



Hacettepe Üniversitesi

FEN FAKÜLTESİ

İSTATİSTİK BÖLÜMÜ

PARAMETRİK OLMAYAN İSTATİSTİKSEL YÖNTEMLER

ÖDEVİ

DOÇ. DR. SEMRA TÜRKAN

EDA YAREN ÖZEL

2200329007

- 1) Bir gıda üreticisinin piyasaya arz etmek üzere 4 farklı formülasyonla işlenmiş salam çeşidi sekiz uzman tarafından 1-10 arasında puanlanmıştır. Puan değerleri aşağıdaki tabloda verilmiştir. Salam çeşitlerinin lezzet bakımından farklı olup olmadığını %5 önem düzeyinde test ediniz.

	SALAM ÇEŞİDİ			
UZMAN	A	B	C	D
1	1	6	8	7
2	8	7	9	8
3	7	5	9	8
4	7	6	10	6
5	8	7	9	8
6	6	7	9	9
7	7	8	8	7
8	6	5	7	5

R Veri Girişi:

```
# Veri Girişi-1:
a_salam <- c(1,8,7,7,8,6,7,6)
b_salam <- c(6,7,5,6,7,7,8,5)
c_salam <- c(8,9,9,10,9,9,8,7)
d_salam <- c(7,8,8,6,8,9,7,5)
salam_cesidi <- factor(rep(c("A","B","C","D"),each=8))
uzman <- factor(rep(c("1","2","3","4","5","6","7","8"),times=4))
puan <- c(a_salam,b_salam,c_salam,d_salam)
veril <- data.frame(puan,salam_cesidi,uzman)
```

Her Grup İçin Normallik Varsayımın Test Edilmesi:

Her gruba ait verilerin sayısı 50'den az olduğu için normallik kontrolü yapılırken Shapiro-Wilk testi uygulanır.

H_0 : A salam çeşidine ait puanların dağılımı ile normal dağılım arasında fark yoktur.
 H_1 : A salam çeşidine ait puanların dağılımı ile normal dağılım arasında fark vardır.

Python Sonuç:

```
In [1]: import numpy as np
        from scipy.stats import shapiro
        import statistics as st

In [2]: A=np.array([1,8,7,7,8,6,7,6])
        shapiro(A)

Out[2]: ShapiroResult(statistic=0.7164304852485657, pvalue=0.0034505012445151806)
```

R Sonuç:

```
> shapiro.test(a_salam)

      Shapiro-Wilk normality test

data:  a_salam
W = 0.71643, p-value = 0.00345
```

Testin sonucuna göre bulunan p değeri önem düzeyinden küçük olduğu için ($p \approx 0.0034 < \alpha = 0.05$) H_0 hipotezi reddedilir. Bu yüzden farklılığın olup olmadığını hesaplamak için parametrik testler kullanılır.

YORUM: Uzmanlar tarafından değerlendirilen A salam çeşidinin puanlarının %5 önem düzeyinde normallik varsayımını sağlamadığını söyleyebiliriz.

H_0 : B salam çeşidine ait puanların dağılımı ile normal dağılım arasında fark yoktur.

H_1 : B salam çeşidine ait puanların dağılımı ile normal dağılım arasında fark vardır.

Python Sonuç:

```
In [3]: B=np.array([6,7,5,6,7,7,8,5])
        shapiro(B)

Out[3]: ShapiroResult(statistic=0.9116305112838745, pvalue=0.36568841338157654)
```

R Sonuç:

```
> shapiro.test(b_salam)

      Shapiro-Wilk normality test

data:  b_salam
W = 0.91163, p-value = 0.3657
```

Testin sonucuna göre bulunan p değeri önem düzeyinden büyük olduğu için ($p \approx 0.36 > \alpha = 0.05$) H_0 hipotezi reddedilemez.

YORUM: Uzmanlar tarafından değerlendirilen B salam çeşidinin puanlarının %5 önem düzeyinde normallik varsayımını sağladığını söyleyebiliriz.

H_0 : C salam çeşidine ait puanların dağılımı ile normal dağılım arasında fark yoktur.

H_1 : C salam çeşidine ait puanların dağılımı ile normal dağılım arasında fark vardır.

Python Sonuç:

```
In [4]: C=np.array([8,9,9,10,9,9,8,7])
        shapiro(C)

Out[4]: ShapiroResult(statistic=0.9055967330932617, pvalue=0.3240951895713806)
```

R Sonuç:

```
> shapiro.test(c_salam)

      Shapiro-Wilk normality test

data:  c_salam
W = 0.9056, p-value = 0.3241
```

Testin sonucuna göre bulunan p değeri önem düzeyinden büyük olduğu için ($p \approx 0.32 > \alpha = 0.05$) H_0 hipotezi reddedilemez.

YORUM: Uzmanlar tarafından değerlendirilen C salam çeşidinin puanlarının %5 önem düzeyinde normallik varsayımını sağladığını söyleyebiliriz.

H_0 : D salam çeşidine ait puanların dağılımı ile normal dağılım arasında fark yoktur.

H_1 : D salam çeşidine ait puanların dağılımı ile normal dağılım arasında fark vardır.

Python Sonuç:

```
In [5]: D=np.array([7,8,8,6,8,9,7,5])
        shapiro(D)

Out[5]: ShapiroResult(statistic=0.9380493760108948, pvalue=0.5919957160949707)
```

R Sonuç:

```
> shapiro.test(d_salam)

      Shapiro-Wilk normality test

data:  d_salam
W = 0.93805, p-value = 0.592
```

Testin sonucuna göre bulunan p değeri önem düzeyinden büyük olduğu için ($p \approx 0.59 > \alpha = 0.05$) H_0 hipotezi reddedilemez.

YORUM: Uzmanlar tarafından değerlendirilen D salam çeşidinin puanlarının %5 önem düzeyinde normallik varsayımını sağladığını söyleyebiliriz.

Her Grup İçin Özet İstatistiklerin Hesaplanması;

Python Sonuç;

```
In [9]: import pandas as pd
```

```
In [10]: df_A_salam=pd.DataFrame(A_salam)
```

```
In [11]: df_A_salam.rename(columns={0:'puanlar'}, inplace=True)  
df_A_salam
```

```
Out[11]:
```

	puanlar
0	1
1	8
2	7
3	7
4	8
5	6
6	7
7	6

```
In [12]: df_A_salam.describe()
```

```
Out[12]:
```

	puanlar
count	8.000000
mean	6.250000
std	2.251983
min	1.000000
25%	6.000000
50%	7.000000
75%	7.250000
max	8.000000

```
In [13]: df_B_salam=pd.DataFrame(B_salam)
```

```
In [14]: df_B_salam.rename(columns={0:'puanlar'}, inplace=True)  
df_B_salam
```

```
Out[14]:
```

	puanlar
0	6
1	7
2	6
3	6
4	7
5	7
6	8
7	6

```
In [15]: df_B_salam.describe()
```

```
Out[15]:
```

	puanlar
count	8.000000
mean	6.375000
std	1.060660
min	5.000000
25%	5.750000
50%	6.500000
75%	7.000000
max	8.000000

A türü salam çeşidine uzmanlar tarafından verilen ortalama lezzet puanı 6.25'dir. Puanların ortanca değeri 7 puandır ve standart sapması 2.25 olduğu için verilerin ortalamaya yakın yerlerde dağıldıklarını söyleyebiliriz.

B türü salam çeşidine uzmanlar tarafından verilen ortalama lezzet puanı 6.375'dir. Puanların ortanca değeri 6.5 puandır ve standart sapması 1.06 olduğu için verilerin ortalamaya yakın yerlerde dağıldıklarını söyleyebiliriz.

```
In [16]: df_C_salam=pd.DataFrame(C_salam)
```

```
In [17]: df_C_salam.rename(columns={0:'puanlar'}, inplace=True)  
df_C_salam
```

```
Out[17]:
```

	puanlar
0	8
1	9
2	9
3	10
4	9
5	9
6	8
7	7

```
In [18]: df_C_salam.describe()
```

```
Out[18]:
```

	puanlar
count	8.000000
mean	8.625000
std	0.916125
min	7.000000
25%	8.000000
50%	9.000000
75%	9.000000
max	10.000000

```
In [19]: df_D_salam=pd.DataFrame(D_salam)
```

```
In [20]: df_D_salam.rename(columns={0:'puanlar'}, inplace=True)  
df_D_salam
```

```
Out[20]:
```

	puanlar
0	7
1	8
2	8
3	6
4	8
5	9
6	7
7	5

```
In [21]: df_D_salam.describe()
```

```
Out[21]:
```

	puanlar
count	8.000000
mean	7.250000
std	1.28174
min	5.000000
25%	6.750000
50%	7.500000
75%	8.000000
max	9.000000

C türü salam çeşidine uzmanlar tarafından verilen ortalama lezzet puanı 8.625'dir. Puanların ortanca değeri 9 puandır ve standart sapması 0.916 olduğu için verilerin ortalamaya yakın yerlerde dağıldıklarını söyleyebiliriz.

D türü salam çeşidine uzmanlar tarafından verilen ortalama lezzet puanı 7.25'dir. Puanların ortanca değeri 7.5 puandır ve standart sapması 1.28 olduğu için verilerin ortalamaya yakın yerlerde dağıldıklarını söyleyebiliriz.

R Sonuçı:

```
> summary(a_salam)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  1.00   6.00   7.00   6.25   7.25   8.00
> summary(b_salam)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 5.000  5.750  6.500  6.375  7.000  8.000
> summary(c_salam)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 7.000  8.000  9.000  8.625  9.000 10.000
> summary(d_salam)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 5.00   6.75   7.50   7.25   8.00   9.00
```

- 2) Dört farklı budama yöntemi ile budanarak malçlı alçak tünelde yetiştirilmiş bir kavun çeşidinin hasat sonunda meyve ağırlıkları gram cinsinden ölçülüp aşağıdaki tabloda verilmiştir. Budama yöntemlerinin meyve ağırlığı üzerinde etkisinin farklı olup olmadığını %5 önem düzeyinde test ediniz.

BUDAMA YÖNTEMİ GRUPLARI			
A	B	C	D
790	820	770	760
810	815	785	770
800	810	765	750
815	820	780	770
805	815	800	760
825	805	770	765

R Veri Girişi:

```
# Veri Girişi-2:
a_budama <- c(790,810,800,815,805,825)
b_budama <- c(820,815,810,820,815,805)
c_budama <- c(770,785,765,780,800,770)
d_budama <- c(760,770,750,770,760,765)
budama_yöntemleri <- factor(rep(c("A","B","C","D"),each=6))
ölçüm <- factor(rep(c("1","2","3","4","5","6"),times=4))
ağırlık <- c(a_budama,b_budama,c_budama,d_budama)
veri2 <- data.frame(ağırlık,budama_yöntemleri,ölçüm)
```

Her Grup İçin Normallik Varsayımın Test Edilmesi:

Her gruba ait verilerin sayısı 50'den az olduğu için normallik kontrolü yapılırken Shapiro-Wilk testi uygulanır.

H_0 : A budama yöntemi grubuna ait meyve ağırlıklarının dağılımı ile normal dağılım arasında fark yoktur.

H_1 : A budama yöntemi grubuna ait meyve ağırlıklarının dağılımı ile normal dağılım arasında fark vardır.

Python Sonuç:

```
In [1]: import numpy as np
        from scipy.stats import shapiro
        import statistics as st

In [2]: A=np.array([790,810,800,815,805,825])
        shapiro(A)

Out[2]: ShapiroResult(statistic=0.9999256134033203, pvalue=0.9999999403953552)
```

R Sonuç:

```
> shapiro.test(a_budama)

      Shapiro-Wilk normality test

data:  a_budama
W = 0.99993, p-value = 1
```

Testin sonucuna göre bulunan p değeri önem düzeyinden büyük olduğu için ($p \approx 0.99 > \alpha = 0.05$) H_0 hipotezi reddedilemez.

YORUM: Dört farklı budama yönteminden biri olan A budama yöntemi ile budanarak yetiştirilmiş kavun ağırlıkları %5 önem düzeyinde normallik varsayımını sağladığını söyleyebiliriz.

H_0 : B budama yöntemi grubuna ait meyve ağırlıklarının dağılımı ile normal dağılım arasında fark yoktur.

H_1 : B budama yöntemi grubuna ait meyve ağırlıklarının dağılımı ile normal dağılım arasında fark vardır.

Python Sonuç:

```
In [3]: B=np.array([820,815,810,820,815,805])
        shapiro(B)

Out[3]: ShapiroResult(statistic=0.9075791835784912, pvalue=0.42066285014152527)
```

R Sonuç:

```
> shapiro.test(b_budama)

      Shapiro-Wilk normality test

data:  b_budama
W = 0.90758, p-value = 0.4207
```

Testin sonucuna göre bulunan p değeri önem düzeyinden büyük olduğu için ($p \approx 0.42 > \alpha = 0.05$) H_0 hipotezi reddedilemez.

YORUM: Dört farklı budama yönteminden biri olan B budama yöntemi ile budanarak yetiştirilmiş kavun ağırlıkları %5 önem düzeyinde normallik varsayımını sağladığını söyleyebiliriz.

H_0 : C budama yöntemi grubuna ait meyve ağırlıklarının dağılımı ile normal dağılım arasında fark yoktur.

H_1 : C budama yöntemi grubuna ait meyve ağırlıklarının dağılımı ile normal dağılım arasında fark vardır.

Python Sonuç:

```
In [4]: ▶ C=np.array([770,785,765,780,800,770])
        shapiro(C)

Out[4]: ShapiroResult(statistic=0.9139266610145569, pvalue=0.46272867918014526)
```

R Sonuç:

```
> shapiro.test(c_budama)

      Shapiro-Wilk normality test

data:  c_budama
W = 0.91393, p-value = 0.4627
```

Testin sonucuna göre bulunan p değeri önem düzeyinden büyük olduğu için ($p \approx 0.46 > \alpha = 0.05$) H_0 hipotezi reddedilemez.

YORUM: Dört farklı budama yönteminden biri olan C budama yöntemi ile budanarak yetiştirilmiş kavun ağırlıkları %5 önem düzeyinde normallik varsayımını sağladığını söyleyebiliriz.

H_0 : D budama yöntemi grubuna ait meyve ağırlıklarının dağılımı ile normal dağılım arasında fark yoktur.

H_1 : D budama yöntemi grubuna ait meyve ağırlıklarının dağılımı ile normal dağılım arasında fark vardır.

Python Sonuç:

```
In [5]: ▶ D=np.array([760,770,750,770,770,767])
        shapiro(D)

Out[5]: ShapiroResult(statistic=0.7749823927879333, pvalue=0.03459442779421806)
```

R Sonuç:

```
> shapiro.test(d_budama)

      Shapiro-Wilk normality test

data:  d_budama
W = 0.90248, p-value = 0.3888
```

Testin sonucuna göre bulunan p değeri önem düzeyinden büyük olduğu için ($p \approx 0.034 < \alpha = 0.05$) H_0 hipotezi reddedilir.

YORUM: Dört farklı budama yönteminden biri olan D budama yöntemi ile budanarak yetiştirilmiş kavun ağırlıkları %5 önem düzeyinde normallik varsayımını sağlamadığını söyleyebiliriz. Bu yüzden farklılığın olup olmadığını hesaplamak için parametrik testler kullanılır.

Her Grup İçin Özet İstatistiklerin Hesaplanması;

Python Sonuç;

```
In [29]: df_A_budama=pd.DataFrame(A_budama)
```

```
In [30]: df_A_budama.rename(columns={0:'puanlar'}, inplace=True)  
df_A_budama
```

Out[30]:

	puanlar
0	790
1	810
2	800
3	815
4	805
5	825

```
In [31]: df_A_budama.describe()
```

Out[31]:

	puanlar
count	6.000000
mean	807.500000
std	12.144958
min	790.000000
25%	801.250000
50%	807.500000
75%	813.750000
max	825.000000

A türü budama yönteminin ortama kavun ağırlığı 807.5'dir. Ağırlıkların ortanca değeri 807.5 gramdır ve standart sapması 12.14 olduğu için verilerin ortalamaya uzak yerlerde dağıldıklarını söyleyebiliriz.

```
In [32]: df_B_budama=pd.DataFrame(B_budama)
```

```
In [33]: df_B_budama.rename(columns={0:'puanlar'}, inplace=True)  
df_B_budama
```

Out[33]:

	puanlar
0	820
1	815
2	810
3	820
4	815
5	805

```
In [34]: df_B_budama.describe()
```

Out[34]:

	puanlar
count	6.000000
mean	814.166667
std	5.845228
min	805.000000
25%	811.250000
50%	815.000000
75%	818.750000
max	820.000000

B türü budama yönteminin ortama kavun ağırlığı 814.16'dır. Ağırlıkların ortanca değeri 815 gramdır ve standart sapması 5.845 olduğu için verilerin ortalamaya yakın yerlerde dağıldıklarını söyleyebiliriz.

```
In [35]: df_C_budama=pd.DataFrame(C_budama)
```

```
In [36]: df_C_budama.rename(columns={0:'puanlar'}, inplace=True)  
df_C_budama
```

```
Out[36]:
```

	puanlar
0	770
1	785
2	785
3	780
4	800
5	770

```
In [37]: df_C_budama.describe()
```

```
Out[37]:
```

	puanlar
count	6.000000
mean	778.333333
std	12.909944
min	765.000000
25%	770.000000
50%	775.000000
75%	783.750000
max	800.000000

```
In [38]: df_D_budama=pd.DataFrame(D_budama)
```

```
In [39]: df_D_budama.rename(columns={0:'puanlar'}, inplace=True)  
df_D_budama
```

```
Out[39]:
```

	puanlar
0	760
1	770
2	760
3	770
4	770
5	767

```
In [40]: df_D_budama.describe()
```

```
Out[40]:
```

	puanlar
count	6.000000
mean	764.500000
std	8.093207
min	750.000000
25%	761.750000
50%	768.500000
75%	770.000000
max	770.000000

C türü budama yönteminin ortama kavun ağırlığı 778.3'dür. Ağırlıkların ortanca değeri 775 gramdır ve standart sapması 12.90 olduğu için verilerin ortalamaya uzak yerlerde dağıldıklarını söyleyebiliriz.

D türü budama yönteminin ortama kavun ağırlığı 764.5'dir. Ağırlıkların ortanca değeri 768.5 gramdır ve standart sapması 8.09 olduğu için verilerin ortalamaya uzak yerlerde dağıldıklarını söyleyebiliriz.

R Sonuçı:

```
> summary(a_budama)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
790.0  801.2   807.5   807.5  813.8   825.0
> summary(b_budama)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
805.0  811.2   815.0   814.2  818.8   820.0
> summary(c_budama)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
765.0  770.0   775.0   778.3  783.8   800.0
> summary(d_budama)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
750.0  760.0   762.5   762.5  768.8   770.0
```

- 3) Bir gıda üreticisinin piyasaya arz etmek üzere 4 farklı formülasyonla işlenmiş salam çeşidi sekiz uzman tarafından 1-10 arasında lezzetlerine göre puanlanmıştır. Bu puanlanan salam çeşitleri arasından A grubu aşağıdaki gibi seçilmiştir. Seçilen salam çeşidinin lezzet puanlarının medyanının 7'den büyük olduğu iddia edilmektedir. %5 önem düzeyinde iddianın doğruluğunu İşaret Testi ve Wilcoxon İşaret Sıra Sayıları Testi ile test ediniz.

SALAM ÇEŞİDİ	PUANLAR
A	1 8 7 7 8 6 7 6

$$H_0: \vartheta (\text{medyan}) = 7$$

$$H_1: \vartheta (\text{medyan}) > 7$$

İşaret Testinin Uygulanması:

Python Sonuçı:

```
In [1]: import numpy as np
        from scipy.stats import binomtest

In [2]: A=np.array([1,8,7,7,8,6,7,6])

In [3]: median_=7

In [4]: nPlus = len(A[A > median_])

In [5]: nNeg = len(A[A < median_])

In [6]: n = nPlus + nNeg

In [7]: binomtest(nPlus, n, p=0.5, alternative='greater')

Out[7]: BinomTestResult(k=2, n=5, alternative='greater', proportion_estimate=0.4, p
value=0.8125)
```

R Sonuç:

```
> library(DescTools)
> SignTest(x=a_salam,mu=7,alternative="greater")

      One-sample Sign-Test

data:  a_salam
S = 2, number of differences = 5, p-value = 0.8125
alternative hypothesis: true median is greater than 7
96.5 percent confidence interval:
 6 Inf
sample estimates:
median of the differences
              7
```

İşaret testinin sonucuna göre bulunan k değeri verilerden medyanı çıkardığımızda çıkan pozitif değerlerin sayısını ifade etmektedir. Bulunan p olasılık değeri ise önem düzeyinden büyük olduğu için ($p \approx 0.81 > \alpha = 0.05$) H_0 hipotezi reddedilemez.

YORUM: Seçilen A grubu salam çeşidinin lezzet puanlarına ilişkin ortanca puanının 7 olduğunu %5 önem düzeyinde söyleyebiliriz.

Wilcoxon İşaret Sıra Sayıları Testinin Uygulanması:

Python Sonuç:

```
In [1]: import numpy as np
        from scipy.stats import wilcoxon

In [2]: A=np.array([1,8,7,7,8,6,7,6])

In [5]: median_=7

In [6]: wilcoxon(A - median_, alternative = 'greater')

Out[6]: WilcoxonResult(statistic=5.0, pvalue=0.7602499389065233)
```

R Sonuç:

```
> library(stats)
> wilcox.test(a_salam,mu=7,alternative="greater")

      wilcoxon signed rank test with continuity correction

data:  a_salam
V = 5, p-value = 0.8019
alternative hypothesis: true location is greater than 7
```

Wilcoxon işaret sıra sayıları testinin sonucuna göre bulunan p olasılık değeri önem düzeyinden büyük olduğu için ($p \approx 0.76 > \alpha = 0.05$) H_0 hipotezi reddedilemez.

YORUM: Seçilen A grubu salam çeşidinin lezzet puanlarına ilişkin ortanca puanının 7 olduğunu %5 önem düzeyinde söyleyebiliriz.

- 4) Bir gıda üreticisinin piyasaya arz etmek üzere 4 farklı formülasyonla işlenmiş salam çeşidi sekiz uzman tarafından 1-10 arasında lezzetlerine göre puanlanmıştır. Bu puanlanan salam çeşitleri arasından B ve D grupları aşağıdaki gibi seçilmiştir. Seçilen salam çeşitlerinin puanlar bakımından farklı oldukları iddiasını %5 önem düzeyinde bağımsız iki örneklem testi olan *Mann-Whitney Testi* ile araştırınız.

SALAM ÇEŞİTLERİ	PUANLAR
B	6 7 5 6 7 7 8 5
D	7 8 8 6 8 9 7 5

$$H_0: \vartheta_B = \vartheta_D$$

$$H_1: \vartheta_B \neq \vartheta_D$$

Mann-Whitney Testinin Uygulanması;

Python Sonuç;

```
In [1]: import numpy as np
        from scipy.stats import mannwhitneyu

In [2]: B=np.array([6,7,5,6,7,7,8,5])

In [3]: D=np.array([7,8,8,6,8,9,7,5])

In [4]: mannwhitneyu(B,D,alternative='two-sided')
Out[4]: MannwhitneyuResult(statistic=18.5, pvalue=0.15999227886797784)
```

R Sonuç;

```
> wilcox.test(b_salam,d_salam,alternative="two.sided",conf.level=0.95,paired=FALSE)

    Wilcoxon rank sum test with continuity correction

data:  b_salam and d_salam
W = 18.5, p-value = 0.16
alternative hypothesis: true location shift is not equal to 0
```

Mann-Whitney testinin sonucuna göre bulunan p olasılık değeri önem düzeyinden büyük olduğu için ($p \approx 0.16 > \alpha = 0.05$) H_0 hipotezi reddedilemez.

YORUM: Seçilen B ve D grubu salam çeşitlerinin lezzet puanları bakımından %5 önem düzeyinde birbirlerinden farklı olmadıklarını söyleyebiliriz.

- 5) Dört farklı budama yöntemi ile budanarak malçlı alçak tünelde yetiştirilmiş bir kavun çeşidinin hasat sonunda meyve ağırlıkları gram cinsinden ölçülmüştür. Bu budama yöntemleri arasında A ve C grupları aşağıdaki gibi seçilmiştir. Seçilen budama yöntemlerinden A grubunun C grubundan daha az etkili olduğu iddiasını %5 önem düzeyinde Wilcoxon İşaretli Sıra Sayıları Testi ile test ediniz.

BUDAMA YÖNTEMLERİ	KAVUN AĞIRLIKLARI
A	790 810 800 815 805 825
C	770 785 765 780 800 770

$$H_0: D_i \text{ (kavun ağırlıkları arasındaki fark)} = 0$$

$$H_1: D_i \text{ (kavun ağırlıkları arasındaki fark)} < 0$$

Wilcoxon İşaretli Sıra Sayıları Testinin Uygulanması:

Python Sonuç:

```
In [1]: import numpy as np
        from scipy.stats import ranksums
```

```
In [2]: A=np.array([790,810,800,815,805,825])
```

```
In [3]: C=np.array([770,785,765,780,800,770])
```

```
In [4]: ranksums(A,C,alternative='less')
```

```
Out[4]: RanksumsResult(statistic=2.6421145377839377, pvalue=0.9958804905871377)
```

R Sonuç:

```
> wilcox.test(a_budama,c_budama,alternative="less",conf.level=0.95,paired=TRUE)

wilcoxon signed rank test with continuity correction

data: a_budama and c_budama
V = 21, p-value = 0.9896
alternative hypothesis: true location shift is less than 0
```

Wilcoxon işaretli sıra sayıları testinin sonucuna göre bulunan p olasılık değeri önem düzeyinden büyük olduğu için ($p \approx 0.99 > \alpha = 0.05$) H_0 hipotezi reddedilemez.

YORUM: Seçilen A ve C grubu budama yöntemlerine ait kavun ağırlıklarının arasında %5 önem düzeyinde farklılık olmadığını söyleyebiliriz.

- 6) Bir gıda üreticisinin piyasaya arz etmek üzere 4 farklı formülasyonla işlenmiş salam çeşidi sekiz uzman tarafından 1-10 arasında puanlanmıştır. Puan değerleri aşağıdaki tabloda verilmiştir. Salam çeşitlerinin lezzet bakımından farklı olup olmadığını %5 önem düzeyinde Kruskal-Wallis Testi ile test ediniz. Gruplar arasında fark varsa Post-Hoc testlerinden birini uygulayınız.

	SALAM ÇEŞİDİ			
UZMAN	A	B	C	D
1	1	6	8	7
2	8	7	9	8
3	7	5	9	8
4	7	6	10	6
5	8	7	9	8
6	6	7	9	9
7	7	8	8	7
8	6	5	7	5

H_0 : Uzmanlar tarafından puanlanan 4 grup salam çeşidinin puanları arasında fark yoktur.

H_1 : Uzmanlar tarafından puanlanan 4 grup salam çeşidinden en az biri farklıdır.

Kruskal-Wallis Testinin Uygulanması;

Python Sonuç;

```
In [1]: import numpy as np
        from scipy.stats import kruskal

In [2]: A=np.array([1,8,7,7,8,6,7,6])

In [3]: B=np.array([6,7,5,6,7,7,8,5])

In [4]: C=np.array([8,9,9,10,9,9,8,7])

In [5]: D=np.array([7,8,8,6,8,9,7,5])

In [6]: kruskal(A,B,C,D)

Out[6]: KruskalResult(statistic=13.076636904761914, pvalue=0.004473770995030885)
```

R Sonuç;

```
> kruskal.test(puan~salam_cesidi,data=veri1)

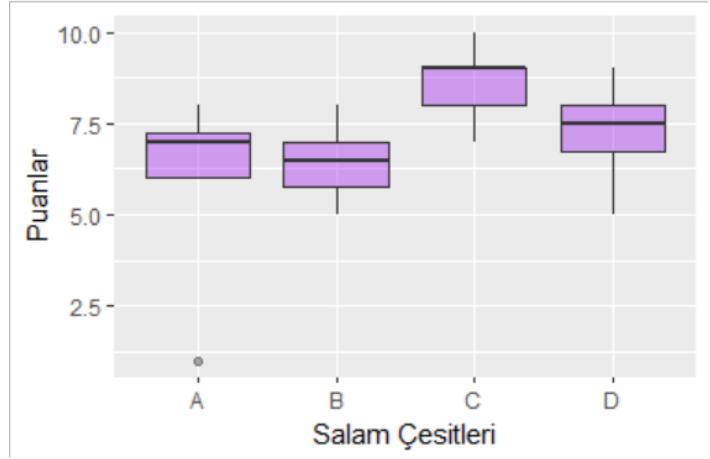
Kruskal-wallis rank sum test

data:  puan by salam_cesidi
Kruskal-Wallis chi-squared = 13.077, df = 3, p-value = 0.004474
```

Kruskal-Wallis testinin sonucuna göre bulunan p olasılık değeri önem düzeyinden küçük olduğu için ($p \approx 0.004 < \alpha = 0.05$) H_0 hipotezi reddedilir.


```
# Bboxplot Grafiği:
```

```
library(ggplot2)
ggplot(veri1, aes(x=salam_cesidi, y=puan)) +
  geom_boxplot(fill="purple", alpha=0.4) +
  xlab("Salam Çeşitleri") + ylab("Puanlar")
```



YORUM: 4 farklı formülasyonla işlenmiş salam çeşitlerine 8 uzman tarafından verilen puanlar arasında %5 önem düzeyinde en az bir salam çeşidinin farklılık gösterdiğini söyleyebiliriz.

Hangi salam çeşidinin farklılık gösterdiğini bulmak için Post-Hoc testleri uygulanır.

Post-Hoc Testinin Uygulanması:

H_0 : Grupların ortalamaları aynıdır.

H_1 : En az bir grup ortalaması diğerlerinden farklıdır.

Python Sonuç:

```
In [1]: pip install scikit-posthocs
```

```
In [2]: import numpy as np
import scikit_posthocs as sp
```

```
In [3]: A=np.array([1,8,7,7,8,6,7,6])
```

```
In [4]: B=np.array([6,7,5,6,7,7,8,5])
```

```
In [5]: C=np.array([8,9,9,10,9,9,8,7])
```

```
In [6]: D=np.array([7,8,8,6,8,9,7,5])
```

```
In [7]: veri=[A,B,C,D]
```

```
In [8]: sp.posthoc_dunn(veri, p_adjust='bonferroni')
```

Out[8]:

	1	2	3	4
1	1.000000	1.000	0.025091	1.000000
2	1.000000	1.000	0.005000	1.000000
3	0.025091	0.005	1.000000	0.287843
4	1.000000	1.000	0.287843	1.000000

R Sonuç:

```
> library(PMCMRplus)
> kwAllPairsNemenyiTest(x=puan,salam_cesidi, method="Chisq")

Pairwise comparisons using Tukey-Kramer-Nemenyi all-pairs test with Tukey-Dist approximation

data: puan and salam_cesidi

  A      B      C
B 0.966 -      -
C 0.026 0.006 -
D 0.822 0.542 0.215

P value adjustment method: single-step
alternative hypothesis: two.sided
```

Post-Hoc testinin sonucuna göre;

1. ve 3. Grup arasındaki fark için düzeltilmiş p değeri önem düzeyinden küçük olduğu için ($p \approx 0.025 < \alpha = 0.05$) A ve C grubu salam çeşitleri arasında fark vardır diyebiliriz.
2. ve 3. Grup arasındaki fark için düzeltilmiş p değeri önem düzeyinden küçük olduğu için ($p = 0.005 < \alpha = 0.05$) B ve C grubu salam çeşitleri arasında fark vardır diyebiliriz.
3. ve 4. Grup arasındaki fark için düzeltilmiş p değeri önem düzeyinden küçük olduğu için ($p \approx 0.29 > \alpha = 0.05$) C ve D grubu salam çeşitleri arasında fark yoktur diyebiliriz.

```
In [8]: p_values=sp.posthoc_dunn(veri, p_adjust='bonferroni')
In [9]: p_values > 0.05

Out[9]:
```

	1	2	3	4
1	True	True	False	True
2	True	True	False	True
3	False	False	True	True
4	True	True	True	True

YORUM: 4 farklı formülasyonla işlenmiş ve uzmanlar tarafından puanlanmış salam çeşitlerinden C grubu salam çeşidinin, D grubu salam çeşidi hariç diğer salam çeşitlerinden farklı olduğunu ve C grubu hariç diğer salam çeşitlerinin de birbirlerinden farklı olmadığını yani gruplar arası farklılığa neden olan salam çeşidinin C grubu olduğunu %5 önem düzeyinde söyleyebiliriz.

- 7) Dört farklı budama yöntemi ile budanarak malçlı alçak tünelde yetiştirilmiş bir kavun çeşidinin hasat sonunda meyve ağırlıkları gram cinsinden ölçülüp aşağıdaki tabloda verilmiştir. Budama yöntemlerinin meyve ağırlığı üzerinde etkisinin farklı olup olmadığını %5 önem düzeyinde Friedman Testi ile test ediniz. Gruplar arasında fark varsa Post-Hoc testlerinden birini uygulayınız.

BUDAMA YÖNTEMİ GRUPLARI			
A	B	C	D
790	820	770	760
810	815	785	770
800	810	765	750
815	820	780	770
805	815	800	760
825	805	770	765

H_0 : 4 farklı budama yönteminden elde edilen kavun ağırlıkları arasında fark yoktur.

H_1 : 4 farklı budama yönteminden en az biri farklıdır.

Friedman Testinin Uygulanması:

Python Sonuç:

```
In [1]: import numpy as np
        from scipy import stats as st

In [2]: A=np.array([790,810,800,815,805,825])

In [3]: B=np.array([820,815,810,820,815,805])

In [4]: C=np.array([770,785,765,780,800,770])

In [5]: D=np.array([760,770,750,770,760,765])

In [6]: st.friedmanchisquare(A,B,C,D)
Out[6]: FriedmanchisquareResult(statistic=17.0, pvalue=0.0007067423923370282)
```

R Sonuç:

```
> friedman.test(agirlik~budama_yontemleri|olcüm,data=veri2)

Friedman rank sum test

data:  agirlik and budama_yontemleri and olcüm
Friedman chi-squared = 17, df = 3, p-value = 0.0007067
```

Friedman testinin sonucuna göre bulunan p olasılık değeri önem düzeyinden küçük olduğu için ($p \approx 0.0007 < \alpha = 0.05$) H_0 hipotezi reddedilir.

YORUM: 4 farklı budama yönteminden elde edilen kavun ağırlıkları arasında %5 önem düzeyinde en az bir budama çeşidinin farklılık gösterdiğini %5 önem düzeyinde söyleyebiliriz.

Hangi budama yönteminin farklılık gösterdiğini bulmak için Post-Hoc testleri uygulanır.

Post-Hoc Testinin Uygulanması;

H_0 : Grupların ortalamaları aynıdır.

H_1 : En az bir grup ortalaması diğerlerinden farklıdır.

Python Sonuç;

```
In [1]: ► pip install scikit-posthocs

In [2]: ► import numpy as np
import scikit_posthocs as sp

In [3]: ► A=np.array([790,810,800,815,805,825])

In [4]: ► B=np.array([820,815,810,820,815,805])

In [5]: ► C=np.array([770,785,765,780,800,770])

In [6]: ► D=np.array([760,770,750,770,760,765])

In [7]: ► veri=np.array([A,B,C,D])

In [8]: ► sp.posthoc_nemenyi_friedman(veri.T)

Out[8]:
```

	0	1	2	3
0	1.000000	0.783106	0.399914	0.019135
1	0.783106	1.000000	0.066476	0.001000
2	0.399914	0.066476	1.000000	0.532081
3	0.019135	0.001000	0.532081	1.000000

R Sonuç;

```
> frdAllPairsNemenyiTest(y=agirlik,groups=budama_yonemleri,blocks=olcum)

Pairwise comparisons using Nemenyi-Wilcoxon-Wilcox all-pairs test for a two-way balanced complete block design
data: y, groups and blocks

  A      B      C
B 0.80776 -      -
C 0.39863 0.06639 -
D 0.01913 0.00083 0.53629

P value adjustment method: single-step
```

Post-Hoc testinin sonucuna göre;

0. ve 1. Grup arasındaki fark için p değeri önem düzeyinden büyük olduğu için ($p \approx 0.78 > \alpha = 0.05$) A ve B grubu budama yöntemleri arasında fark yoktur diyebiliriz.

0. ve 2. Grup arasındaki fark için p değeri önem düzeyinden büyük olduğu için ($p \approx 0.40 > \alpha = 0.05$) A ve C grubu budama yöntemleri arasında fark yoktur diyebiliriz.

0. ve 3. Grup arasındaki fark için p değeri önem düzeyinden küçük olduğu için ($p \approx 0.019 < \alpha = 0.05$) A ve D grubu budama yöntemleri arasında fark vardır diyebiliriz.

1. ve 2. Grup arasındaki fark için p değeri önem düzeyinden büyük olduğu için ($p \approx 0.066 > \alpha = 0.05$) B ve C grubu budama yöntemleri arasında fark yoktur diyebiliriz.

1. ve 3. Grup arasındaki fark için p değeri önem düzeyinden küçük olduğu için ($p \approx 0.001 < \alpha = 0.05$) B ve D grubu budama yöntemleri arasında fark vardır diyebiliriz.

2. ve 3. Grup arasındaki fark için p değeri önem düzeyinden büyük olduğu için ($p \approx 0.53 > \alpha = 0.05$) C ve D grubu budama yöntemleri arasında fark yoktur diyebiliriz.

```
In [8]: p_values=sp.posthoc_nemenyi_friedman(veri.T)
```

```
In [9]: p_values > 0.05
```

Out[9]:

	0	1	2	3
0	True	True	True	False
1	True	True	True	False
2	True	True	True	True
3	False	False	True	True

YORUM: 4 farklı budama yöntemi ile elde edilen kavun ağırlıklarından D grubu budama yönteminin, C grubu budama yöntemi hariç diğer budama yöntemlerinden farklı olduğunu ve D grubu hariç diğer budama yöntemlerinin de birbirlerinden farklı olmadığını, yani gruplar arası farklılığa neden olan budama yönteminin D grubu olduğunu %5 önem düzeyinde söyleyebiliriz.

KAYNAKÇA

"Zeynel CEBECİ (2019). R İle Parametrik Olmayan İstatistik Analiz. Abaküs Yayıncılık."

1. Veri (bağımsız veri) : Sayfa 255
2. Veri (bağımlı veri) : Sayfa 213

PYTHON KODLARI

1

Her bir salam çeşidi için normallik varsayımının kontrolü:

```
import numpy as np
```

```
from scipy.stats import shapiro
```

```
import statistics as st
```

```
A_salam=np.array([1,8,7,7,8,6,7,6])
```

```
shapiro(A_salam)
```

```
B_salam=np.array([6,7,5,6,7,7,8,5])
```

```
shapiro(B_salam)
```

```
C_salam=np.array([8,9,9,10,9,9,8,7])
```

```
shapiro(C_salam)
```

```
D_salam=np.array([7,8,8,6,8,9,7,5])
```

```
shapiro(D_salam)
```

Her bir salam çeşidi için özet istatistikler:

```
import pandas as pd
```

```
df_A_salam=pd.DataFrame(A_salam)
```

```
df_A_salam.rename(columns={0:'puanlar'}, inplace=True)
```

```
df_A_salam
```

```
df_A_salam.describe()
```

```
df_B_salam=pd.DataFrame(B_salam)
```

```
df_B_salam.rename(columns={0:'puanlar'}, inplace=True)
```

```
df_B_salam
```

```
df_B_salam.describe()
```

```
df_C_salam=pd.DataFrame(C_salam)
```

```
df_C_salam.rename(columns={0:'puanlar'}, inplace=True)
```

```
df_C_salam
```

```

df_C_salam.describe()

df_D_salam=pd.DataFrame(D_salam)

df_D_salam.rename(columns={0:'puanlar'}, inplace=True)

df_D_salam

df_D_salam.describe()

# 2

# Her bir budama yöntemi için normallik varsayımının kontrolü:

A_budama=np.array([790,810,800,815,805,825])

shapiro(A_budama)

B_budama=np.array([820,815,810,820,815,805])

shapiro(B_budama)

C_budama=np.array([770,785,765,780,800,770])

shapiro(C_budama)

D_budama=np.array([760,770,750,770,770,767])

shapiro(D_budama)

# Her bir budama yöntemi için özet istatistikler:

df_A_budama=pd.DataFrame(A_budama)

df_A_budama.rename(columns={0:'puanlar'}, inplace=True)

df_A_budama

df_A_budama.describe()

df_B_budama=pd.DataFrame(B_budama)

df_B_budama.rename(columns={0:'puanlar'}, inplace=True)

df_B_budama

df_B_budama.describe()

df_C_budama=pd.DataFrame(C_budama)

df_C_budama.rename(columns={0:'puanlar'}, inplace=True)

df_C_budama

```



```
df_C_budama.describe()

df_D_budama=pd.DataFrame(D_budama)

df_D_budama.rename(columns={0: 'puanlar'}, inplace=True)

df_D_budama

df_D_budama.describe()
```

3

A salam çeşidi için İşaret testi:

```
from scipy.stats import binomtest

median_hipotez=7

nPlus=len(A_salam[A_salam > median_hipotez])

nNeg=len(A_salam[A_salam < median_hipotez])

n=nPlus+nNeg

binomtest(nPlus, n, p=0.5, alternative='greater')
```

A salam çeşidi için Wilcoxon testi:

```
from scipy.stats import wilcoxon

wilcoxon(A_salam - median_hipotez, alternative='greater')
```

4

B ve D salam çeşitleri için Mann-Whitney testi:

```
from scipy.stats import mannwhitneyu

mannwhitneyu(B_salam, D_salam, alternative='two-sided')
```

5

A ve C budama yöntemi için Wilcoxon testi:

```
from scipy.stats import ranksums

ranksums(A_budama, C_budama, alternative='less')
```

6

Salam çeşitleri için Kruskal-Wallis testi:

```
from scipy.stats import kruskal
```

```
kruskal(A_salam, B_salam, C_salam, D_salam)
```

```
# Salam çeşitleri için Post-Hoc testi:
```

```
pip install scikit-posthocs
```

```
import scikit_posthocs as sp
```

```
veri_salam=np.array([A_salam, B_salam, C_salam, D_salam])
```

```
sp.posthoc_dunn(veri_salam, p_adjust='bonferroni')
```

```
# 7
```

```
# Budama yöntemleri için Friedman testi:
```

```
from scipy import stats as st
```

```
st.friedmanchisquare(A_budama, B_budama, C_budama, D_budama)
```

```
# Budama yöntemleri için Post-Hoc testi:
```

```
veri_budama=np.array([A_budama, B_budama, C_budama, D_budama])
```

```
sp.posthoc_nemenyi_friedman(veri_budama.T)
```

R STUDIO KODLARI

Veri Girişi-1:

```
a_salam <- c(1,8,7,7,8,6,7,6)
```

```
b_salam <- c(6,7,5,6,7,7,8,5)
```

```
c_salam <- c(8,9,9,10,9,9,8,7)
```

```
d_salam <- c(7,8,8,6,8,9,7,5)
```

```
salam_cesidi <- factor(rep(c("A","B","C","D"),each=8))
```

```
uzman <- factor(rep(c("1","2","3","4","5","6","7","8"),times=4))
```

```
puan <- c(a_salam,b_salam,c_salam,d_salam)
```

```
veri1 <- data.frame(puan,salam_cesidi,uzman)
```

Normallik Testi-1:

```
shapiro.test(a_salam)
```

```
shapiro.test(b_salam)
```

```
shapiro.test(c_salam)
```

```
shapiro.test(d_salam)
```

Özet İstatistikler-1:

```
summary(a_salam)
```

```
summary(b_salam)
```

```
summary(c_salam)
```

```
summary(d_salam)
```

Veri Girişi-2:

```
a_budama <- c(790,810,800,815,805,825)
```

```
b_budama <- c(820,815,810,820,815,805)
```

```
c_budama <- c(770,785,765,780,800,770)
```

```
d_budama <- c(760,770,750,770,760,765)
```

```
budama_yöntemleri <- factor(rep(c("A","B","C","D"),each=6))
```

```
ölçüm <- factor(rep(c("1","2","3","4","5","6"),times=4))
```

```

agirlik <- c(a_budama,b_budama,c_budama,d_budama)

veri2 <- data.frame(agirlik,budama_yöntemleri,ölçüm)

# Normallik Testi-2:

shapiro.test(a_budama)

shapiro.test(b_budama)

shapiro.test(c_budama)

shapiro.test(d_budama)

# Özet İstatistikler-2:

summary(a_budama)

summary(b_budama)

summary(c_budama)

summary(d_budama)

# İşaret Testi-1:

library(DescTools)

SignTest(x=a_salam,mu=7,alternative="greater")

# Wilcoxon Testi-1:

library(stats)

wilcox.test(a_salam,mu=7,alternative="greater")

# Mann-Whitney Testi-1:

wilcox.test(b_salam,d_salam,alternative="two.sided",conf.level=0.95,paired=FALSE)

# İki Örneklem Wilcoxon Testi-2:

wilcox.test(a_budama,c_budama,alternative="less",conf.level=0.95,paired=TRUE)

# Kruskal-Wallis Testi-1:

library(stats)

kruskal.test(puan~salam_cesidi,data=veri1)

# Boxplot Grafiği:

library(ggplot2)

```

```
ggplot(veri1, aes(x=salam_cesidi, y=puan)) +  
  geom_boxplot(fill="purple", alpha=0.4) +  
  xlab("Salam Çeşitleri") + ylab("Puanlar")  
  
# Post-Hoc Testi-1:  
  
library(PMCMRplus)  
  
kwAllPairsNemenyiTest(x=puan,salam_cesidi, method="Chisq")  
  
# Friedman Testi-2:  
  
friedman.test(agirlik~budama_yöntemleri|ölçüm,data=veri2)  
  
# Post-Hoc Testi-2:  
  
frdAllPairsNemenyiTest(y=agirlik,groups=budama_yöntemleri,blocks=ölçüm)
```