

Hipotez Testleri ve Güven Aralıkları

Başlamadan önce;

α : Significance level (anlamlılık düzeyi)

* α , Hipotez testinin kalitesini gösterir

Anlamlılık düzeyi + Güven Düzeyi = 1

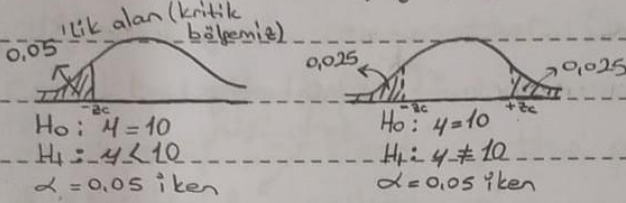
* $1 - \alpha$: Güven Düzeyi

Örnek vermek gerekirse; $\alpha = 0.01$ iken bir hipoteze doğru diyorsak; ora %99 güvenle doğru diyoruz.

(α arttıkça hipotezimizin kalitesi azalır.)

→ α arttıkça hata analizimiz büyümüş olur.

α , bize Kritik bölgeyi (Red bölgesini) verir.



→ α , kritik bölge ve kritik değeri bulmamıza yardımcı olur.

Kritik Bölge: Anlamlılık düzeyinden elde edilir.
(Red Bölge)

→ H_0 hipotezinin reddedildiği bölgedir.

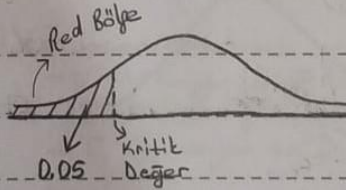
! Hipotezler popülasyon istatistikleri üzerinden kurulur.

(asla ve asla örneklem istatistikleri

ile kurulmazlar) → $H_0: \mu = 10$
 $H_1: \mu < 10$

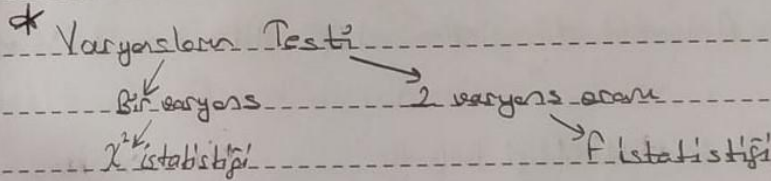
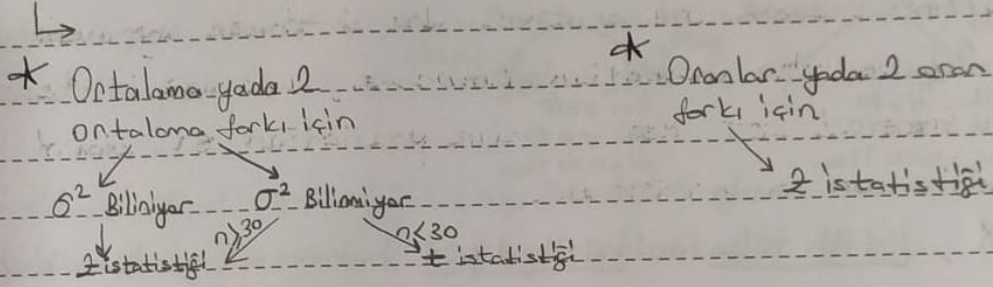
Örnek;

$H_0: \mu = 10$
 $H_1: \mu < 10$



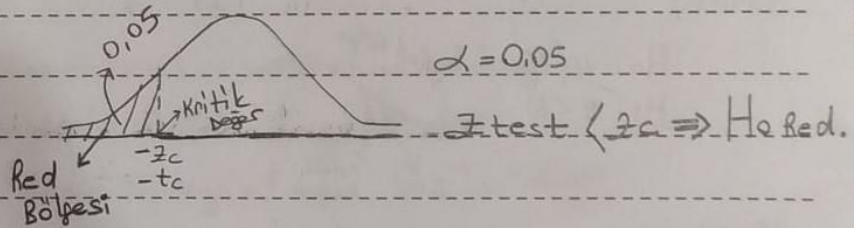
Kritik değer; 4 tane örneklem dağılımından biri ile hesaplanır.
Z tablosundan, t tablosundan, χ^2 ve F tablosundan
bulunur. (Hipotez testi hangi dağılım ile çözülecek ise, onu
tespit ettikten sonra o dağılıma ait tablodan
kritik değer bulunur)

Test İstatistiğinin Belirlenmesi



>> Test İstatistiğinin Kritik Değerin neyesinde kaldığına bakılarak "Hipotez Red veya Kabul edilir"

Örnek



| Parametrik Test | T-testi | Z-testi |
|-----------------|---|---|
| Anlam | Varyans verilmediğinde 2 veri setinin karşılaştırılması nasıl farklı olduğunu belirler. | Varyans verildiğinde 2 veri setinin karşılaştırılması birbirlerinden farklı olup olmadığını belirler. |
| Dağılım | Student-t dağılımı (normal dağılıma çok yakın) | Normal Dağılım |

> Z-testinde; popülasyon varyansı bilinmelidir.

> t-testinde; popülasyon varyansı bilinmez.

Kısaca; Z ve t İstatistikleri; Test İstatistiğinin belirlenmesinde kullanılır (z_{test} , t_{test}). Buna bağlı Hipotez Red veya Kabul edilir.

→ In t-test, we measure if 2 samples are different from one another. One of these samples could be the population, however, we use "T-test" in place of "Z-test" if the population standard deviation is unknown.

| | | Hipotez Test Sonucu | | α (Önem Seviyesi) |
|------------------------------------|--------|--|---|---------------------------------|
| | | significant finding RED | non-significant finding RED DEĞİL | |
| Ho Hipotezi (popülasyon sonucu) | Doğru | I. tip Hata (α) (Önem seviyesi) Doğru Karar 1- β | Doğru karar 1- α (Güven Aralığı) | Ho doğru iken; Ho Red olasılığı |
| | Yanlış | II. tip Hata β (Testin Gücü) | II. tip Hata β (Testin Zayıflığı) | |

↳ (Power of Test)

X Hipotez Testinin Aşamaları

- 1) Hipotez Kurulur
- 2) α ve β belirlenir
- 3) Test istatistikliği belirlenir

3.1) Popülasyon ortalaması ve Popülasyon oranı için

- σ^2 biliniyorsa Z istatistikliği
- σ^2 bilinmiyorsa ve $n > 30$ ise Z istatistikliği
- σ^2 bilinmiyorsa ve $n < 30$ ise t istatistikliği

3.2) Popülasyon varyansı (σ^2) için χ^2 istatistikliği

3.3) 2 popülasyon varyansının karşılaştırılması için F istatistikliği

4) Kritik değer bulunur

5) Test istatistikliği hesaplanır

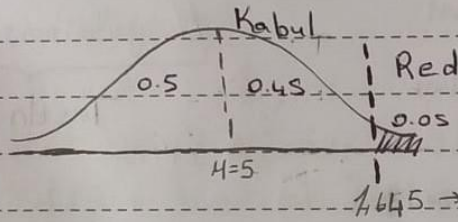
6) Kritik tablo değeri ile Test istatistikliği karşılaştırılır

Testin Gücü

ör Bir işletmede bilyelerin ortalama ağırlığı 5 gr.;
 standard sapması 0.1'dir.
 σ^2 Biliniyor (Büyük örneklerde Popülasyon Ortalaması)
 İşletmede bir değişiklik yapılıyor ve daha sonra üretilen
 16 rasal örnek alınıyor. Bu rasal örneğin ort = 5.038 gr
 $\rightarrow \alpha = 0.05$ iken popülasyon ortalamasının 5 gr'dan
 artıp veya artmadığını test edin.

* Sonu Z-testi sorusu. $H_0: \mu = 5$

$H_1: \mu > 5$



Güven Aralığı

Diyelim ki bir ciyate üreten firma ortalaması 2mm
 ciyatalar üretsın ve bızde bunu test etmek isteyelim
 $\alpha = 0.05$ iken; Eğer ki problemin sonunda
 hesapladığımız güven aralığı
 $P(1.91 < \mu < 2.12) = 0.95$ ise

Bunun yorumu şu şekildedir;

1) %95 emniyetle söyleyebiliriz ki tüm popülas-
 yonda ciyataların ortalaması 1.91 ile 2.12
 arasında olacaktır.

Veya

2) Aynı çalışmayı aynı popülasyondan seçilmiş
 100 farklı örneklem ile tekrarlarsam %95'inde
 $\mu \Rightarrow 1.91$ ile 2.12 arasında olur.

(ör) Bir araştırmaya 186 erkek cevap vermiş ve cevapların ortalaması 4,059, standard sapması 0,839 bulunmuş. 172 kadın katılmış ve cevapların ortalaması 3,68, standard sapması 0,966 bulunmuş. $\alpha = 0,01$ iken erkeklerin cevap ortalamasının kadınlarınkinden Yüksek olup, olmadığını test edin.

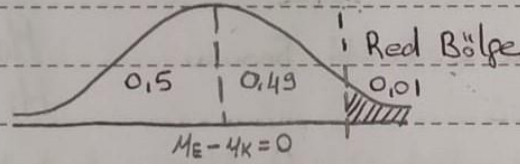
Varyanslar biliniyor ve büyük örnek (2 popülasyon arasındaki fark: $\mu_1 - \mu_2$)

* Sonuç Z testi Sonucu

$$H_0: \mu_E = \mu_K$$

$$H_1: \mu_E > \mu_K$$

$$\alpha = 0,01$$



(ör) Bir markette 802 müşterinin 378'i bir malı alıp veriş arabasına koyduktan sonra doğru fiyatı söyleyebilmiştir. Toplam müşterilerin en az yarısının doğru fiyatı söyleyebileceği sıfır hipotezini $0,1 = \alpha$ test edin

Popülasyon oranı için σ^2 biliniyor veya $n > 30$ (Büyük Örneklerde Popülasyon Oranı)

* Sonuç Z testi Sonucu

$$H_0: p \geq 0,5$$

$$H_1: p < 0,5$$

$$\alpha = 0,1$$

(ör) Russell olarak seçilen 203 İngiliz ticari dergi reklamından 52'sinde gülme unsuru var iken; 270 Amerikan ticari dergi reklamından 56'sında gülme unsuru işlenir. Tüm İngiliz ve ABD ticari dergi reklamlarında gülme unsuru işlenme oranlarının aynı olup olmadığına testi için kritik değeri bulun ve buna karşılık gelen H_0 Red olasılığının limit değerini hesaplayın.

(2 popülasyon oranı arasındaki fark) σ^2 known veya $n > 30$

* Soru 2 testi Sorusu

$$H_0: p_x - p_y = 0$$

$$H_1: p_x - p_y \neq 0$$

(ör) Bir mağazada Aralık ayı satışlarının Kasım ayı satışlarından 0.20 daha fazla olduğu bilinir. Popülasyon dağılımının normal olduğu varsayımı altında Aralık ayı satış ortalamalarındaki gerçek artışın 0.2 olduğu hipotezini $\alpha = 0.1$ iken test edin.

σ^2 bilinmiyor ve $n < 30$ (küçük örneklerde popülasyon Ortalaması)

$$H_0: \mu = \mu_0 = 20$$

$$H_1: \mu \neq 20$$

* Soru 1 testi Sorusu

(ÖR) 2 ayır Tv programında 10 ayır reklam kişilere
izletilmiş ve 24 saat sonra 2 ayır kışaktaki
reklamlardan hatırlanma indeksleri elde edilmiştir.
Buna göre; Sabah kışağın akşam kışağından daha
yüksek olup olmadığını $\alpha = 0.05$ için test ediyoruz.

σ^2 bilinmiyor ve $n \leq 30$ (2 popülasyon ortalaması arasındaki
fark: $\mu_1 - \mu_2$)

* Sonu t testi Sonusu

$$H_0: \mu_1 - \mu_2 = 0$$

$$H_1: \mu_1 - \mu_2 > 0$$

(ÖR) Bir çimento fabrikasında üretilen çimentodan yapılan
betonların sağlamlığının varyansının 10 kg/cm^2 den fazla olduğu iddia ediliyor.
Bu nedenle 10 beton örneği alınmış ve bu örneklerin sağlamlılıkları
saptanmıştır. Normal dağılıp gösteren bir popülasyondan alınan
bu örneklerin ortalaması $(\bar{x}) = 312$ ve varyansı $= 195$ bulunmuştur.
İddiaya 0.05 için sırayın

(Bir popülasyon varyansı için hipotez Testi)

$$H_0: \sigma^2 = 10 \text{ kg/cm}^2$$

$$H_1: \sigma^2 > 10 \text{ kg/cm}^2$$

$$\alpha = 0.05$$

* Sonu χ^2 (Ki-kare) testi Sonusu

ÖR) Pazarı sunulan 2 aynı bağımsız hisse senedinin değişkenliklerinin birbirine eşit olup, olmadığı test edin

(2 popülasyon varyansının karşılaştırılması)

$$H_0: \sigma_1^2 = \sigma_2^2$$

$$H_1: \sigma_1^2 \neq \sigma_2^2$$

* Soru "F" testi Sorusu

< Kapsamlı Bir Örnek (p-value included) >

→ Diyelim ki, çalışmamıza başında I. tip hata olasılığını $\alpha = 0.05$ olarak öngördük. Bunun anlamı "H₀ gerçekte doğru iken onu yanlışlıkla red etme olasılığımız maksimum %5 olmalı."

→ → Hipotez testi sonucunda hesaplanan I. tip hata miktarına "p değeri" diyebiliriz.

Eğer;

1) " $p < \alpha$ " ise, H₀ Red edilir. Bunun anlamı, H₀'ı red etmekle gerçekleşen yanlış "öngörülenden küçüktür." ($\alpha = 0.05$ iken; Öngördüğümüz maksimum yanlış: 0.05 p değeri=0.01 olsun; Gerçekleşen yanlış: 0.01)

* Yani H₀'ı red ettiğimizdeki yanlış öngörülenden çok daha az oldu.

2) " $p > \alpha$ " ise H₀ Red edilemez. Bunun anlamı, Gerçekleşen yanlış öngörülenden küçük olmadığı için H₀ Red edilemez.

ör

Varsayalım ki, Uygun hipotez testini kullandık ve p değerini 0.26 elde ettik

H_0 : A ve B diyetleri arasında fark yoktur

H_1 : B diyeti 'A' dan başarılıdır

$\alpha = 0.05$ olsun.

$p > \alpha$ için H_0 kabul ettik.

ör

||| Bir internet sitesi olsun ve insanların ortalama 20 dakika sitemizde kaldığını varsayalım. Eğer web sitenin arka plan rengini sarı yaparsam bu ortalama süre artacak mı bunu test edelim.

$H_0: \mu = 20$

$H_1: \mu > 20$

$\alpha = 0.05$ olsun

ilk arka plan



$\mu = 20$

2. arka plan



Buradan bir örneklem alacağız

Örnekleme aldık (Sarı arka planlı siteye ziyaret edenler arasında 100 = n kişi seçtik)



$n = 100$, $\bar{x} = 25$ dk, $sd = 2$ olsun

p value: $p(\bar{x} \geq 25 \text{ dk} \mid H_0 \text{ doğru}) \Rightarrow H_0 \text{ doğru iken, } \bar{x}'\text{'in 25 veya daha büyük olmasının olasılığı.}$

Mantık şu; Bir örneklem aldık ve H_0 hipotezinin doğru olduğunu varsaydık; eğer $p = 0.03$ ise bu örnekleme elde etme olasılığımız $\alpha = 0.03$ deriz.

önceki sayfa için $\bar{X}=25, sd=2$

* p değeri \rightarrow Örneklem istatistiklerini H_0 doğru iken elde etmenin olasılığıdır.

Sonuç olarak;

$\alpha \Rightarrow H_0$ doğru iken onu yanlışlıkla red etme olasılığı

$p \Rightarrow H_0$ doğru iken örneklem istatistiklerini elde etme olasılığıdır.

$\alpha = 0.05$; $p = 0.01$ olsun.

Bu durumda H_0 Red edilir.

Yani; H_0 'ı Red etme olasılığım, H_0 doğru iken alınan örnekleme dair istatistikleri elde etme olasılığım dan daha fazla oluyor oldu.