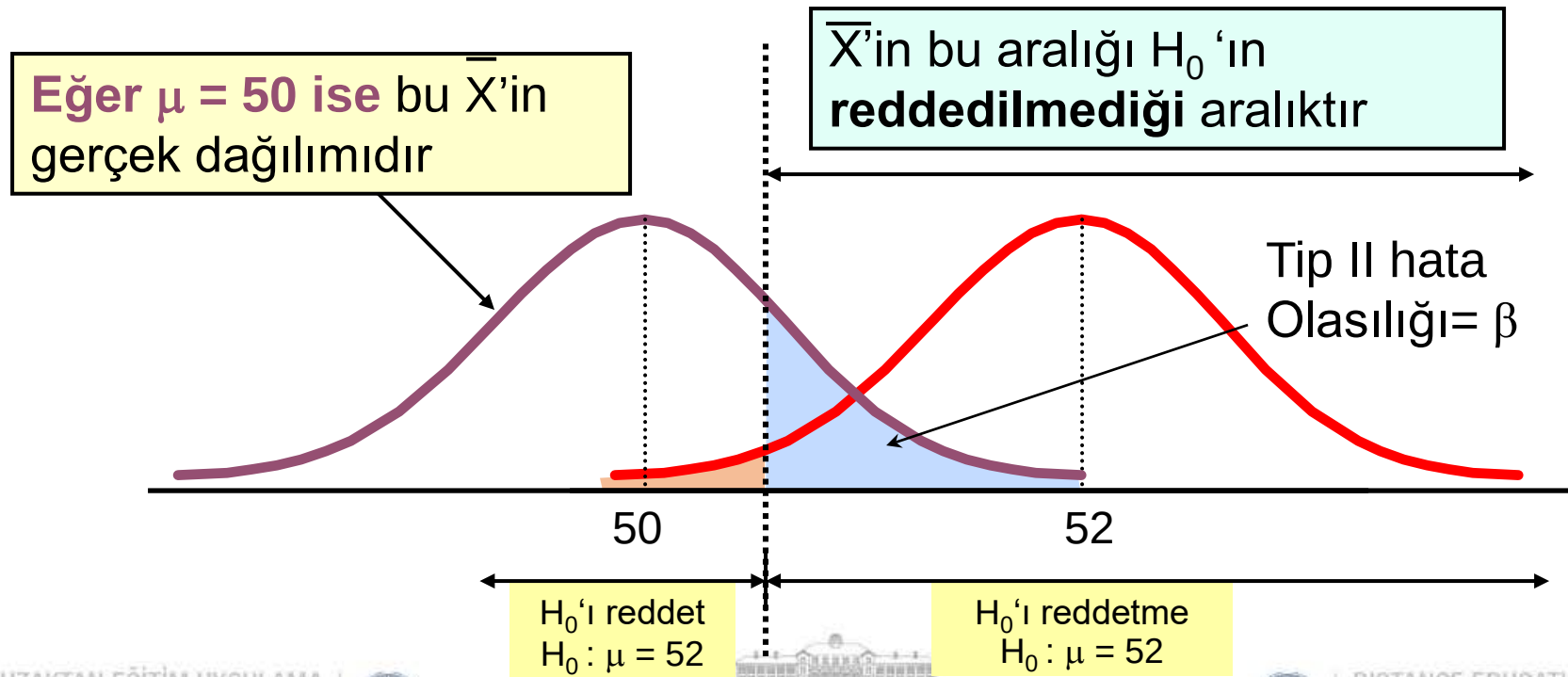
- Testin gücü yanlış bir H_0 hipotezini doğru bir şekilde reddetme olasılığıdır

Gerçek ortalama $\mu = 50$ iken $H_0: \mu = 52$ hipotezini doğru olarak reddettiğimizi varsayalım



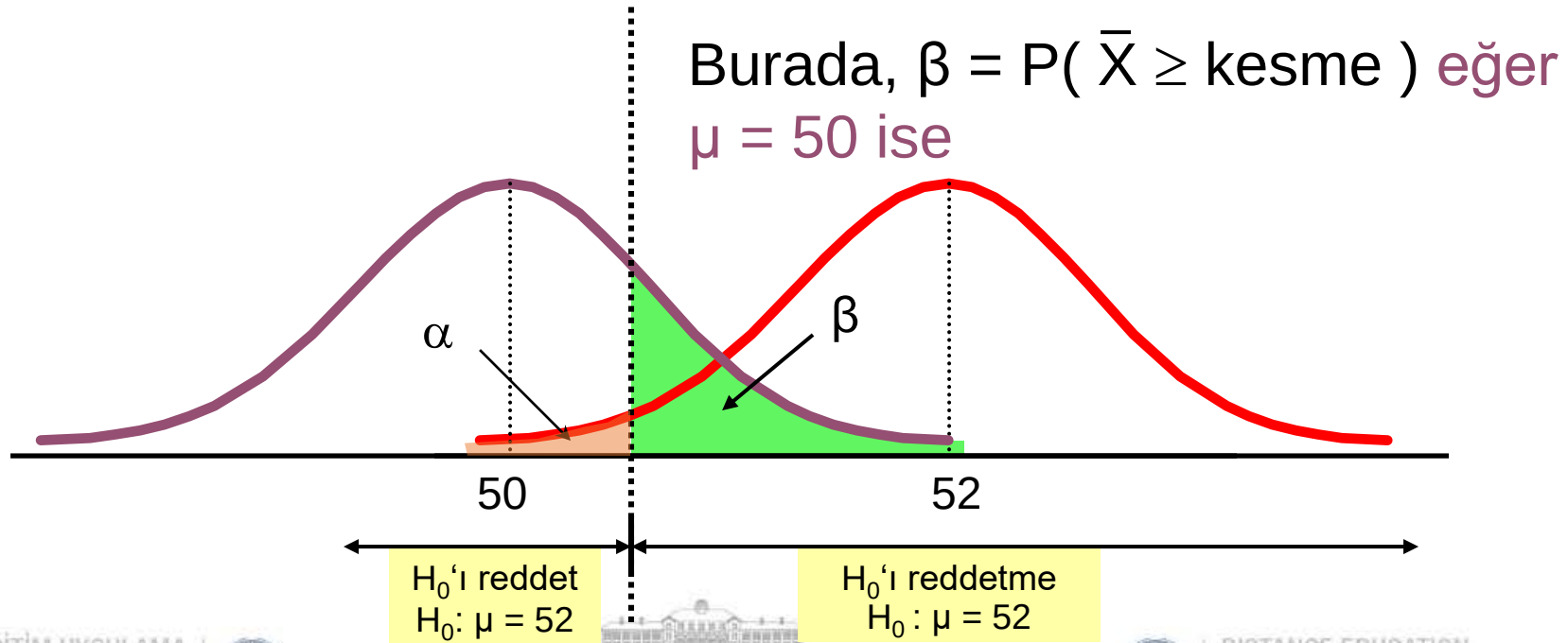
Tip II hata

- Gerçek ortalamamanın $\mu = 50$ olduğu halde $H_0: \mu = 52$ hipotezini reddetmediğimizi varsayalım



Tip II hata

- Gerçek ortalama $\mu = 50$ iken $H_0: \mu = 52$ hipotezinin reddedilmediğinde



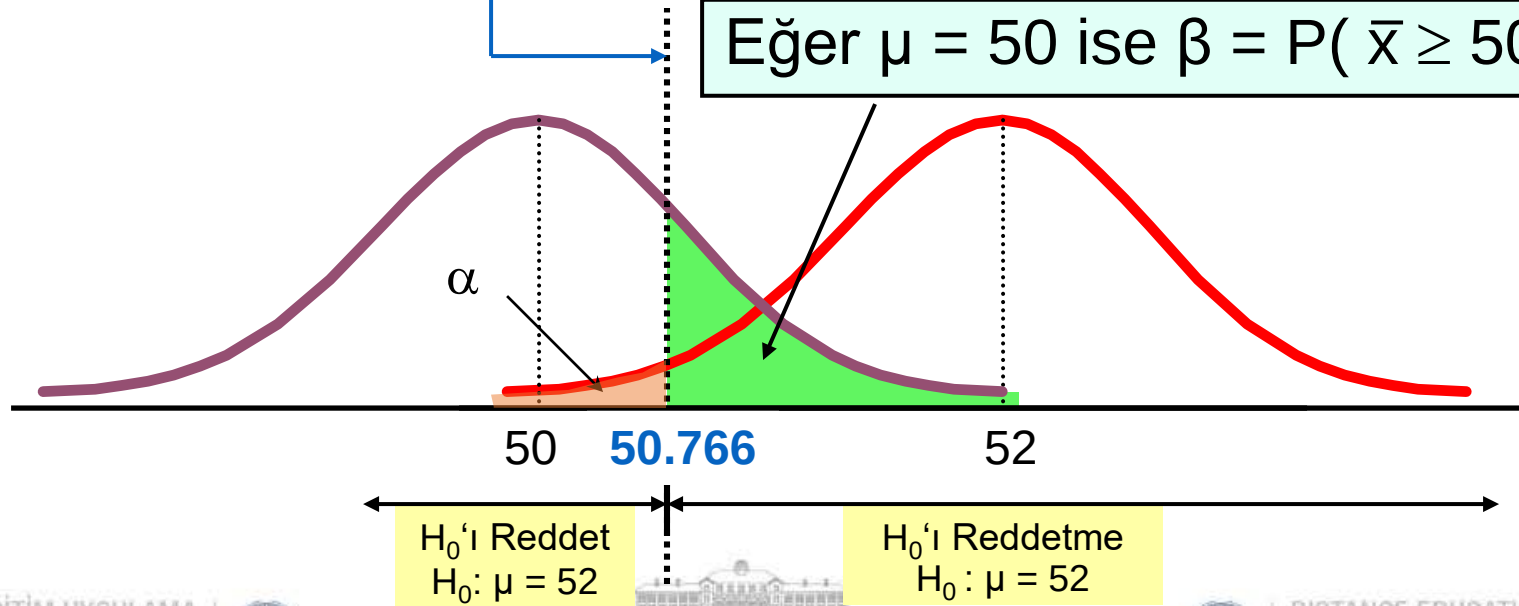
β 'nin hesaplanması

- $n = 64$, $\sigma = 6$, ve $\alpha = .05$ olduğunda

$$\text{kesme} = \bar{X}_\alpha = \mu - Z_\alpha \frac{\sigma}{\sqrt{n}} = 52 - 1.645 \frac{6}{\sqrt{64}} = 50.766$$

($H_0 : \mu = 52$ için)

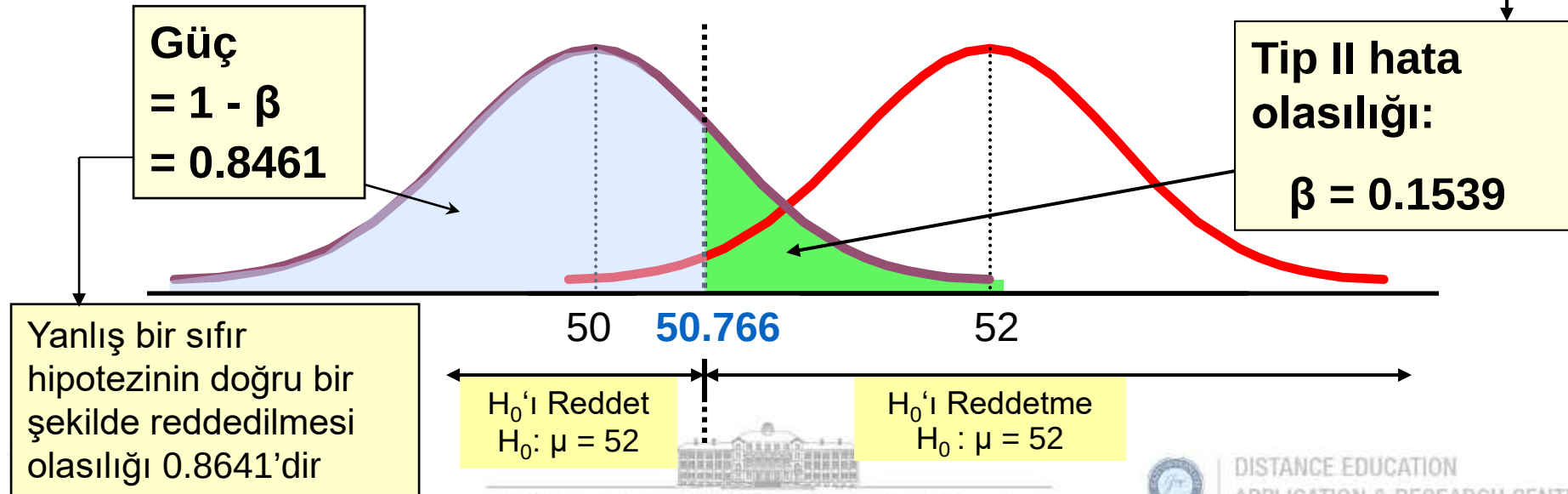
Eğer $\mu = 50$ ise $\beta = P(\bar{x} \geq 50.766)$



β 'nin ve testin gücünün hesaplanması

- $n = 64$, $\sigma = 6$, ve $\alpha = 0.05$ iken

$$P(\bar{X} \geq 50.766 | \mu = 50) = P\left(Z \geq \frac{50.766 - 50}{6/\sqrt{64}}\right) = P(Z \geq 1.02) = 1.0 - 0.8461 = 0.1539$$



Testin gücü

- Testin gücü ile ilgili elde edilen çıkarımlar:
 - Tek kuyruklu bir test iki kuyruklu bir testten daha güçlüdür
 - Önem seviyesindeki (α) artış, beraberinde güçteki artışı da getirecektir
 - Örnek boyutundaki artış, beraberinde güçteki artışı da getirecektir

Konu özetİ

Bu konu içerisinde

- Bir hipotez testinin gücünün nasıl hesaplanacağı ve bundan nasıl çıkarım yapılacağı işlenmiştir.