



HİPOTEZ TESTLERİ

Ramazan Tetik

- 01** Hakkımda
- 02** Normallik Testleri
- 03** Korelasyon Testleri
- 04** Sabitlik Testleri
- 05** Parametrik Testler
- 06** Parametrik Olmayan Testler

SUNUM PLANI



HAKKIMDA

Ben Ramazan Tetik. Aydın Adnan Menderes
Üniversitesi Bilgisayar Mühendisliği
bölümünde 3. sınıfı geçtim.

1. NORMALLİK TESTLERİ

- Normallik sınavlarının amacı verilmiş bir veri dizisinin **normal dağılıp dağılmadığını** incelemektir.
- Hipotezler şu şekilde kurulur:
- **H₀**: Veri normal dağılıma uygundur.
- **H₁**: Veri normal dağılıma uygun değildir.
- Bu çalışmada 3 çeşit normallik testi uygulayacağız.
 - 1.Shapiro-Wilk Testi
 - 2.D'Agostino's K² Tesi
 - 3.Anderson-Darling Testist

SEÇİLEN VERİ SETLERİ VE PROBLEMLER

1

Shapiro-Wilk Testi~Hava Kalitesi

- Bu veri setindeki AQI (Hava Kalitesi Endeksi) değerlerinin normal dağılıma uyup uymadığını belirleyiniz. $\alpha=0.05$
- CO (Karbon Monoksit) konsantrasyonlarının normal dağılıma uyup uymadığını test ediniz. $\alpha=0.05$

2

D'Agostino's K^2 Testi~Amerika'da İşsizlik

- İşsizlik oranının normal dağılıp dağılmadığını test ediniz. $\alpha=0.05$
- İstihdam oranının normal dağılıma uygun olup olmadığını test ediniz.

3

Anderson-Darling Testi ~ Diyabet

- Glikoz seviyelerinin normal dağılıma uygun olup olmadığını test ediniz.
 - Vücut kitle indeksinin normal dağılıma uygun olup olmadığını test ediniz.

ÇÖZÜMLER

Hava Kalitesi çözümü

```
AQI = df['AQI'].dropna()
```

```
stat,p = stats.shapiro(AQI)
print('stat=%.3f, p=%.3f' % (stat, p))
```

```
stat=0.977, p=0.263
```

```
alpha = 0.05
if p>alpha:
    print("H0 kabul edilir, veri normal dağılıma sahiptir.")
else:
    print("H1 kabul edilir, veri normal dağılıma sahip değildir.")
```

```
H0 kabul edilir, veri normal dağılıma sahiptir.
```

```
stat, p = normaltest(unemployment_rates)
print('stat=%.3f, p=%.3f' % (stat, p))
```

```
stat=5637.280, p=0.000
```

```
if p > 0.05:
    print('H0 kabul edilir. Veri normal dağılmamıştır.')
else:
    print('H0 reddedilir. Veri normal dağılmamıştır.')
```

```
H0 reddedilir. Veri normal dağılmamıştır.
```

İşsizlik Oranı Çözümü

ÇÖZÜMLER

Vücut Kitle Endeksi Çözümü

```
ad_test = stats.anderson(bmi_values, dist='norm')
print("Anderson-Darling Testi: Test istatistiği = {}".format(ad_test.statistic))
```

```
Anderson-Darling Testi: Test istatistiği = 4.1891562244597935
```

```
for i in range(len(ad_test.critical_values)):
    sl, cv = ad_test.significance_level[i], ad_test.critical_values[i]
    if ad_test.statistic < ad_test.critical_values[i]:
        print('%.3f: %.3f, data looks normal (fail to reject H0)' % (sl, cv))
    else:
        print('%.3f: %.3f, data does not look normal (reject H0)' % (sl, cv))
```

```
15.000: 0.573, data does not look normal (reject H0)
10.000: 0.653, data does not look normal (reject H0)
5.000: 0.783, data does not look normal (reject H0)
2.500: 0.913, data does not look normal (reject H0)
1.000: 1.086, data does not look normal (reject H0)
```

2.KORELASYON TESTLERİ

- Bu bölümde, iki örneğin **ilişkili olup olmadığını** kontrol etmek için kullanabileceğiniz istatistiksel testler listelenmektedir.
 - Parametrik bir testtir.
 - Normal dağılım şartı vardır.
-
- **Kullanılan Korelasyon Testleri**
 - Pearson Korelasyon Katsayısı
 - Spearman'ın Sıra Korelasyonu
 - Kendall'ın Sıra Korelasyonu
 - Ki-Kare Testi

SEÇİLEN VERİ SETLERİ VE PROBLEMLER

1 Pearson Korelasyon Katsayısı ~ Hava Kalitesi

- Azot oksitlerin (NOx) konsantrasyonu ile hava kalitesi endeksi (AQI) arasındaki ilişki nedir? Bu iki değişken arasında anlamlı bir doğrusal ilişki var mı? $\alpha=0.05$

2 Spearman Sıra Korelasyon Katsayısı ~ Öğrenci performans tahmini

- Matematik Notu (math score) ve Okuma Notu (reading score) arasındaki ilişki miktarını ve yönünü 0,05 anlamlılık düzeyinde araştırın.

3 Kendall'ın Sıra Korelasyonu Diyabet ~ Kalp Hastalıkları

- Yaş ile ST depresyonu arasındaki ilişki miktarını ve yönünü 0,05 anlamlılık düzeyinde araştırın.
- Yaş ile Kolesterol arasındaki ilişki miktarını ve yönünü 0,05 anlamlılık düzeyinde araştırın.

4 Ki-kare Testi ~ Kalp Hastalıkları

- Cinsiyet ile göğüs ağrısı türü arasındaki ilişki miktarını ve yönünü 0,05 anlamlılık düzeyinde araştırın.

ÇÖZÜMLER

Hava Kalitesi çözümü

```
AQI = df['AQI']
Nox = df["NOx  in ug/m3"]
```

```
from scipy.stats import pearsonr
stat, p = pearsonr(AQI, Nox)
print('stat=%.3f, p=%.3f' % (stat, p))
if p > 0.05:
    print('H0 hipotezi kabul edilir.AQI ile NOx arasında anlamlı bir doğrusal ilişki yoktur.')
else:
    print('H1 hipotezi kabul edilir.AQI ile NOx arasında anlamlı bir doğrusal ilişki vardır.')

stat=0.255, p=0.038
H1 hipotezi kabul edilir.AQI ile NOx arasında anlamlı bir doğrusal ilişki vardır.
```

```
from scipy.stats import spearmanr
stat, p = spearmanr(df["math score"], df["reading score"])
print('stat=%.3f, p=%.3f' % (stat, p))
if p > 0.05:
    print('H0 kabul edilir.Matematik Notu (math score) ve Okuma Notu (reading score) arasında anlamlı bir ilişki yoktur.')
else:
    print('H1 kabul edilir.Matematik Notu (math score) ve Okuma Notu (reading score) arasında anlamlı bir ilişki vardır.')

stat=0.803, p=0.000
Matematik Notu (math score) ve Okuma Notu (reading score) arasında anlamlı bir ilişki vardır.
```

Öğrenci
performans
tahmini

ÇÖZÜMLER

Kalp Hastalıkları ➤

```
from scipy.stats import kendalltau
data1 = df["age"]
data2 = df["oldpeak"]
stat, p = kendalltau(data1, data2)
print('stat=%.3f, p=%.3f' % (stat, p))
if p > 0.05:
    print('H0 kabul edilir.Yaş ile ST depresyonu arasında sıralı bir ilişki yoktur.')
else:
    print('H1 kabul edilir.Yaş ile ST depresyonu arasında sıralı bir ilişki vardır.')

stat=0.179, p=0.000
H1 kabul edilir.Yaş ile ST depresyonu arasında sıralı bir ilişki vardır.
```

```
from scipy.stats import chi2_contingency
contingency_table = pd.crosstab(df['sex'], df['cp'])
stat, p, dof, expected = chi2_contingency(contingency_table)
print('stat=%.3f, p=%.3f' % (stat, p))
if p > 0.05:
    print('H0 kabul edilir.Cinsiyet ile göğüs ağrısı türü arasında anlamlı bir ilişki yoktur.')
else:
    print('H0 reddedilir.Cinsiyet ile göğüs ağrısı türü arasında anlamlı bir ilişki vardır.')

stat=6.991, p=0.072
H0 kabul edilir.Cinsiyet ile göğüs ağrısı türü arasında anlamlı bir ilişki yoktur.
```

➤ Kalp Hastalıkları

3.SABİTLİK TESTLERİ

- Durağanlık bileşeni kısaca serinin istatistiksel özelliklerinin zaman içerisinde değişmemesidir.
- Zaman serilerinde durağanlığı Agumented Dickey-Fuller (ADF) veya Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test ile istatistiksel olarak test edebiliriz.
- Bir zaman serisinin ortalamasının veya varyansının zaman içinde değişip değişmediğini test etmek için kullanılır.
- Örneğin Dickey-Fuller bir zaman serisi için null hipotezi (H_0) olarak serinin durağan olmadığını ve alternatif hipotezi olarak serinin durağan olduğunu kabul eder. Eğer test sonucunda null hipotezi reddedilirse, seri durağan olarak kabul edilir

SEÇİLEN VERİ SETLERİ VE PROBLEMLER

1 ADF testi ~ Güneş Lekeleri

- H_0 : Serinin birim kökü vardı
- H_1 : Serinin birim kökü yoktur.

2 KPSS Testi ~ Güneş Lekeleri

- H_0 : Süreç trend durağandır.
- H_1 : Seri birim köke sahiptir (seri durağan değildir).

ÇÖZÜMLER

ADF ➤

```
adf_test(sunspots["SUNACTIVITY"])
```

Results of Dickey-Fuller Test:

Test Statistic	-2.837781
p-value	0.053076
#Lags Used	8.000000
Number of Observations Used	300.000000
Critical Value (1%)	-3.452337
Critical Value (5%)	-2.871223
Critical Value (10%)	-2.571929
dtype: float64	

```
from statsmodels.tsa.stattools import adfuller

def adf_test(timeseries):
    print("Results of Dickey-Fuller Test:")
    dftest = adfuller(timeseries, autolag="AIC")
    dfoutput = pd.Series(
        dftest[0:4],
        index=[
            "Test Statistic",
            "p-value",
            "#Lags Used",
            "Number of Observations Used",
        ],
    )
    for key, value in dftest[4].items():
        dfoutput["Critical Value (%s)" % key] = value
    print(dfoutput)
```

-0,05'lik anlamlılık düzeyi ve
ADF testinin
p-değeri temel alındığında,
sıfır hipotezi reddedilemez.
Dolayısıyla seri durağan değildir.

ÇÖZÜMLER

KPSS

```
from statsmodels.tsa.stattools import kpss

def kpss_test(timeseries):
    print("Results of KPSS Test:")
    kpsstest = kpss(timeseries, regression="c", nlags="auto")
    kpss_output = pd.Series(
        kpsstest[0:3], index=["Test Statistic", "p-value", "Lags Used"]
    )
    for key, value in kpsstest[3].items():
        kpss_output["Critical Value (%s)" % key] = value
    print(kpss_output)
```

```
kpss_test(sunspots["SUNACTIVITY"])
```

Results of KPSS Test:

Test Statistic	0.669866
p-value	0.016285
Lags Used	7.000000
Critical Value (10%)	0.347000
Critical Value (5%)	0.463000
Critical Value (2.5%)	0.574000
Critical Value (1%)	0.739000
dtype: float64	

0,05'lik anlamlılık düzeyi ve KPSS testinin p-değeri temel alındığında, alternatif lehine sıfır hipotezini reddetmek için kanıt vardır. Bu nedenle, seri KPSS testine göre durağan değildir.

4. PARAMETRİK TESTLER

- Parametrik testler, bir veri setindeki dağılımın belirli bir parametreye göre modellendiği istatistiksel testlerdir. Bu testler, veri setinin belirli bir dağılıma (genellikle normal dağılım) uyduğunu varsayar ve bu varsayımdan doğrultusunda hipotez testleri yapar. Parametrik testler, verilerin belirli bir dağılıma uygun olduğu durumlarda daha güçlü ve hassas sonuçlar verir.
- **Kullanılan Parametrik Testleri**
- Öğrencinin t-testi
- Eşleştirilmiş öğrenci t-testi:
- ANOVA Testi
- Tekrarlı Ölçümler ANOVA

SEÇİLEN VERİ SETLERİ VE PROBLEMLER

1 Öğrenci-T Testi ~ Iris

- Bir biyolog, setosa ve versicolor türleri arasında sepal length (çanak uzunluğu) bakımından anlamlı bir fark olup olmadığını incelemek istemektedir.

2 Eşleştirilmiş Öğrenci-T Testi ~ Akademik Slaytlar

- Aşağıdaki tabloda, tedavi gören bir grup hastada, ilaç almadan önce ve 6 hafta sonrası için ölçülen toplam kolesterol değerleri verilmiştir. İlacın toplam kolesterolde azalmaya yol açacağı sonucuna varmamız için $\alpha=0.01$ anlamlılık düzeyinde yeterli kanıt var mı?

3 ANOVA Testi ~ Akademik Slaytlar

- Üç farklı bölgede oturan kişilerin boy ve kilo ölçümleri yapılarak vücut kitle indeksi (VKİ) değerleri hesaplanmıştır. Bu üç farklı bölgede yaşayan kişilerin VKİ ortalamaları arasında fark olup olmadığını $\alpha=0,05$ yanılma düzeyinde araştırınız.

4 Tekrarlı Ölçümlerle ANOVA Testi ~ Akademik Slaytlar

- Bir grup katılımcıya 3 çeşit (bira-şarap-su) içeceğini 3 farklıbetimlemesi(negatif-pozitif ve nötr) gösterildiği ve katılımcıların bu içeceklerle karşı olan tutumlarının ölçüldüğü verimiz yanda verilmiştir.Toplamda aynı bireylere gösterilen 9 farklı durumu içeren 2 faktöre sahip olduğu için 2 faktör tekrarlı ölçüler ANOVA ile verimizi analiz edin.

ÇÖZÜMLER

Iris çözümü



```
# T-Testi uygulama
t_stat, p_value = stats.ttest_ind(setosa, versicolor)

# Sonuçları yazdırma
print('T-Statistic:', t_stat)
print('P-Value:', p_value)

# Sonuçları yorumlama
alpha = 0.05
if p_value < alpha:
    print("Null Hipotez reddedilir; iki tür arasında anlamlı bir fark vardır.")
else:
    print("Null Hipotez reddedilemez; iki tür arasında anlamlı bir fark yoktur.")
```

```
# Perform the paired t-test
t_stat, p_value = stats.ttest_rel(data['Before'], data['After'])
T-Statistic: -10.52098626754911
P-Value: 8.985235037487079e-18
Null Hipotez reddedilir; iki tür arasında anlamlı bir fark vardır.
```

```
# Print the results
print(f'Test statistic: {t_stat:.4f}')
print(f'P-value: {p_value:.4f}')
```

```
Test statistic: 15.4134
P-value: 0.0000
```

```
# Sonuçları yorumlama
alpha = 0.01
if p_value > alpha:
    print("H0 kabul edilir.  $\mu_1 = \mu_2$  (esleştirilmiş popülasyon ortalamaları eşittir")
else:
    print("H1 kabul edilir; esleştirilmiş popülasyon ortalamaları eşit değildir.")
```

H1 kabul edilir; esleştirilmiş popülasyon ortalamaları eşit değildir.

Tedavi Çözümü

ÇÖZÜMLER

Vücut Kitle Endeksi



```
# 2-faktör Tekrarlı Ölçümler ANOVA'yı uygulama
aovrm2 = AnovaRM(df_long, 'tutum', 'kat1|mc1', within=['icecek_durum'])
res2 = aovrm2.fit()
```

```
# Sonuçların özeti
print(res2.summary())
```

Anova

F Value	Num DF	Den DF	Pr > F
30.8348	8.0000	152.0000	0.0000

```
# F-değeri
fvalue = ms_between / ms_within
print("ANOVA F değeri:", fvalue)
```

ANOVA F değeri: 10.723384818632466

```
# Kritik değer
alpha = 0.05
kritik_deger = stats.f.ppf(1 - alpha, df_between, df_within)
print("Kritik değer:", kritik_deger)
```

Kritik değer: 3.1907273359284987

```
# Sonucu yorumlayalım
if fvalue > kritik_deger:
    print("Sonuç: Bölgeler arasında istatistiksel olarak anlamlı fark vardır ( $H_0$  reddedilir).")
else:
    print("Sonuç: Bölgeler arasında istatistiksel olarak anlamlı fark yoktur ( $H_0$  kabul edilir).")
```

Sonuç: Bölgeler arasında istatistiksel olarak anlamlı fark vardır (H_0 reddedilir).



Bira-Şarap-Su

Null hipotezi reddediyoruz,
icecek durumu arasında
anlamlı bir fark var.

5. PARAMETRİK OLMAYAN TESTLER

- Parametrik olmayan testler, verilerin belirli bir dağılıma (örneğin, normal dağılım) uyduğu varsayımlını gerektirmeyen istatistiksel testlerdir. Bu testler, verilerin dağılımı hakkında daha az varsayımda bulunur ve özellikle veri seti küçük, normallik varsayımlı sağlanmıyorsa veya veriler sıralama (ordinal) ölçüğünde ise kullanılır.
- **Kullanılan Parametrik Olmayan Testleri**
 - Mann – Whitney U Testi
 - Wilcoxon Signed Rank
 - Friedman Testi
 - Kruskal-Wallis Testi

SEÇİLEN VERİ SETLERİ VE PROBLEMLER

1 Mann – Whitney U Testi ~ Kalp Hastalıkları

- Erkek ve kadın bireyler arasında serum kolesterol seviyeleri açısından anlamlı bir fark olup olmadığını inceleyin.

2 Wilcoxon Signed Rank ~ Akademik Slaytlar

- Sürpriz bir sınav yapılmış ve 6 öğrencinin notları aşağıdaki gibi verilmiştir. Bir aylık uygulama süresi verildikten sonra aynı sınav tekrar yapılmış ve aşağıdaki puanlar alınmıştır. Parametrik olmayan bir test kullanarak elde edilen notlarda bir fark olup olmadığını belirleyin. Anlamlılık düzeyi 0,05'tir.

3 Friedman Testi ~ Akademik Slaytlar

- 7 rastgele kişiye 3 farklı ilaç verildi ve her kişi için ilaçlara karşılık gelen tepki süresi not edildi. Tüm 3 ilaçın aynı olasılık dağılımına sahip olduğu iddiasını %5 önem düzeyinde test edin.

4 Kruskal-Wallis Testi ~ Akademik Slaytlar

- Araştırma ve Geliştirme ekibinin üç farklı motor yağı uygulamasının otomobillerin kilometre performansında farka yol açıp açmadığını belirlemek istiyor. Her grup tam olarak bir motor yağıyla katkılanyor. Daha sonra aynı pistte 20 kilometre yol kat etmelerine izin veriliyor ve yolculukları sona erdiğinde kilometre not ediliyor.

ÇÖZÜMLER

```
# Mann-Whitney U testi uygulama
stat, p_value = mannwhitneyu(male_chol, female_chol)
print('Mann-Whitney U Testi Statistiği:', stat)
print('P-Degeri:', p_value)
alpha = 0.05
if p_value < alpha:
    print("Null Hipotez reddedilir; Erkek ve kadın bireyler arasında serum kolesterol seviyeleri açısından anlamlı bir fark var")
else:
    print("Null Hipotez reddedilemez; Erkek ve kadın bireyler arasında serum kolesterol seviyeleri açısından anlamlı bir fark yok")

Mann-Whitney U Testi Statistiği: 7852.0
P-Degeri: 0.006728094084218456
Null Hipotez reddedilir; Erkek ve kadın bireyler arasında serum kolesterol seviyeleri açısından anlamlı bir fark var
```

Kalp Hastalıkları

```
# İki test arasındaki farkları hesaplayalım
farklar = test_2 - test_1
# Wilcoxon Signed Rank Testini uygulayalım
statistic, p_value = wilcoxon(farklar)
print(f"Test İstatistiği (W2): {statistic}")
print(f"p-değeri: {p_value}")
alpha = 0.05
if p_value > alpha:
    print("H0 hipotezini reddedemiyoruz, iki test arasındaki fark istatistiksel olarak anlamlı değildir.")
else:
    print("H0 hipotezini reddediyoruz, iki test arasındaki fark istatistiksel olarak anlamlıdır.")
```

Test İstatistiği (W2): 3.5

p-değeri: 0.21875

H0 hipotezini reddedemiyoruz, iki test arasındaki fark istatistiksel olarak anlamlı değildir.

Sınav Farkı

ÇÖZÜMLER

İlaç Etkisi



```
# Perform the Friedman test
stat,p = friedmanchisquare(y[:, 0], y[:, 1], y[:, 2])

# Display the result
print(f'Statistics : {stat}')
print(f'P value : {p}')
# Compare p-value with alpha
if p <= alpha:
    print(f"Reject the null hypothesis (p-value: {p:.3f} ≤ α: {alpha})")
else:
    print(f"Fail to reject the null hypothesis (p-value: {p:.3f} > α: {alpha})")

Statistics : 8.857142857142847
P value : 0.011931522535756207
Reject the null hypothesis (p-value: 0.012 ≤ α: 0.05)
```

```
# Import libraries
from scipy import stats
# Conduct the Kruskal-Wallis Test
result = stats.kruskal(data_group1, data_group2, data_group3)
print(result)
if result.pvalue <= alpha:
    print(f"Reject the null hypothesis (p-value: {result.pvalue:.3f} ≤ α: {alpha})")
else:
    print(f"Fail to reject the null hypothesis (p-value: {result.pvalue:.3f} > α: {alpha})")

KruskalResult(statistic=3.492418772563175, pvalue=0.17443390338074047)
Fail to reject the null hypothesis (p-value: 0.174 > α: 0.05)
```

Motor Yağı
Etkisi

TEŞEKKÜRLER