

## AB Testi (AB Testing)

2.11.22

"Without a grounding in Statistics, a Data Scientist is a Data Lab Assistant."

### Sampling (Örnekleme)

Bir ana kütle içerisinde, bu ana kütlelerin özelliklerini iyi taşıdığı temsil ettiği varsayılan bir alt kümedir. Population → Örneklem (Sample)

"The Future of AI will Be About Less Data, Not More"

HB2

Örneğin bir anket yapacaksın 1000 kişi üzerinden, tek tek bu kullanıcıları bulmak yerine bu 1000 kişiyi temsil eden bir alt gruba işlem yapabilirsin. Bu alt gruba örneklem diyoruz.

\* Anıya daha az veriyle genellemeler yapılabilir. Bıçırık bir yarıma yaklaşık tabiri zaman, para, işgücü vb. açıdan fayda sağlar.

Not : Elindeki veri setinde aykırı değerler varsa dağılımı çarpırsa bu değerleri temsil etmek için mean() 'e değil median() 'a bakılır. Aksi takdirde bir miktar yanıltıcı olabilir.

### Confidence Intervals (Güven Aralıkları)

Ana kütle parametresinin tahmini değerini (istatistik) kapsayabilecek 70 sayıdan oluşan bir aralık bulunmasıdır.

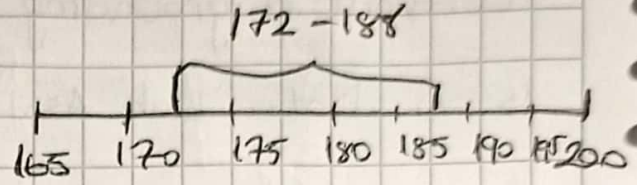


\*  $\text{mean} \pm \text{std}$  : ilkel olarak kullanılabildiği, bilimsel olarak güven aralığını tercih etmeliyim.

std : ortalamadan sapma demek Date : ...../...../.....

Örnek : Web sitesinde geçirilen ortalama süre için güven aralığı?

Ortalama : 180 sn  
Standart sapma : 40 sn



Kullanıcının 180 sn civarında hangi aralıkta, belirli bir kapsayıcılıkla, bir atalık verilseydi.  
\* Senin kullanıcıların web sitende geçirdiği ortalama süre %95 güven ile 172 sn ile 188 sn arasındadır.

Adım 1 :  $n$ , ortalama ve standart sapmayı bil.

$n = 100$  ortalama = 180 , standart sapma = 40

Adım 2 : Güven aralığına karar ver. 95 mi 99 mu?

Z tablo değerini hesapla.

%95 için Z değeri = 1,96

%99 için Z değeri = 2,57

Genelde bu kullanılır.  
Öntammi değerdir.

Kapsayıcılığı ifade eder.

Bir ortalama değerim var aralık istiyorum, bunu göstermek için bir genelleme yetkinliği varlığını istiyordum.

%95 bu aralıkta olacaktır.

Adım 3 : Yukarıdaki değerleri kullanarak güven aralığı hesapla. sapmayı standartlaştırır

$$\bar{X} \pm z \frac{s}{\sqrt{n}} = 180 \pm 1,96 \times \frac{40}{\sqrt{100}}$$

Sonuç :  $180 \pm 7,84$  yani 172 ile 188

arasındadır. %95 güven ile.

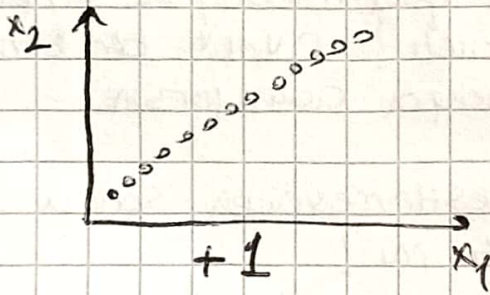


## Korelasyon (Correlation)

Değişkenler arasındaki ilişki, bu ilişkinin yönü ve güçlüğü ile ilgili bilgiler sağlayan istatistiksel bir yöntemdir.

Korelasyon  $-1$  ile  $+1$  arasında değerler alır.  $0$  korelasyon olmadığı anlamına gelir.  $+1$ 'e yaklaştıkça "mükemmel pozitif korelasyon",  $-1$ 'e doğru gidildikçe "mükemmel negatif korelasyon" söz konusu olur.

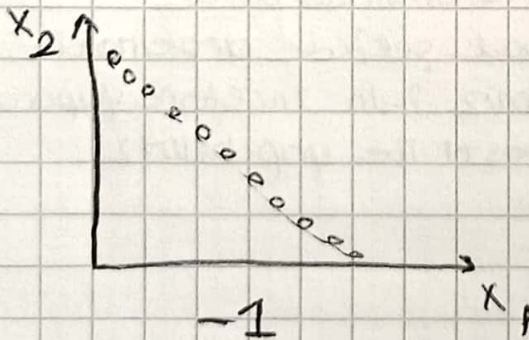
\* Mükemmel pozitif korelasyon bir değişken değeri arttıkça, diğerinde değeri artacağı anlamına gelir.



Mükemmel pozitif yönlü korelasyon.

Eğitim  $\rightarrow$  Gelir seviyesi  
arttıkça artar

\* Mükemmel negatif yönlü korelasyon bir değişken değerleri arttıkça, diğer değişken değerleri azalıyor olduğu anlamına gelir.



Mükemmel Negatif yönlü korelasyon.

Bilgisayar  
Sıklığı  $\rightarrow$  Performansı  
arttıkça azalır



## Hipotez Testleri (Hypothesis Testing)

3.11.2022

\* Bir hipotezi bir sayı test etmek için kullanılan istatistiksel yöntemlerdir.

\* Grup karşılaştırmalarında temel amaç, olası farklılıkların sans eseri ortaya çıkıp çıkmadığını göstermeye çalışmaktır.

Örneğin: Mobil uygulamalarda yapılan arayüz değişikliği sonrasında kullanıcıların uygulamada geçirdikleri ortalama süre arttı mı?

Tasarım 1: 1.55 dk

Tasarım 2: 2.58 dk vakit geçirsede, bu farklılık sans eseri mi ortaya çıktı? Çünkü bir örneklem üzerinden genelleme yapmaya çalışıyoruz.

Örneğin: Kayıt ekranı sadeleştirilmesinden sonra karpet kaydeden oranı arttı mı?

Tasarım 1 = 0.42

Tasarım 2 : 0.38

) bu farklılık sans

eseri ortaya çıkıp

olabilir. İstatistiksel olarak

şansa yer bırakmayacak şekilde aralarında fark vardır dememiz için inceleme yapacağız. Bunu da hipotez testi ile yapabiliriz.

kayıt ekranına  
acmak, ilk temas  
sonrası kaydeden



Subject:   
 kontrol grubu   
 deney grubu

Date : ...../...../.....

## AB Testi (Bağımsız İki Örneklem T Testi)

Ya İki grubun ortalaması kıyaslanıyordur. Ya da İki gruba ilişkin oranlar kıyaslanıyordur. 2 grup ortalaması arasında karşılaştırma yapılmak istenildiğinde kullanılır.

### İki Grup Ortalamasını Karşılaştırma

$$H_0 = \mu_1 = \mu_2$$

$$H_1 = \mu_1 \neq \mu_2$$

\*

$$H_0 = \mu_1 \leq \mu_2$$

$$H_1 = \mu_1 > \mu_2$$

$$H_0 = \mu_1 \geq \mu_2$$

$$H_1 = \mu_1 < \mu_2$$

$H_0$  = yokluk hipotezidir.

2 grup arasında fark yoktur. 2 grup ortalaması birbirine eşittir der.

$H_0$  hipotezini test ederiz.

Red ediliş edilmeme durumuna göre  $H_1$  hipotezi doğur.

$H_1$  = alternative hipotezdir.

\* hipotezi ile ilgilenmeyiz.



Subject :

Date : ...../...../.....

2 örneklem ortalamasını kıyaslarken kullanılan test istatistikleri vardır.

\* Örnek sayıları aynı, varyansları homojen ise,  
(Genelde bu kullanılır)

$$t = \frac{\bar{x}_1 - \bar{x}_2}{S_p \sqrt{\frac{2}{n}}} \quad S_p = \sqrt{\frac{S^2_{x_1} + S^2_{x_2}}{2}}$$

\* Örnek sayıları farklı, varyanslar homojen ise,

$$t = \frac{\bar{x}_1 - \bar{x}_2}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad S_p = \sqrt{\frac{(n_1 - 1)S^2_{x_1} + (n_2 - 1)S^2_{x_2}}{n_1 + n_2 - 2}}$$

\* Örnek sayıları farklı varyanslar homojen değil ise;

$$t = \frac{\bar{x}_1 - \bar{x}_2}{S_{\bar{A}}} \quad S_{\bar{A}} = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

Yukarıdaki formülasyonları kullanmak yerine

p value değerine bakılacaktır

$p < 0.05$  ise  $H_0$  redd edilir.



Bağımsız iki örneklem T testinin 2 tane varsayımı vardır

■ Normalite = 2 grubunda normal dağılması gerekir

■ Varyans Homojenliği = 2 grubunda dağılımlarının benzer olup olmaması demektir.

### Adımlar

1- Hipoteziyi kur

2- Varsayım kontrolü

- Normalite varsayımı
- Varyans homojenliği

3- Hipotezin uygulanması

a) Varsayımlar sağlanıyorsa bağımsız iki örneklem t testi (parametrik test)

b) Varsayımlar sağlanmıyorsa mann whitneyi testi (nonparametrik test)

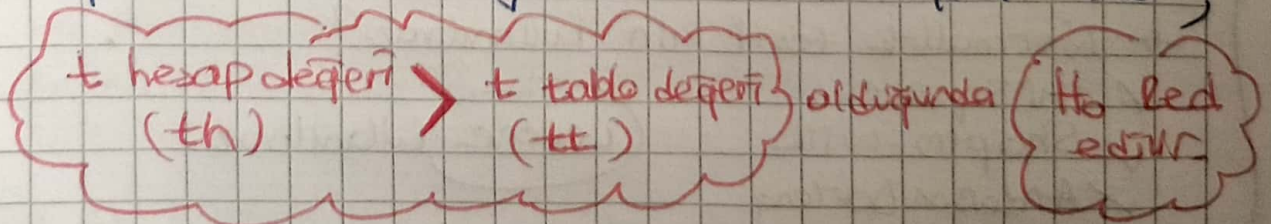
4- p-value değerine göre sonuçları yorumla

Not: Normalite sağlanmıyorsa direkt (b)'ye geçilir

Varyans homojenliği sağlanmıyorsa (a) arg'manı girilir

Normalite İncelemesi önceki aykırı değer incelemesi ve düzeltilmesi yapmak faydalı olabilir

Eğer t testi değerine batma ihtiyacı olursa





Subject :

Date : ...../...../.....

Mesela,  $t_h < t_t$  çıkarsa  $H_0$  Red edilemez.

Yani  $H_0 = H_1$  dir. Anlamlı bir fark olamaz.

Bu fark şans eseri ortaya çıkmıştır şeklinde yorumlanır.

\* Grupların birbirine bağımlı olması, daha fazla grup olması ile ilgili başka diğer testlerde vardır. Bu kapsamda değildir.

\* AB Testi denildiğinde 2 grubun ortalaması da kıyaslanabilir ya da dâhilîm oranları kıyaslanır.

### İki Örneklem Oran Testi

İki oran arasında karşılaştırma yapmak için kullanılır. Önceden ellinde ortalamalar vardı şimdi ise oranlar var.

$$H_0 : P_1 = P_2$$

P = Oran

$$H_1 : P_1 \neq P_2$$

$$Z_h = \frac{P_1 - P_2}{\sqrt{P(1-P) \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

İki ilgili grup içinde örneklem sayısı 30'dan büyüktür.



## İkiden fazla Grup Ortalaması Karşılaştırma (ANOVA - Analysis of Variance)

\* İkiden fazla grup olduğunda bu grupların ortalamları arasında fark var mıdır? Yok mudur? bununla ilgilenir.

\* Varyans analizi yapmaktan ziyade, grup ortalamalarını kıyaslayacağız.

Ör.  $H_0 = \mu_1 = \mu_2 = \mu_3$

$H_1$  = Eşit değildir (en az birisi farklıdır)

p value değeri = 0.002 olsun. o zaman,

↓  $H_0$  = Reddedilir

Gruplar arası anlamlı bir farklılık vardır. Bu farklılık hangi gruptan kaynaklanır buna da bakılabilir.

\* Significance = p-value değerine eşittir.

### Sorular

→ ANOVA testlerinin temel amaçlarından biri nedir?

\* Belirsizlik durumuna karar vermek.

→ Normalite koşulunun sağlanması için yapılan veri dönüştürme işlemlerinden biri değildir?

Karekökünü almak

Karesini almak

Ortalamasını almak

Logaritmasını almak

cevap

→ Herhangi bir spesifik dağılım şartı aranmayan yöntemlere ↓ yöntemler adı verilir.  
nonparametrik

→ Bir popülasyondan seçilen k adet örneklemenin varyanslarının eşit olması durumuna verilen ad nedir?  
Varyans Homojenliği



→ İki grup ortalaması arasında anlamlı fark olup olmadığı test edilmet istenmektedir. Grupların varyans homojenliği varsayımını sağladığı ancak normalite varsayımını sağlamadığı bilindiğine göre, iki grubun ortalamaları arasında istatistiksel olarak anlamlı bir fark olup olmadığı hangi test ile ölçilebilir?

Çerçp (Bağımsız İki Örneklem T Testi) - Non Parametrik  
İki Grup Ortalaması Karşılaştırma

## Bağımsız İki Örneklem T testi Adımları Özet

1) Hipotezi wr

$H_0: \mu_1 = \mu_2$  "Anlamı bir fark yoktur"

$$H_1: \mu_1 \neq \mu_2 \quad \text{"Anlamlıdır fark vardır."}$$

## 2) Voraussetzung Kontrolle

→ Normalite Varzayımlı: Bir değışkenin dağılımı standart normal dağılıma benzer olup olmadığının hipotez testidir.

$H_0$  : Normal dağılım varsayımı sağlanmaktadır  
 $H_1$  : " " " sağlanmamaktadır

Kullanılabilircek Testler :

- Shapiro-wilk
- Anderson Darling
- Pearson Chi-square
- Cramer-Von Mises
- Shapiro-Franco

$H_0 < 0.05$  ise  $H_0$  reddedilir.

Normalite testiye göre veriler normal dağılmıyorsa verileri LOGARİTMA, KAREKÜK, KARE ALMA gibi



matematiksel fonksiyonlar uygulayarak normalite sağlanmaya çalışılır.

★ Eğer bu şekilde sonuç vermezse parametrik olmayan (non-parametrik) testler uygulanır.

→ Varyans Homojenliği Varsayımı

$H_0$  = Varyans homojendir.

$H_1$  = Varyans homojen değildir.

Kullanılabilecek test : LEVENE (2 gruplar)

p-value  $< 0.05$  ise  $H_0$  Reddedilir.

3) Hipotezin Uyulanması

a) Varsayımlar sağlanıyorsa bağımsız iki örneklem t-testi (parametrik test)

b) Varsayımlar sağlanmıyorsa Mann-Whitney U test (non-parametrik test) uygulanır.

p-value  $< 0.05$  ise  $H_0$  Reddedilir.

Sonuç olarak, p-value değeri 0.05'ten küçük gelirse  $H_0$  reddedilir ve "gruplar arasında istatistik olarak anlamlı bir fark vardır" denir.