

BATCH: 223

DT/NT : DATA SCIENCE
LESSON : STATISTICS -2
SUBJECT: SIGNIFICANCE TEST
HYPOTHESIS TESTS
TYPE ERRORS



TECHPRO
EDUCATION

 techproeducation.com

 +1 (585) 304 29 59



STATISTICS - 2



Data Science Program

Statistics Session -8

Session - 8 Content

Content

- Significance Test
- Hypothesis test
- Type I – II Error
- One – Two tail tests



RECAP

**Herkes önceki dersten hatırladığı
1 cümle yazabilir mi?**



Recap – Previous Lesson

Simple Random Sampling (SRS)

- Örneklem Dağılımı
- Bir istatistik örneklem dağılımı
- Popülasyon dağılımı farklı
- Örneklem büyüklüğüne bağlıdır
- Mevcut örneklem yöntemlerinden Simple Random Samplingi kullanılmıyacak

Sampling Distributions of \bar{X}

Örneklem Dağılımı

- 2 zar atışının ortalaması, $n=2$

$$\mu_1 = \sum_{i=1}^n x_i \cdot P(x_i) = 1 \cdot \frac{1}{36} + 2 \cdot \frac{2}{36} + \dots + 12 \cdot \frac{1}{36} = 7$$

$$\sigma_1^2 = \sum_{i=1}^n x_i^2 \cdot P(x_i) - (\mu_1)^2 = 1 \cdot \frac{1}{36} + 4 \cdot \frac{2}{36} + \dots + 144 \cdot \frac{1}{36} - 49 = 5.83$$

$$\mu_2 = \sum_{i=1}^n x_i^2 \cdot P(x_i) = 1 \cdot \frac{1}{36} + 4 \cdot \frac{2}{36} + \dots + 144 \cdot \frac{1}{36} = 55.83$$

$$\sigma_2^2 = \sum_{i=1}^n x_i^2 \cdot P(x_i) - (\mu_2)^2 = 55.83 - 31.11 = 24.72$$

$$\sigma_2 = \sqrt{24.72} = 4.97$$

Genel kural

- $E(\bar{X}) = \mu_1 = \mu$
- $V(\bar{X}) = \sigma_1^2 = \sigma^2/n$
- $\sigma_{\bar{X}} = \sigma / \sqrt{n}$

Central Limit Theorem

Merkezi Limit Teoremi

- Eğer sample size yeterince büyük olursa, örneklem ortalamasının dağılımı normal dağılıma yaklaşır.
- normal popülasyonun ortalaması çok yakın olur.

Sample Distribution

Örneklem Dağılımı

- Örneklem dağılımı, hipotez testi yapmak için gerekli bilgiyi sağlar.
- Standard Error Kavramı
- Random Statistics kavramı

Sampling Distributions

Örneklem Dağılımı

- Popülasyon dağılımı ile bundan oluşan Sample Dağılımı birbirinden farklıdır.

Confidence Interval (CI)

Göven Aralığı

- İstatistiksel tahminler
- Aralık tahmini
- Göven aralığı kavramı ile konuşmak
- Örneğin: 95% güven ile
- Aralıkta dedi
- ortalama 0.05 yani %5'te 95% CI esen olur

SIGNIFICANCE TESTS

Anlamlılık Testleri

Significance Test

Anlamlılık Testi

- Diyetisyeni önerdiği program gerçekten kilo kaybına neden oluyor mu?
- İki farklı antibiyotiğin tedavi etkinliklerinin aynı mı değil mi?

Anlamlılık testi; bir popülasyon hakkında istatistiksel çıkarım yapmak için kullanılan ikinci ana yöntemdir.

Bunu Hipotez testiyle anlayabiliriz.

Kilo verdim ama bu kilo kaybı istatistik olarak anlamlı mı ?

Farklı diyetler uyguladım, her birinden ayrı ayrı kilolar verdim.. Peki bunlar istatistiksel olarak anlamlı mı ?

3 kişi A diyetisyenine gitti, kilo verdi

3 kişi B diyetisyenine gitti, kilo verdi

Bu 2 diyetisyen ile kilo verenlerin birbirlerine göre farkları istatistiksel olarak anlamlı mı?

Hypothesis Test

Hipotez Testi Tanımlar

- Hipotez testi, bir kitle parametresinin değeri hakkındaki bir iddiayı test etmek için örnek istatistikleri kullanan bir işlemdir.
- Gözlem ya da deneme sonucu elde edilmiş sonuçların, **rastlantıya bağlı olup olmadığının incelenmesinde** kullanılan istatistiksel yöntemlere HİPOTEZ TESTLERİ denir
- Ortaya atılan iddiaların, örnekten elde edilen bilgilere bağlı olarak, belirli bir hata payı ile doğrulanmasına HİPOTEZ TESTİ denir

Hipotez testi; Örneklerden elde edilen bilgiler ile teoriyi kıyaslayıp **KARAR VERMEYİ** içerir

Bu Sorulara cevap bulmalıyız:

Örneklerden yapılan çıkarımın Kurulan hipotez ile uyumsuz olduğuna nasıl karar verebiliriz?

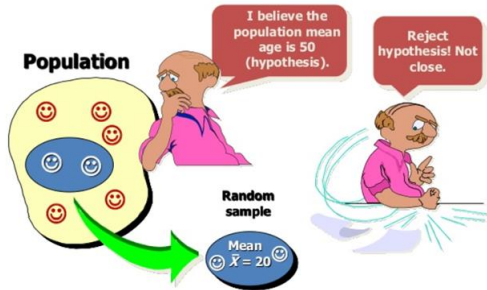
Kurulan hipotezi ne zaman ret/kabul edebiliriz?

Hatalı karar verme ihtimalimiz nedir?

Hypothesis Test

Hipotez Testi

- Bir kitle parametresiyle ilgili bir sözlü ifadeye veya iddiaya istatistiksel hipotez adı verilir.



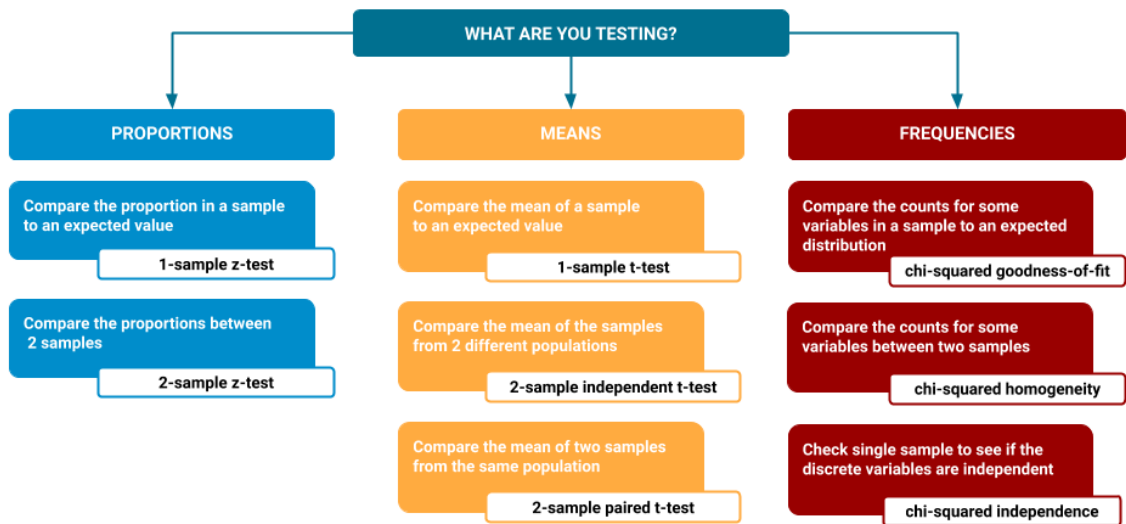
Şarj edilebilir pil üreticisi, ürettikleri pillerin ortalama en az 1.000 şarj için iyi olduğunu iddia ederse, bu iddiayı test etmek için bir örneklem alınacaktır

Ortalama 1000 saatlik süreyi test etmek için, bir iddiayı temsil eden, diğeri ise tamamlayıcısını temsil eden bir çift hipotez belirtilmiştir. Bu hipotezlerden biri yanlış olduğunda diğeri doğru olmalıdır



Simple Hypothesis Testing

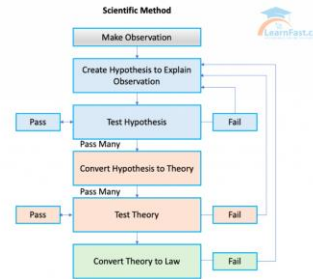
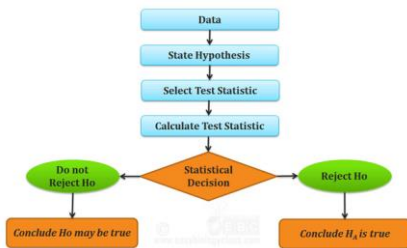
Choosing a simple test for comparing differences in populations



Significance Test Steps



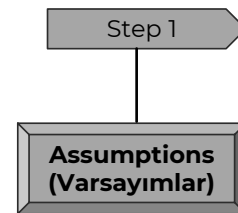
STEPS IN HYPOTHESIS TESTING



Hypothesis Test Steps – 1 : Assumptions

1. Varsayımlar

- değişkeni ve parametreyi belirtin
- Varsayımlar veri üretim yöntemi (randomization), sample size ve popülasyon distribution ile ilgilidir
- Burada z-test için gerekli varsayımlar geçerlidir



Örnekler random olmalı

Gözlemler Independent

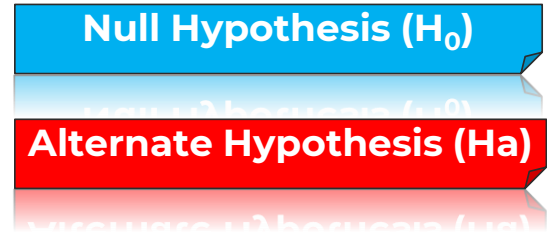
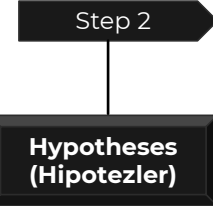
σ biliniyor veya $n > 30$



Hypothesis Test Steps- 2 : Hypotheses

2. Hipotezler

- Parametrelere belirli değerler verilerek kurulan hipotezlere **null hypothesis** (Sıfır hipotezi) denir
- Reject
- fail to reject
- Null hypothesis'in reddedilmesi (yanlış olduğunun ispatlanması) durumunda **alternate hypothesis** oluşturulur.



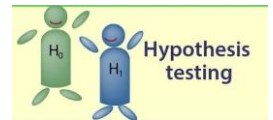
Hypothesis Test Steps – 2 : Hypotheses

H_0 Hipotezi

- Null (sıfır) hipotezidir
- başlangıçta doğru olduğu varsayılan iddiadır.
- Bu hipotez farksızlığı esas alır.
- İki ortalama arasında fark yoktur.
- İki grup arasında ilişki yoktur gibi

H_a Hipotezi

- Alternative hipotezidir
- H_0 'a aykırı olarak geliştirilen bir iddiadır
- H_1 alternatif hipotezi farklılık üzerine kurulur.



- Accept or Reject Null hipotez
- Hipotezin muhtemel sonuçları, H_0 ı reddetme (Reject) veya hipotezinin yanlış olduğunu ispatlayamamak
- Bazı kaynaklarda; retain – fail to reject gibi terimler de kullanılır



Örnek - 1

- A kolejinde mezunlarının % 94'ü mezun olduktan sonra altı ay içinde iş buluyor.

$$H_0: p = 0.94 \text{ (Claim)}$$

$$H_a: p \neq 0.94$$

$p = 0.94$ → Eşitlik durumu

→ Boş hipotezinin tamamlayıcısı

Örnek-2

- Bir üretici, şarj edilebilir pillerinin **ortalama en az 1.000** şarj ömrüne sahip olduğunu iddia eder.

$$H_0: \mu \geq 1000 \text{ (İddia)}$$

$$H_a: \mu < 1000$$

$\mu \geq 1000$ → Eşitlik durumu

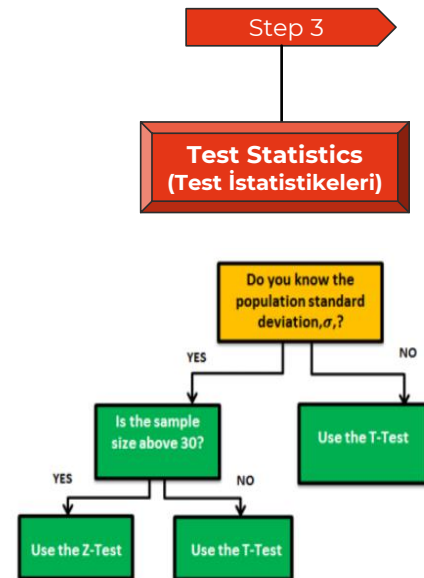
→ Boş hipotezinin tamamlayıcısı



Hypothesis Test Steps -3 : Test Statistic

3. Test İstatistikleri

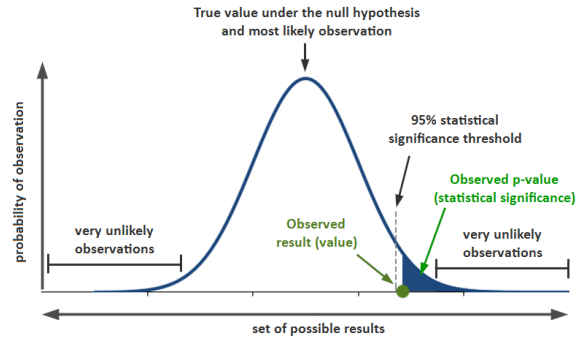
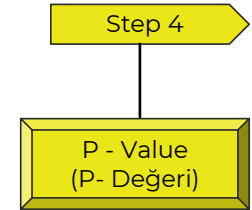
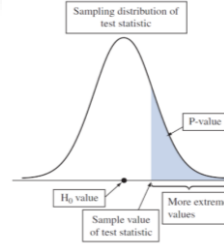
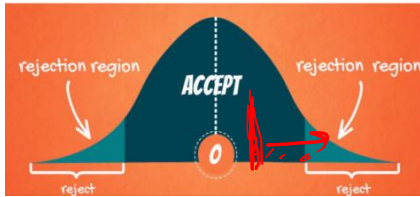
- Yandaki Algoritmaya göre hareket edebiliriz
- Pratik hayatta Z-test mi yapmalıyım yoksa t test mi diye düşünüyorsak, t testi ile işe başlamak tavsiye edilir.



Hypothesis Test Steps – 4 : P - Value

4. P Değeri

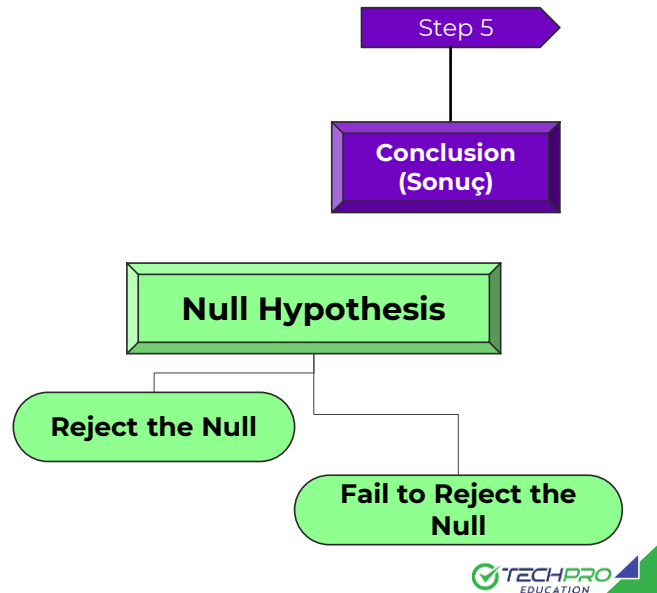
- gözlemlenen örneklem sonuçlarının ne kadar aşırı olduğunu ölçmek için
- Eğer $p\text{-value} < 0.05$ (alfa) ise ele aldığımız H_0 hipotezi reddedilmiş olunur



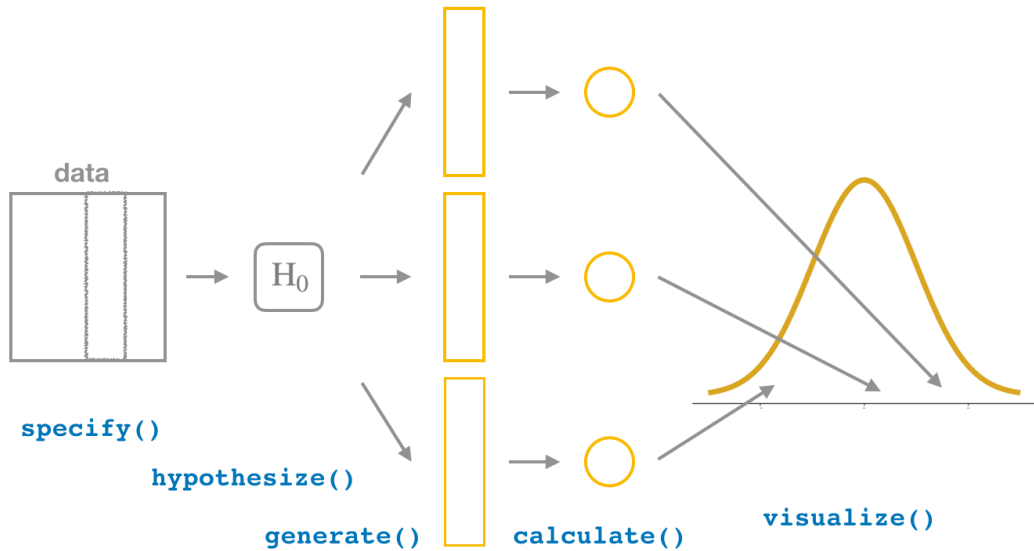
Hypothesis Test Steps- 5 : Conclusions

5. Sonuç

- Sonuç 2 ihtimal olabilir
 - Null hipotezi reddet
 - Null Hipotezi kabul et (veya fail to reject denir)



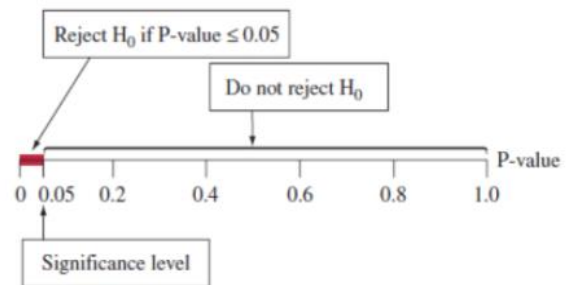
Steps



Significance Level (α - alpha)

► Anlamlılık seviyesi –Önem Seviyesi

- En çok kullanılan significance level $\alpha = 0,05$ (%95 CI ya göre)
- Ayrıca 0,01 de çok kullanılabilir (problemin hassasiyetine göre)



Common significance levels

0.10

0.05

0.01



Significance Level (α – alpha)

Anlamlılık seviyesi –Önem Seviyesi

- α , Normal dağılımın kuyruklarında toplam alanı ifade eder.
- $P < \alpha$ ise etki istatistiksel olarak anlamlıdır

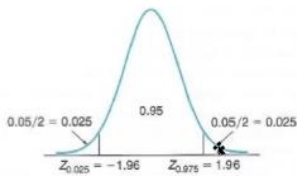


Table Values of $Z_{1-(\alpha/2)}$ for Confidence Intervals

Confidence Level (%)	$Z_{1-(\alpha/2)}$	α^*
99.99	3.819	0.0001
99.9	3.291	0.001
99	2.576	0.01
95	1.960	0.05
90	1.645	0.10
80	1.282	0.20

One-sided (one-tailed) test

Used when the null doesn't contain equality (=) or inequality sign ($<$, $>$, \leq , \geq)



Two-sided (two-tailed) test

Used when the null contains an equality (=) or an inequality sign (\neq)



Question41: What do you understand by Hypothesis Testing?

In Statistics, Hypothesis Testing is mainly used to see if a certain experiment generates meaningful results. It helps assess the statistical **significance** of insight by finding the odds of the results occurring by chance. In Hypothesis Testing, the first thing is to know the null hypothesis and then specify it. After that, the p-value is calculated, and if the null hypothesis is true, the other values are also determined. The alpha value specifies the significance, and you can adjust it accordingly.

If the p-value is less than the alpha value, the null hypothesis is rejected, but the null hypothesis is accepted if the p-value is greater than the alpha value. If the null hypothesis is rejected, it indicates that the results obtained are statistically significant.

Question52: What is the relationship between the **significance** level and the confidence level in Statistics?

In Statistics, the significance level is the probability of getting a completely different result from the condition where the null hypothesis is true. On the other hand, the confidence level is used as a range of similar values in a population.

We can specify the similarity between the significance level and the confidence level by the following formula:

$$\text{Significance level} = 1 - \text{Confidence level}$$

Question58: How do you assess the statistical **significance** of an insight?

You would perform hypothesis testing to determine statistical significance. First, you would state the null hypothesis and alternative hypothesis. Second, you would calculate the p-value, the probability of obtaining the observed results of a test assuming that the null hypothesis is true. Last, you would set the level of the significance (alpha) and if the p-value is less than the alpha, you would reject the null — in other words, the result is statistically significant.



117. What is the significance of p-value?

p-value typically ≤ 0.05

This indicates strong evidence against the null hypothesis; so you reject the null hypothesis.

p-value typically > 0.05

This indicates weak evidence against the null hypothesis, so you accept the null hypothesis.

p-value at cutoff 0.05

This is considered to be marginal, meaning it could go either way.

Q10. What is p-value?

When you perform a hypothesis test in statistics, a p-value can help you determine the strength of your results. p-value is a number between 0 and 1. Based on the value it will denote the strength of the results. The claim which is on trial is called the Null Hypothesis.

Low p-value (≤ 0.05) indicates strength against the null hypothesis which means we can reject the null Hypothesis. High p-value (≥ 0.05) indicates strength for the null hypothesis which means we can accept the null Hypothesis p-value of 0.05 indicates the Hypothesis could go either way. To put it in another way,

High P values: your data are likely with a true null. Low P values: your data are unlikely with a true null.

**Question21: What is Hypothesis Testing?**

Hypothesis testing is a form of statistical inference that uses data from a sample to draw conclusions about a population parameter or a population probability distribution.

There are 3 steps in Hypothesis Testing:

- State Null and Alternate Hypothesis
- Perform Statistical Test
- Accept or reject the Null Hypothesis

Question22: What is the Null and Alternate Hypothesis?

A null and alternate hypothesis is used in statistical hypothesis testing.

Null Hypothesis

- It states that the population parameter is equal to the assumed value
- It is an initial claim based on previous analysis or experience

Alternate Hypothesis

- It states that population parameters are equal or different to the assumed value
- It is what you might believe to be true or want to prove true



Question23: What are a p-value and its role in Hypothesis Testing?

P-value is the probability that a random chance generated the data or something else that is equal or rare.

P-values are used in hypothesis testing to decide whether to reject the null hypothesis or not.

- $p\text{-value} < \alpha$ – value

Means results are not in favor of the null hypothesis, reject the null hypothesis

Question41: What do you understand by Hypothesis Testing?

In Statistics, Hypothesis Testing is mainly used to see if a certain experiment generates meaningful results. It helps assess the statistical significance of insight by finding the odds of the results occurring by chance. In Hypothesis Testing, the first thing is to know the null hypothesis and then specify it. After that, the p-value is calculated, and if the null hypothesis is true, the other values are also determined. The alpha value specifies the significance, and you can adjust it accordingly.

If the p-value is less than the alpha value, the null hypothesis is rejected, but the null hypothesis is accepted if the p-value is greater than the alpha value. If the null hypothesis is rejected, it indicates that the results obtained are statistically significant.

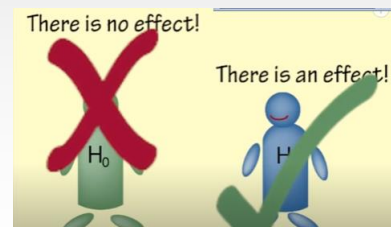


YOUTUBE ONERI VIDEO

<https://www.youtube.com/watch?v=0zZYBALbZgg>

<https://www.youtube.com/watch?v=eyknGvncKLw>

- › Hypothesis testing: step-by-step, p-value, t-test for difference of two means - Statistics Help
- › P-value in statistics: Understanding the p-value



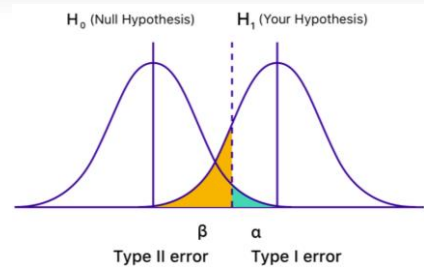
TYPE I - TYPE II ERROR

Hata Tipleri

Type I – II Error

Hata Tipleri

- Testin sonunda iki karardan biri verilecek:
 1. Null hipotezini reddetmek (reject)
 2. Null hipotezini reddedememek (fail to reject)
 Null hipotezi doğru olduğunda reddedilirse bir I. tip hata oluşur.
 Null hipotez yanlış olduğunda reddedilmezse, bir II. Tip hata oluşur
- TYPE I:** H_0 null hipotezi doğru olduğu halde reject yapmak (suçsuz birinin suçlu ilan edilmesi)
- TYPE II:** H_0 false olduğu halde H_0 'ı reddetmemek (not reject) (suçlu birini suçlu ilan etmede başarısız olma)



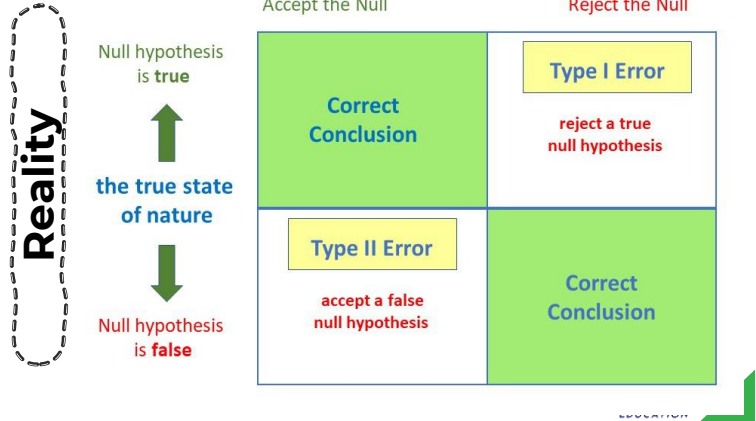
Hatalı Pozitif

Hatalı Negatif



Type I – II Error

Karar	H_0	
	H_0 doğru	H_0 yanlış
H_0 reddedilemez H_0 red	Doğru karar	II. Tip Hata
	I. Tip Hata	Doğru karar



Question34: What is the difference between type I vs. type II errors?

A type I error occurs when the null hypothesis true in the population is rejected. It is also known as false-positive.

A type II error occurs when the null hypothesis that is false in the population fails to get rejected. It is also known as a false-negative.



YOUTUBE ONERI VIDEO

<https://www.youtube.com/watch?v=edzQQFNzFjM>

► Type 1 and Type 2 errors -

		Reality (unknown)	
		Null hypothesis is true	Null hypothesis is false
Decision	Do not reject Null hypothesis (H_0)		
	Reject Null hypothesis		

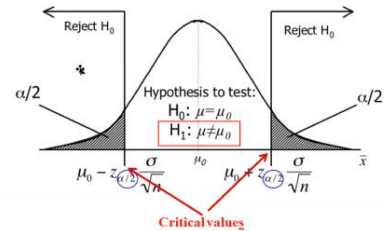
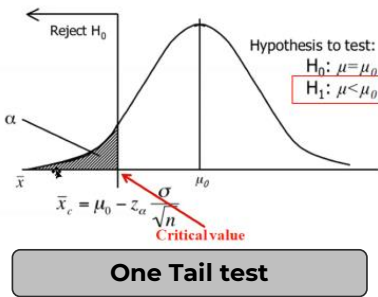
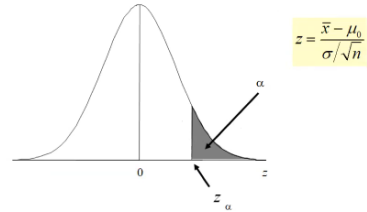
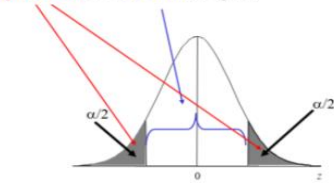
Note: The image shows a green figure in the bottom-right cell (Reject H_0 when Null hypothesis is true) and a cartoon character in the bottom-right cell (Reject H_0 when Null hypothesis is false).

ONE AND TWO TAILED TEST

Sol kuyruk-Sağ kuyruk - 2 Yönlü Test

One and Two Tailed Test

A **Rejection** and **Retain** Region



Two Tail test



One and Two Tailed Test

Tek – Çift Kuyruklu Test

Sol Kuyruklu Test :

Alternatif hipotez eşitsizlikten küçük sembolü (<) içeriyorsa, hipotez testi sol kuyruklu bir testtir

Sağ Kuyruklu Test :

Alternatif hipotez, büyük simgesini (>) içeriyorsa, hipotez testi, sağ kuyruklu bir testtir.

İki Yönlü Test :

- Alternatif hipotez, eşit olmayan (\neq) sembolü içeriyorsa, hipotez testi iki kuyruklu bir testtir. İki kuyruklu bir testte her kuyruğun 1/2 P alanı vardır

Two-sided (two-tailed) test

Used when the null contains an equality (=) or an inequality sign (\neq)

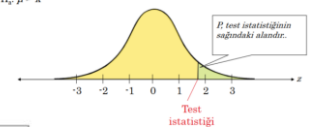
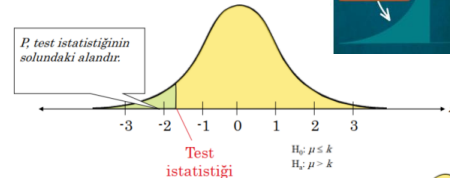


Used when the null doesn't contain equality or inequality sign (<, >, ≤, ≥)



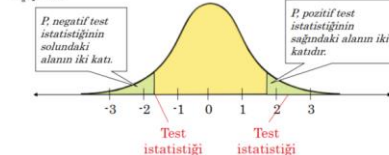
$$H_0: \mu \geq k$$

$$H_1: \mu < k$$



$$H_0: \mu = k$$

$$H_1: \mu \neq k$$



Steps for one-two tail test

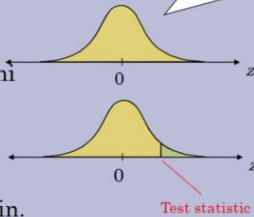
- İddiayı matematiksel ve sözlü olarak belirtin. Sıfır ve alternatif hipotezleri tanımlayın.

$$H_0: ? \quad H_a: ?$$

- Anlam düzeyini belirtin.
 $\alpha = ?$

- Standart örneklem dağılımını belirler ve grafiğini çizin.

Bu örneklem dağılımı H_0 'nin doğru olduğu varsayımına dayanmaktadır.



- Test istatistiğini ve standartlaştırılmış değerini hesaplayın. Çiziminize ekleyin.

- p -değerini bulun.

- Aşağıdaki karar kuralını kullanın.

p değeri anlamlılık seviyesine eşit veya bu değerden düşük mü?

hayır

H_0 reddedilemez

evet

H_0 red

- Kararı yorumlayın.

Example

Örnek

- Bir Denizde olması gereken kurşun seviyesi 10 p/m seviyesindedir. (μ_0)
- Popülasyon Normal dağılıma uygun olmakla beraber std.dev. $\sigma=1,5$ dir.
- Bizim alacağımız 40 farklı örneklerde kurşun seviyesi ölçüldü ve ortalama 10,5 (sample mean) bulundu
- Bu ortalamadaki fark $\alpha=0,05$ için (%95 güven için) istatistiksel olarak anlamlı mıdır ?



Solution

Step 1

**Assumptions
(Varsayımlar)**

Örnekler random olmalı ✓

Gözlemler Independent ✓

σ biliniyor veya $n > 30$ ✓

Step 2

**Hypotheses
(Hipotezler)**

Null Hypothesis ($H_0: \mu = 10$)

Alternate Hypothesis ($H_a: \mu > 10$)



Solution

Step 3

**Test Statistics
(Test İstatistikeleri)**

Z test

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

$$\bar{x} = 10.5$$

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{10.5 - 10}{1.5 / \sqrt{40}} = 2.1$$

Table Values of $Z_{1-(\alpha/2)}$ for Confidence Intervals

Confidence Level (%)	$Z_{1-(\alpha/2)}$
99.99	3.819
99.9	3.291
99	2.576
95	1.960
90	1.645
80	1.282

Z score = 2,1 > 1,96 olduğundan Null hipotez reject edilebilir.

Step 4

**P - Value
(P- Değeri)**

P-value =

$P(z > 2.1 \mid H_0 \text{ true}) = .0179$

Table 4 Normal Curve Areas
Standard normal probability to right hand tail
(the negative values of z are found by symmetry)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4051	.4011	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1333	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0903	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0722	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0352	.0344	.0336	.0329	.0322	.0314	.0307	.0300	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0175	.0171	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084

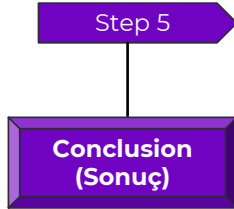
In [1]: `import scipy.stats as stats`

In [2]: `1-stats.norm.cdf(2.1)`

Out[2]: 0.017864420562816563



Solution



$$P - \text{Value} = 0,0179$$

$$\alpha = 0,05$$

$$P\text{-Value} < \alpha = 0,05$$

P-değeri önceden belirlenen α değerinden küçük olduğu için **Null hipotezi reject** olur (reddedilir)

Kurşun oranı önerildiği gibi 10 seviyesinde değildir, daha fazladır. Bu deniz sahilini kapatmak için yeterli data toplanmış demektir.



Example - 2

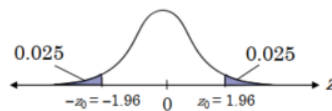
Red Bölgeleri ile Test Etme

Örnek:

Bir telefon şirketi, telefon görüşmesinin ortalama uzunluğunun 8 dakika olduğunu iddia ediyor. 58 telefon görüşmesinin rasgele seçilmesi sonucu, örneklem ortalaması 7.8 dakika ve standart sapma 0.5 dakika olarak bulunuyor. Bu iddiayı $\alpha = 0.05$ olarak desteklemek için yeterli kanıt var mı?

$$H_0: \mu = 8 \text{ (iddia)} \quad H_a: \mu \neq 8$$

Anlam düzeyi $\alpha = 0.05$.



Example - 2

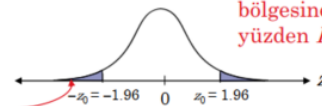
Örneğin devamı:

Bir telefon şirketi, telefon görüşmesinin ortalama uzunluğunun 8 dakika olduğunu iddia ediyor. 58 telefon görüşmesinin rasgele seçilmesi sonucu, örneklem ortalaması 7.8 dakika ve standart sapma 0.5 dakika olarak bulunuyor. Bu iddiayı $\alpha = 0.05$ olarak desteklemek için yeterli kanıt var mı?

$H_0: \mu = 8$ (iddia) $H_a: \mu \neq 8$

Standartlaştırılmış test istatistiği

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{7.8 - 8}{0.5 / \sqrt{58}} \approx -3.05.$$



Test istatistiği ret bölgesine düşüyor bu yüzden H_0 reddedilir.

% 5 anlamlılık düzeyinde, bir telefon görüşmesinin ortalama süresinin 8 dakika olduğu iddiasını reddetmek için yeterli kanıt vardır.



Hypothesis Testing by t test

- Şu an kadar Z testi ile yaptık ama eğer **$n < 30$** durumu söz konusu ise ve popülasyona ait dağılımla alakalı bilgimiz yok ise **t testine** yönleniyoruz

- If $H_A: \mu < \mu_0$, $p_{\text{obs}} = P(T \leq t)$,
- If $H_A: \mu > \mu_0$, $p_{\text{obs}} = P(T \geq t)$,
- If $H_A: \mu \neq \mu_0$, $p_{\text{obs}} = 2 \times P(T \geq |t|)$,

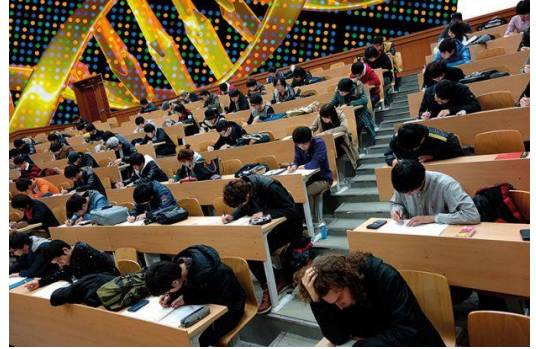


Example for t test

Örnek:

Bir eğitim kurumunun 1000 öğrencisi vardır. Öğrencilerin ortalama IQ seviyeleri daha önce 110 bulunmuştur. Yeni durumda 20 randomly öğrencinin IQ bakılmıştır. Bu 20 öğrencinin

- Ortalama IQ: 108 Std.Dev: 10 ise
- $\alpha = 0,01$ için bu sonuçlar anlamlı mıdır ?



Solution

Step 1

Assumptions
(Varsayımlar)

Değişken quantitative olmalı ✓

Data üretimi randomly ✓

Popülasyon distr. Yaklaşık normal dağılıma uygun ✓

Step 2

Hypotheses
(Hipotezler)

Null Hypothesis ($H_0: \mu \geq 110$)

Alternate Hypothesis ($H_a: \mu < 110$)



Solution

Step 3

t test

**Test Statistics
(Test İstatistikeleri)**

$$\bar{x} = 108$$

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

$$t = \frac{108 - 110}{10 / \sqrt{20}} = \frac{-2}{2.236} = -0.894$$

t score = -0,894 > -2,539 olduğundan Null hipotez fail to reject denilebilir. .

Step 4

**P - Value
(P- Değeri)**

P-value =

$$P(t < -0.894 \mid H_0 \text{ true}) = .1913$$

```
In [1]: import scipy.stats as stats
```

```
In [2]: stats.t.cdf(-0.894, 19)
```

```
Out[2]: 0.19125344283171025
```



Solution

Step 5

**Conclusion
(Sonuç)**

P - Value = 0,1913

$\alpha = 0,01$

P-Value > α

P-değeri önceden belirlenen α değerinden büyük olduğu için Null hipotezi **fail to reject olur (kabul edilebilir)**

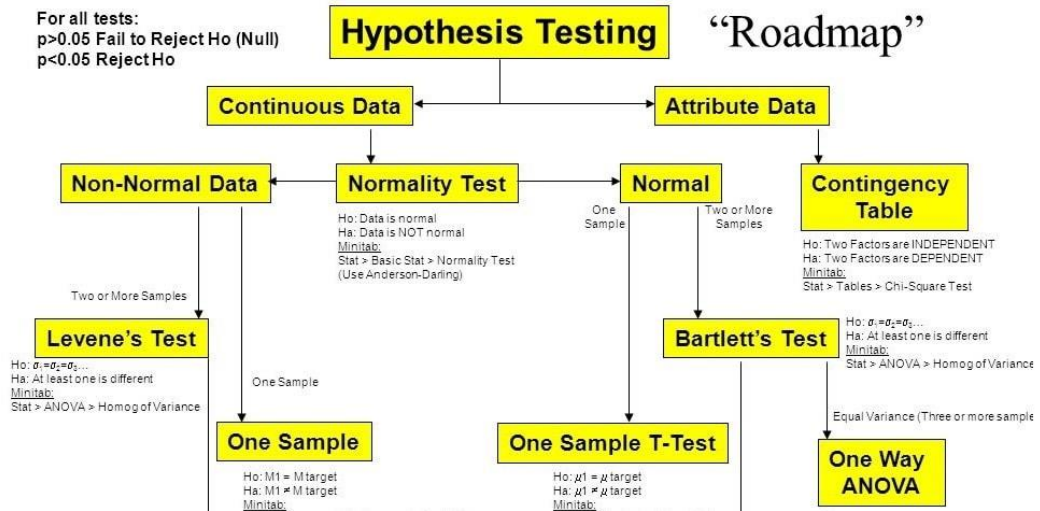
20 öğrenciden alınan ortalama her ne kadar düşük çıksa da bu istatistiksel olarak anlamlı değildir, okulun ortalamasının 110 olduğu yine söylenebilir





Roadmap

For all tests:
 $p > 0.05$ Fail to Reject H_0 (Null)
 $p < 0.05$ Reject H_0



Python Coding

- › One-Two tail test
- › Z test
- › t test

Hypothesis_Testing-One-two_tail.ipynb