HİPOTEZ TESTLERİ

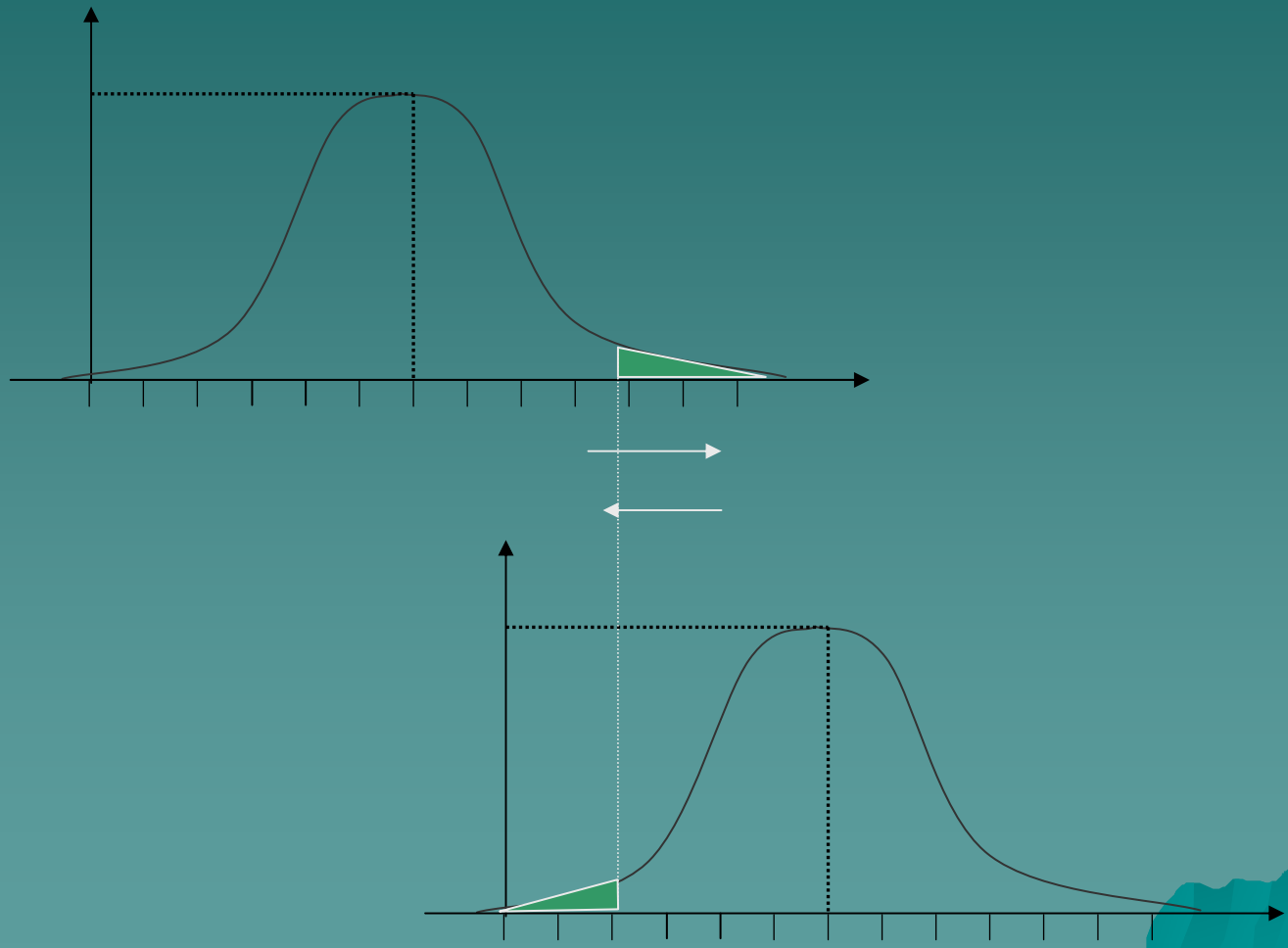
- Populasyonlar gösterdikleri olasılık dağılışları ile tanımlanırlar. Bu dağılışlar bilindiği taktirde populasyonlar hakkında verilecek kararlar kesinlik kazanır. Fakat populasyon dağılışları genellikle bilinmez. Dolayısıyla da bu tip kararların verilebilmesi zordur.
- Populasyonlar hakkında bilgi edinmenin diğer bir yolu örneklemedir. Uygun bir şekilde seçilen şans örnekleri yardımıyla populasyonun gösterdiği dağılışa ait parametreler tahmin edilir. Belirli varsayımlara dayanarak örneklerden elde edilen bu tahminler yardımıyla belli bir risk karşılığında populasyonlar veya bunların gösterdiği dağılışlar hakkında çeşitli kararlar verilir. Bu kararlar verilirken ya bir tahmin yapılır yada konu ile ilgili belirli bir varsayımda bulunulur. Gerçekleşsin veya gerçekleşmesin ileri sürülen bu tip varsayımlara *HİPOTEZ* denir.

- Sıfır hipotezi ve alternatif hipotez olmak üzere iki tip hipotez mevcuttur. Parametrelere belli değerler vererek kurulan hipotezlere H_0 (Null, Sıfır) hipotezi denir. H_0 hipotezi tek değerli bir hipotezdir. H_0'ın dışında kalan hipotezler ise Alternatif hipotez olarak adlandırılır ve H_1 ile gösterilir.

Bir hipotez testinde iki tip hata söz konusudur

	GERÇEKTE	
	Doğru	Yanlış
H_o RED	I. TİP HATA α	-
H_o Red Etmeme	-	II. TİP HATA β

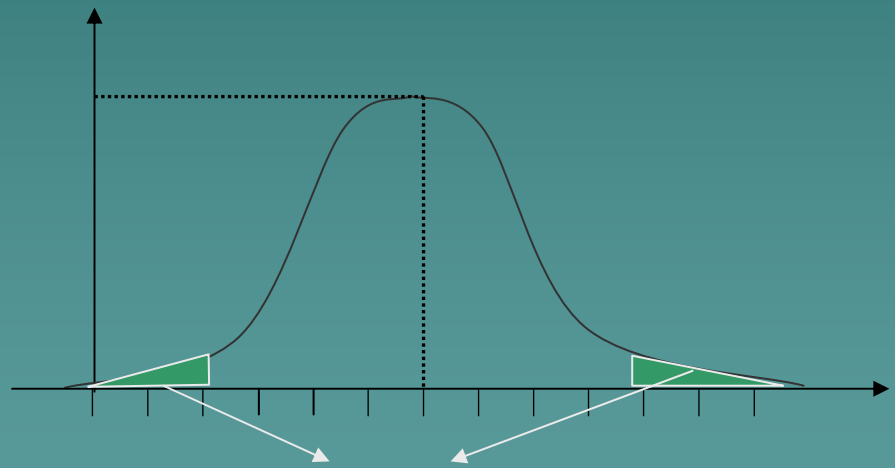
- Bu iki tip hata arasında ters bir ilişki mevcuttur



Örneğin, bir para denemesinde hipotezler üç farklı şekilde kurulabilir

I- H_0: P=0.5 para hilesizdir

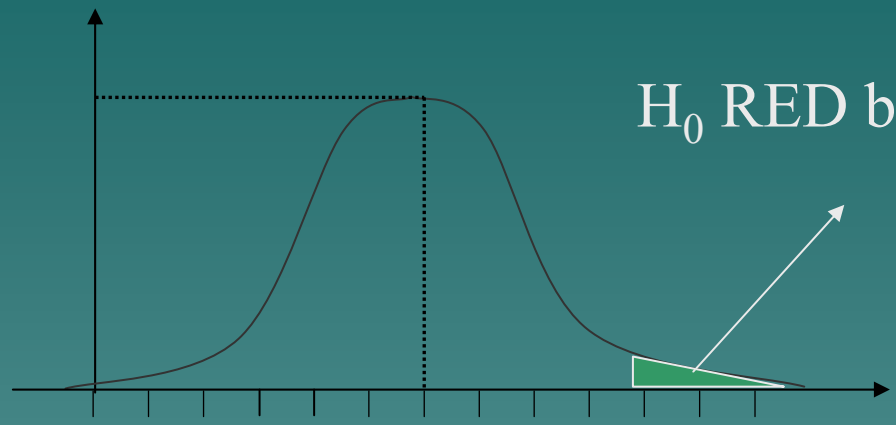
H_1: P ≠ 0.5 (Çift yönlü test)



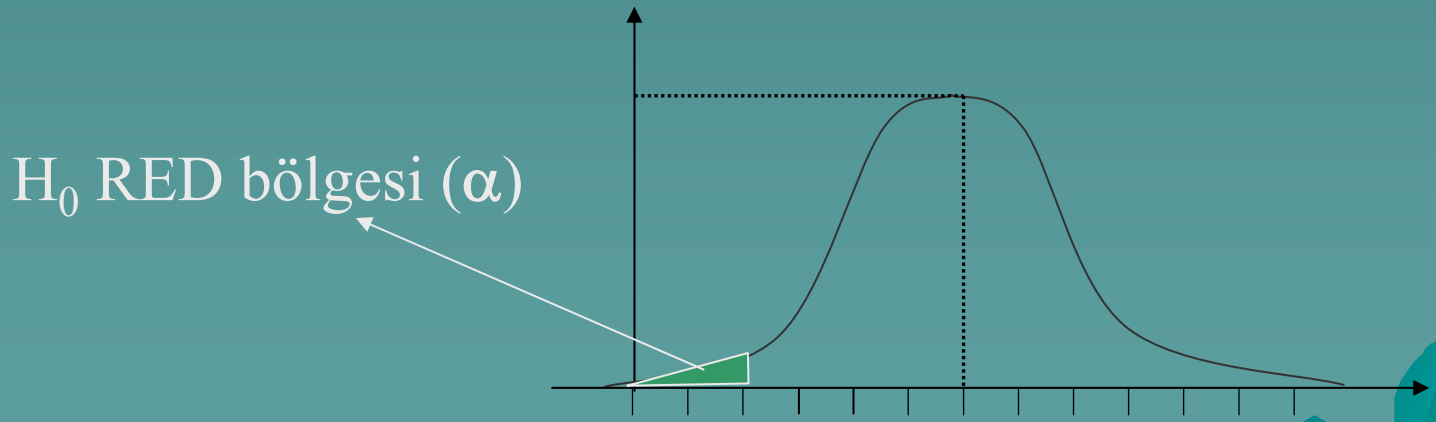
H_0 RED bölgeleri

$(\alpha/2)$

II- H_0:p=0.5 para hilesizdir
H_1:p> 0.5 para hilelidir (Tek yönlü)



III- H_0:p=0.5 para hilesizdir
H_1:p<0.5 para hilelidir (Tek yönlü)



- Biyolojik çalışmalarda birinci tip hata olarak adlandırılan α sabittir. Genellikle %5 ve %1 olarak kullanılır. Çift yönlü hipotezlerde α ikiye bölünür

HİPOTEZ TESTİ YAPILIRKEN İZLENECEK YOLLAR

1- Hipotezler kurulur

2- α önem seviyesi belirlenir

3- Uygulanacak test istatistiği seçilir

test istatistiği genel olarak

$$Testistatistiği = \frac{istatistik - parametre}{std.hata}$$

Test istatistikleri ve kullanıldıkları durumlar

1) Eğer σ^2 biliniyorsa Z istatistiği kullanılır

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}}$$

2) Eğer σ^2 bilinmiyor ve n≥30 ise bu örnek populasyonmuş gibi düşünülür ve yine Z istatistiği kullanılır. Fakat burada örneğin standart hatası kullanılır.

$$z = \frac{\bar{x} - \mu}{s_{\bar{x}}}$$

3) Eğer σ^2 bilinmiyor ve n<30 ise t istatistiği kullanılır

$$t = \frac{\bar{x} - \mu}{s_{\bar{x}}}$$

4) Bu eşitlikler yardımıyla Ho'ın doğru olduğu varsayılarak test istatistiklerinin değerleri hesaplanır.

5) Seçilen önem seviyesine göre hesaplanan istatistikler cetvel değerleri ile kıyaslanır. Eğer

hesap değeri > cetvel değeri

ise H_0 red edilir.

Normal Dağılışın Ortalaması için Yapılan hipotez testleri

- Örnek: Bir bölgede yetiştirilmekte olan buğday çeşidinin ortalama veriminin 275 kg/da ve varyansının 250 olduğu bilinmektedir. Bu buğday çeşidi 10 çiftçi tarafından ekilerek dekara verimler aşağıdaki gibi bulunmuştur. Buna göre daha önceki iddianın doğru olup olmadığını test ediniz?

Verim (kg/da)
325
240
260
280
300
285
310
345
295
310
$\bar{x} = 295$

Çözüm

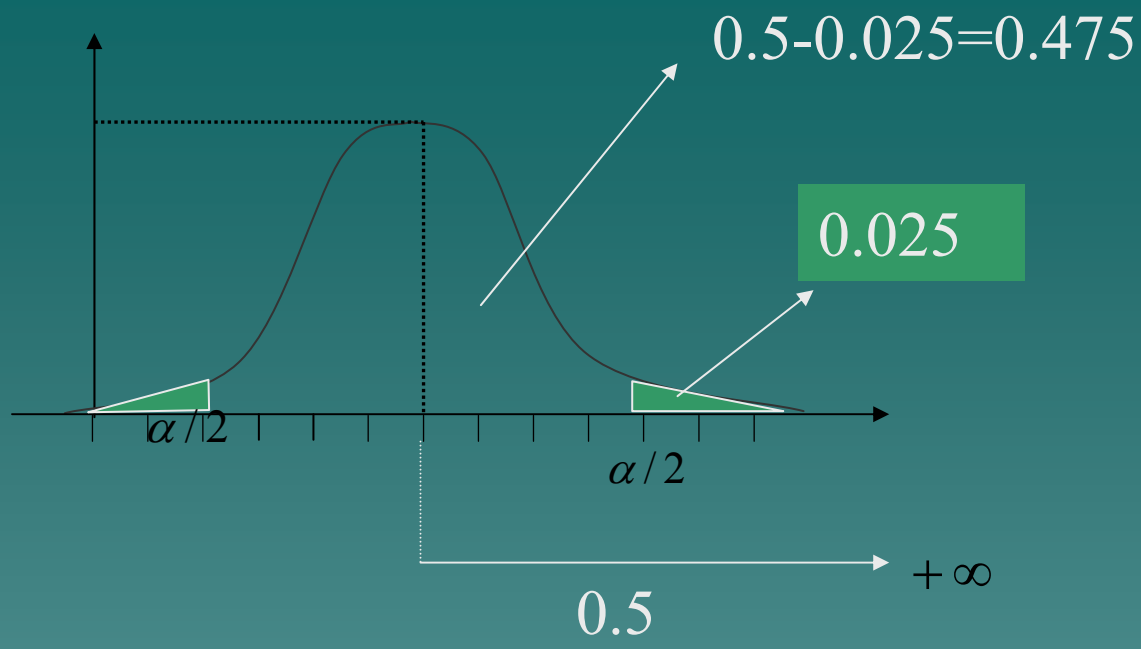
- Hipotezler

$H_0 : \mu = 275kg$ Çift yönlü hipotez

$H_1 : \mu \neq 275kg$

σ^2 bilindiğine göre Z testi uygulanır

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{295 - 275}{\sqrt{\frac{250}{10}}} = 4$$



Z cetvelinden 0.475 alana denk gelen nokta 1.96 değeridir. Buna göre Hesap değeri(4)>Cetvel değeri(1.96) olduğu için H_0 hipotezi RED edilir. Buna göre buğday çeşidinin verimi iddia edildiği gibi 275 kg/da değildir. Z hesap değerinin işaretinin + olması verimin 275 kg/da dan daha fazla olduğunu göstermektedir.

- Örnek 2: Yapılan bir çalışmada bir çilek çeşidinin 10 parseldeki verimleri kg olarak aşağıdaki gibi tespit edilmiştir. Buna göre bu çilek çeşidinin ortalama veriminin 10 kg/parsel olup olmadığını %5 önem seviyesinde test edip sonuçlarını yorumlayınız.

Verim (kg/da)
5
8
6
7
12
9
4
7
8
10
$\bar{x} = 7.6kg$

Çözüm

- Hipotezler

$$H_0 : \mu = 10kg$$

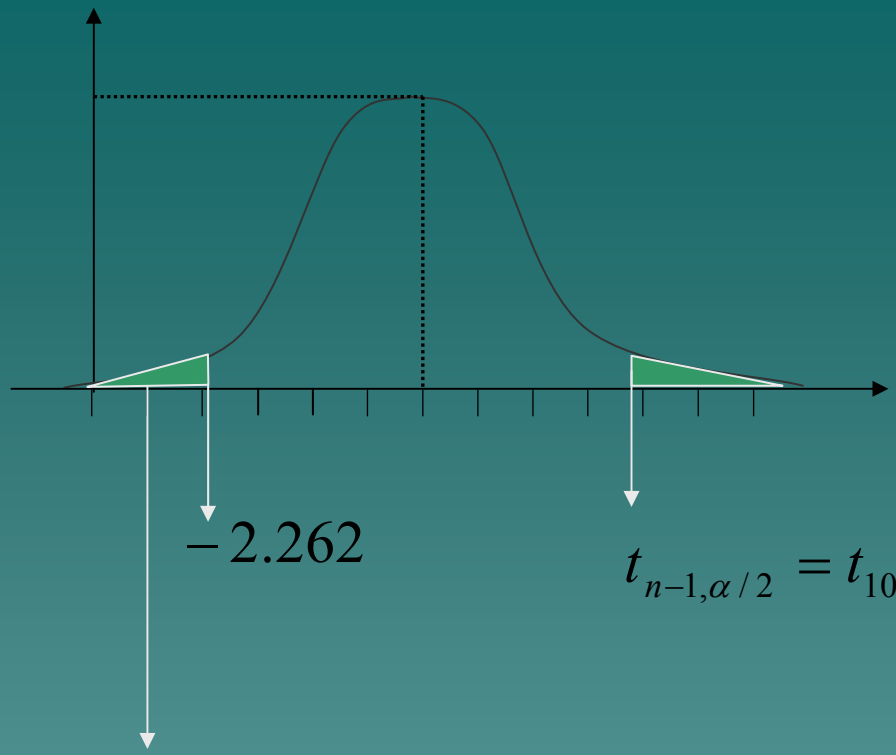
$$H_1 : \mu \neq 10kg$$

Çift yönlü hipotezler

σ^2 bilinmiyor ve n<30 olduğu için t istatistiği kullanılacaktır.

$$s^2 = \frac{\sum x^2 - \left(\sum x\right)^2 / n}{n-1} = \frac{628 - (76)^2 / 10}{9} = 5.6$$

$$t = \frac{\bar{x} - \mu}{s_{\bar{x}}} = \frac{7.6 - 10}{\sqrt{\frac{5.6}{10}}} = -3.209$$



HD=-3.209

Hesap değeri H_0 RED bölgesine düşmektedir. Bu çilek çeşidi parsel başına 10 kg verim vermemektedir.

- Örnek: bir mısır çeşidinin 8 parseldeki verimleri aşağıdaki gibi bulunmuştur. Bu mısır çeşidinin ortalama veriminin 20 kg/parsel olan bir mısır çeşidi populayonundan alınmış bir örnek olup olmadığını test ediniz.

Verim (kg/parsel)
15
18
16
17
22
19
14
17