**İST 367- İSTATİSTİKSEL YÖNTEMLER ÖDEVİ**

**Soru 1:** Değişik türde yeşil fasulyeler ile ilgili bir deney yapılmak istenmektedir. Bunun için her biri farklı gen ailesine sahip 4 çeşit yeşil fasulye 3 farklı gübre çeşidi kullanılarak yetiştiriliyor. Sonuç olarak birlikte kullanıldığında en fazla verimi veren gübre ve fasulye türü belirlenmek isteniyor. Buna göre gübre ve fasulye türlerine ait aşağıdaki verim ölçümleri elde edilmiştir.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Gübre** | | | |
| **Aile** |  | **1** | **2** | **3** |
| **a** | 4,7 | 9,4 | 6,3 |
| **b** | 3,5 | 7,6 | 5,1 |
| **c** | 0,1 | 5,3 | 1,8 |
| **d** | 1,6 | 6,2 | 3,6 |

**Varyans Analizi**

**Bağımlı değişken:** Verim

**Faktörler:** Fasulye çeşitleri (sabit)

**Blok:** Gübre çeşitleri (sabit)

**Model Denklemi:**

yij= +i j +ij

yij: i. satır, j. sütuna karşılık gelen gözlem değeri

: Sabit Etki

i : i. blok etkisi

j :j. Deneme etkisi

ij : i. satır, j. sütuna karşıklık gelen hata

**Varsayımlar**

**1- Normallik Varsayımı**

H0: Fasulye çeşitlerine ait verim değerleri ile normal dağılım arasında fark yoktur.

Hs: Fasulye çeşitlerine ait verim değerleri ile normal dağılım arasında fark vardır.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Tests of Normality** | | | | | | | |
|  | fas\_ailesi | Kolmogorov-Smirnova | | | Shapiro-Wilk | | |
|  | Statistic | df | Sig. | Statistic | df | Sig. |
| verim | a | ,250 | 3 | . | ,967 | 3 | ,652 |
| b | ,224 | 3 | . | ,984 | 3 | ,759 |
| c | ,256 | 3 | . | ,962 | 3 | ,623 |
| d | ,201 | 3 | . | ,994 | 3 | ,856 |
| a. Lilliefors Significance Correction | | | | | | | |

**Yorum:** α=0,05<p-value = (0.652, 0.759, 0.623, 0.856) olduğundan H0 hipotezi kabul edilir.

Fasulye çeşitlerine ait verim değerleri ile normal dağılım arasında fark olmadığı %95 güven düzeyinde söylenebilir.

**2- Varyansların Homojenliği Varsayımı**





|  |  |  |  |
| --- | --- | --- | --- |
| **Test of Homogeneity of Variances** | | | |
| verim | | | |
| Levene Statistic | df1 | df2 | Sig. |
| ,092 | 3 | 8 | ,962 |

**Yorum:**

1. α =0,05<Sig. = 0.962 olduğundan H0 hipotezi kabul edilir.
2. Ftablo=Fα=0.05;3;8= 4,0662

Fhesap=0.092<Ftablo=4,0662 olduğundan H0 hipotezi kabul edilir.

Fasulye çeşitleri için 3 farklı denemenin değerlerine ait varyansların homojen olduğu %95 güven düzeyinde söylenebilir. Varsayımlar sağlandığından iki yönlü varyans analizine geçebiliriz.

## Deneme ve blokların her ikisi de sabit seçimlidir.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Tests of Between-Subjects Effects** | | | | | |
| Dependent Variable: verim | | | | | |
| Source | Type III Sum of Squares | df | Mean Square | F | Sig. |
| Corrected Model | 77,085a | 5 | 15,417 | 203,301 | ,000 |
| Intercept | 253,920 | 1 | 253,920 | 3348,396 | ,000 |
| gubre\_turu | 44,205 | 2 | 22,103 | 291,462 | ,000 |
| fas\_ailesi | 32,880 | 3 | 10,960 | 144,527 | ,000 |
| Error | ,455 | 6 | ,076 |  |  |
| Total | 331,460 | 12 |  |  |  |
| Corrected Total | 77,540 | 11 |  |  |  |
| a. R Squared = ,994 (Adjusted R Squared = ,989) | | | | | |

* Fasulye çeşitleri (denemeler) sabit seçimli olduğu için hipotezler aşağıdaki gibidir;





**Yorum:** Tabloya bakıldığında α =0,05>Sig. = 0.00 olduğundan H0 hipotezi reddedilir. Yani, verim değerleri bakımından fasulye çeşitleri arasında fark vardır.

* Gübre çeşitleri (bloklar) sabit seçimli olduğu için hipotezler aşağıdaki gibidir;



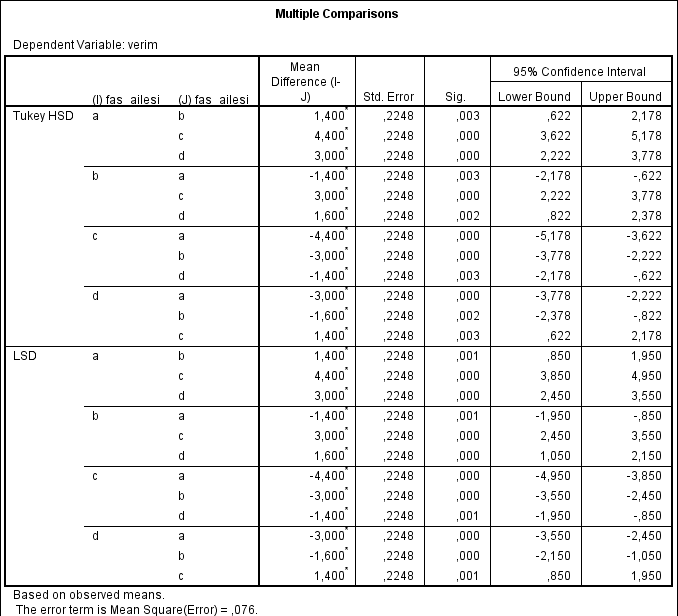


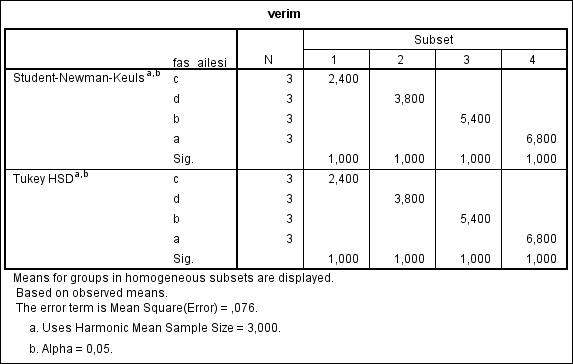
**Yorum:** Tabloya bakıldığında α =0,05>Sig. = 0.00 olduğundan H0 hipotezi reddedilir. Yani, verim değerleri bakımından gübre çeşitleri arasında fark vardır.

Sonuç olarak; verim değerleri bakımından hem fasulye çeşitleri arasında hem de gübre çeşitleri arasında farklılık bulunduğundan çoklu karşılaştırma testi yapmalıyız.

**Fasulye Çeşitleri için Çoklu Karşılaştırma**

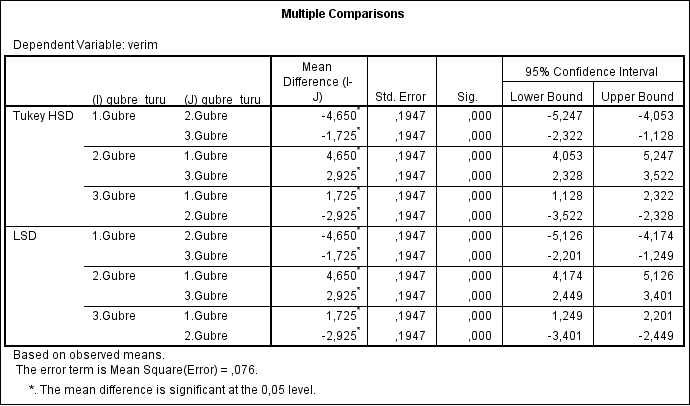
****

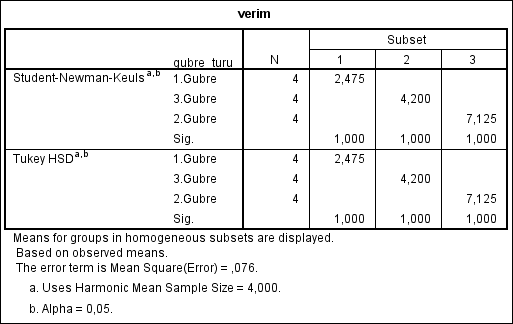
****

Çoklu karşılaştırma testlerinden(LSD, SNK, Tukey) SNK tablosunu yorumlayalım. SNK tablosuna baktığımızda 4 tane alt grup oluştuğunu görülmektedir. Yani fasulye çeşitlerinin verim açısından hepsinin birbirinden farklı olduğu bulunmuştur.

**Gübre Çeşitleri için Çoklu Karşılaştırma**



****



Çoklu karşılaştırma testlerinden(LSD, SNK, Tukey) SNK tablosunu yorumlayalım. SNK tablosuna baktığımızda 3 tane alt grup oluştuğunu görülmektedir. Yani gübre çeşitlerinin verim açısından hepsinin birbirinden farklı olduğu bulunmuştur.

**Soru 2:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Elektrik Gerecinin Yaşam Testi Verileri** | | | | |
| **Fırında Kalma  Süresi  (S) (Dak.)** | **Fırın Isısı (F)** | | | |
| **580** | **600** | **620** | **640** |
| **5** | 217 | 158 | 229 | 223 |
| 188 | 126 | 160 | 201 |
| 162 | 122 | 167 | 182 |
| **10** | 233 | 138 | 186 | 227 |
| 201 | 130 | 170 | 181 |
| 170 | 185 | 181 | 201 |
| **15** | 175 | 152 | 155 | 156 |
| 195 | 147 | 161 | 172 |
| 213 | 180 | 182 | 199 |

Tablo’daki veriler bir elektrik gerecinin ömrüne (Y), fırın ısısı (F) ve fırında kalma süresi (S)’nin etkisini test etmek için yapılan deneye aittir. Burada ürünün ömrü bağımlı, Fırın ısısı ve Fırında kalma süreleri bağımsızdır değişkenlerdir. Ayrıca Fırın ısısı ve Fırında kalma süreleri özel seçimlidir.

**İki Yönlü Varyans Analizi**

**Bağımlı değişken:** Ürünün ömrü

**Bağımsız değişken:** Fırın Isısı ve Fırında kalma süreleri

**Faktörler:** Fırın Isısı (sabit), Fırında Kalma Süresi (sabit)

**Model Denklemi:**

yij= + i + j + ij + ijl

**Varsayımlar**

**1- Normallik Varsayımı**

H01: 1. deneme kombinasyonuna ait verilerin dağılımı ile normal dağılım arasında fark yoktur.

Hs1: 1. deneme kombinasyonuna ait verilerin dağılımı ile normal dağılım arasında fark vardır.

H02: 2. deneme kombinasyonuna ait verilerin dağılımı ile normal dağılım arasında fark yoktur.

Hs2: 2. deneme kombinasyonuna ait verilerin dağılımı ile normal dağılım arasında fark vardır.

H03: 3. deneme kombinasyonuna ait verilerin dağılımı ile normal dağılım arasında fark yoktur.

Hs3: 3. deneme kombinasyonuna ait verilerin dağılımı ile normal dağılım arasında fark vardır.

H04: 4. deneme kombinasyonuna ait verilerin dağılımı ile normal dağılım arasında fark yoktur.

Hs4: 4. deneme kombinasyonuna ait verilerin dağılımı ile normal dağılım arasında fark vardır.

H05: 5. deneme kombinasyonuna ait verilerin dağılımı ile normal dağılım arasında fark yoktur.

Hs5: 5. deneme kombinasyonuna ait verilerin dağılımı ile normal dağılım arasında fark vardır.

H06: 6. deneme kombinasyonuna ait verilerin dağılımı ile normal dağılım arasında fark yoktur.

Hs6: 6. deneme kombinasyonuna ait verilerin dağılımı ile normal dağılım arasında fark vardır.

H07: 7. deneme kombinasyonuna ait verilerin dağılımı ile normal dağılım arasında fark yoktur.

Hs7: 7. deneme kombinasyonuna ait verilerin dağılımı ile normal dağılım arasında fark vardır.

H08: 8. deneme kombinasyonuna ait verilerin dağılımı ile normal dağılım arasında fark yoktur.

Hs8: 8. deneme kombinasyonuna ait verilerin dağılımı ile normal dağılım arasında fark vardır.

H09: 9. deneme kombinasyonuna ait verilerin dağılımı ile normal dağılım arasında fark yoktur.

Hs9: 9. deneme kombinasyonuna ait verilerin dağılımı ile normal dağılım arasında fark vardır.

H010: 10. deneme kombinasyonuna ait verilerin dağılımı ile normal dağılım arasında fark yoktur.

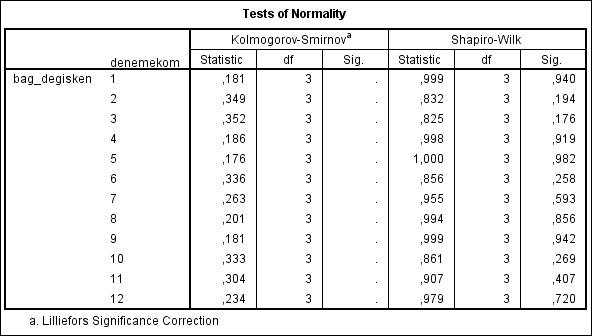
Hs10: 10. deneme kombinasyonuna ait verilerin dağılımı ile normal dağılım arasında fark vardır.

H011: 11. deneme kombinasyonuna ait verilerin dağılımı ile normal dağılım arasında fark yoktur.

Hs11: 11. deneme kombinasyonuna ait verilerin dağılımı ile normal dağılım arasında fark vardır.

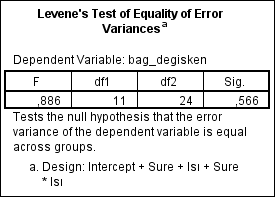
H012: 12. deneme kombinasyonuna ait verilerin dağılımı ile normal dağılım arasında fark yoktur.

Hs12: 12. deneme kombinasyonuna ait verilerin dağılımı ile normal dağılım arasında fark vardır.

****

Tabloya bakıldığında (Sig.> α=0,05) bütün deneme kombinasyonlarına ait verilerin dağılımı ile normal dağılım arasında fark yoktur.

**2- Varyansların Homojenliği Varsayımı**

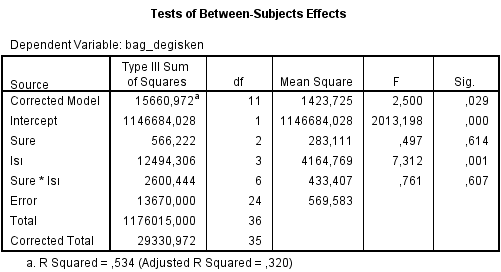


H0: Deneme kombinasyonlarına ait verilerin varyansları homojendir.

HS: Deneme kombinasyonlarına ait verilerin varyansları homojen değildir.

**Yorum:** Sig.=0,566>α=0,05 olduğundan deneme kombinasyonlarına ait veriler için varyanslar homojendir. Varsayımlar sağlandığından iki yönlü çok etkenli varyans analizine geçebiliriz.

**#** Fırın ısısı ve Fırında kalma süresi özel seçimlidir.



H01: Fırında kalma sürelerinin elektrik gerecinin ömrüne etkisi yoktur.

HS1: Fırında kalma sürelerinin elektrik gerecinin ömrüne etkisi vardır.

Sig.=0,614 > α=0,05 olduğundan H01 hipotezi kabul edilir.

**Yorum:** Fırında kalma sürelerinin elektrik gerecinin ömrüne etkisi olmadığı %95 anlamlılık düzeyinde söylenir.

H02: Fırın ısılarının elektrik gerecinin ömrüne etkisi yoktur.

HS2: Fırın ısılarının elektrik gerecinin ömrüne etkisi vardır.

Sig.=0,01 < α=0,05 olduğundan H02 hipotezi reddedilir.

**Yorum:** Fırın ısılarının elektrik gerecinin ömrüne etkisi olduğu %5 anlamlılık düzeyinde söylenir. Fırın ısısı özel seçimli olduğu için **Çoklu Karşılaştırma Testi** yapacağız.

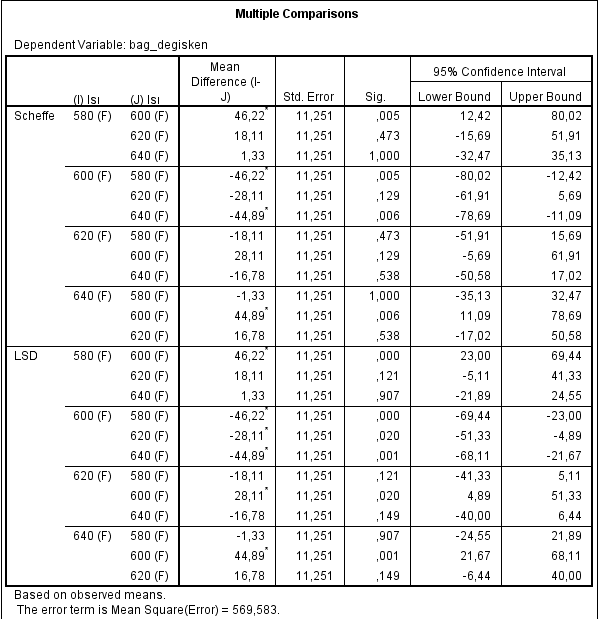
H03: Fırında kalma süresi ve fırın ısısı etkileşiminin elektrik gerecinin ömrüne etkisi yoktur.

HS3: Fırında kalma süresi ve fırın ısısı etkileşiminin elektrik gerecinin ömrüne etkisi vardır.

Sig.=0,607 > α=0,05 olduğundan H03 hipotezi kabul edilir.

**Yorum:** Fırında kalma süresi ve fırın ısısı etkileşiminin elektrik gerecinin ömrüne etkisi olmadığı %95 anlamlılık düzeyinde söylenir.

**Fırın Isısı için Çoklu Karşılaştırma Testleri (Scheffe, LSD, SNK)**

****

**Fırın Isısı için Çoklu Karşılaştırma**

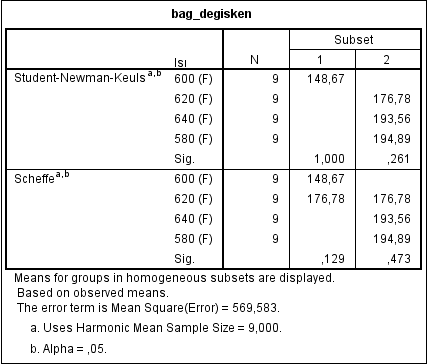


**Yorum:** LSD tablosundan;

580F -600F 🡪 Sig.=0,00 < α=0,05 olduğundan H01 hipotezi reddedilir.

600F - 620F 🡪 Sig.=0,20 < α=0,05 olduğundan H04 hipotezi reddedilir.

600F - 640F 🡪 Sig.=0,01 < α=0,05 olduğundan H05 hipotezi reddedilir.

****

Çoklu karşılaştırma testlerinden SNK ve Scheffe tablosunu yorumlayalım.

* SNK tablosuna baktığımızda 2 tane alt grup oluştuğunu görülmektedir. Alt gruba bakıldığında 1. alt grupta fırın ısısı olarak 600F yer almıştır, 2. Alt grupta ise 620F, 640F, 580F yer almıştır.
* Scheffe tablosuna baktığımızda 2 tane alt grup oluştuğunu görülmektedir. Alt gruba bakıldığında 1. alt grupta fırın ısısı olarak 600F ve 620F yer almıştır, 2. alt grupta ise 640F, 580F yer almıştır. Buradan alt gruplar arasında farklılığın olduğunu ama kendi alt grup içleri arasında fark olmadığı gözlemlenmiştir. **Not:** Scheffe tablosundan 620F değeri iki alt grupta gözlemlenmiştir.

**Sonuç:** Fırın ısısının elektrik gerecinin ömrüne etkisi olduğu %95 güven düzeyinden söylenebilir.

**Not:** Problem SPSS ile çözülmüştür ve SPSS sonuçlarında E(KO) sonuçları gözükmemektedir. Verilen problem özel seçimli olduğu için E(KO) sütununa gerek yoktur, çünkü E(KO) kolonu karışık seçimli problemlerde F testinin paydasında ne olacağını(KOHATA/SD) bulmamıza yardımcı olur.

**Soru 3:** 19,5 metre kare büyüklüğündeki parsellere yapılan buğday ekim sonuçları kg. olarak verilmiştir. Parselleri rasgele seçimli olan karışık seçimli deney düzeninde çözümlemeyi gerçekleştiriniz.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Parsel** | **0** | **Na** | **2Na** | **3Na** |
| **1** | 3,4 | 4,7 | 4,9 | 4,3 |
| **2** | 3,5 | 4,6 | 3,9 | 4,3 |
| **3** | - | 4,8 | 4,6 | 4,8 |
| **4** | 3,4 | 4,4 | 4,7 | 4,3 |

**Bağımlı değişken:** kilogram

**Faktörler:** Deneme (düzey) (sabit)

**Blok:** Parsel (rasgele)

**Model Denklemi:**

yij= +i j +ij

**Kayıp Gözlem Değeri**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Parsel** | **0** | **Na** | **2Na** | **3Na** | **Toplam** |
| **1** | 3,4 | 4,7 | 4,9 | 4,3 | 17,3 |
| **2** | 3,5 | 4,6 | 3,9 | 4,3 | 16,3 |
| **3** | **X31** | 4,8 | 4,6 | 4,8 | **14,2+X31** |
| **4** | 3,4 | 4,4 | 4,7 | 4,3 | 16,8 |
| **Toplam** | **10,3+X31** | 18,5 | 18,1 | 17,7 | **64,6+X31** |

Tablodaki X31 kayıp gözlem değerini tahmin edelim. R’ da bunun için bir fonksiyon yazalım.

**Kod Kısmı**

kgozlem<-function(n,k,T,Ti,Tj)

{

x=((n\*Ti)+(k\*Tj)-T)/((n-1)\*(k-1))

print("Kayip Gözlem Verisi")

print("-------------------")

return(x)

}

kgozlem(4, 4, 64.6, 10.3, 14.2)

**Değişkenlerin Anlamları**

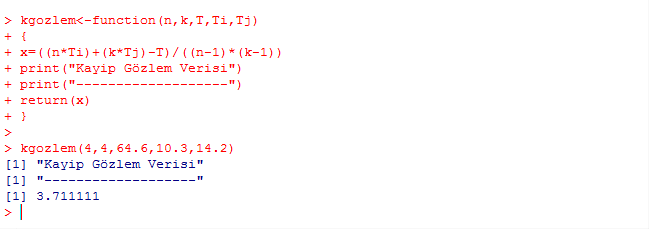
**n:** Satır Sayısı (değeri: 4)

**k:** Sütun Sayısı (değeri: 4)

**T:** Tüm satırların yada tüm sütunların toplamı (değeri: 64,6)

**Ti:** Kayıp gözlem verisinin bulunduğu satırın toplamı (değeri: 10,3)

**Tj:** Kayıp gözlem verisinin bulunduğu sütunun toplamı (değeri: 14,2)



Kayıp gözlem verisi tahmin edilmiş ve 3.7 olarak alınmıştır, artık analize geçebiliriz.

**## R Kodları ##**

**Verinin Dışarıdan R Programına Girilişi**

veri= read.table("C:/IstYontFinal/veri.txt", header=TRUE)

veri

attach(veri)

parsel\_1=veri[veri[,1]=="1",3]

parsel\_2=veri[veri[,1]=="2",3]

parsel\_3=veri[veri[,1]=="3",3]

parsel\_4=veri[veri[,1]=="4",3]

**> parsel\_1**

[1] 3.4 4.7 4.9 4.3

**> parsel\_2**

[1] 3.5 4.6 3.9 4.3

**> parsel\_3**

[1] 3.7 4.8 4.6 4.8

**> parsel\_4**

[1] 3.4 4.4 4.7 4.3

**Normallik Kontrolü**

**##** parsel\_1 için

> shapiro.test(parsel\_1)

Shapiro-Wilk normality test

data: parsel\_1

W = 0.9072, p-value = 0.4679

H0: 1. parsele ait buğday ekim sonuçları ile normal dağılım arasında fark yoktur.

HS: 1. parsele ait buğday ekim sonuçları ile normal dağılım arasında fark vardır.

**Yorum:** Sig.=0,05 < p-value = 0.4679 olduğundan H0 hipotezi kabul edilir. 19,5 metre kare büyüklüğündeki 1. parsele ait buğday ekim sonuçları ile normal dağılım arasında fark olmadığı %95 güven düzeyinde söylenebilir.

**##** parsel\_2 için

> shapiro.test(parsel\_2)

Shapiro-Wilk normality test

data: parsel\_2

W = 0.9841, p-value = 0.9254

H0: 2. parsele ait buğday ekim sonuçları ile normal dağılım arasında fark yoktur.

HS: 2. parsele ait buğday ekim sonuçları ile normal dağılım arasında fark vardır.

**Yorum:** Sig.=0,05 < p-value = 0.9254 olduğundan H0 hipotezi kabul edilir.

19,5 metre kare büyüklüğündeki 2. parsele ait buğday ekim sonuçları ile normal dağılım arasında fark olmadığı %95 güven düzeyinde söylenebilir.

**##** parsel\_3 için

> shapiro.test(parsel\_3)

Shapiro-Wilk normality test

data: parsel\_3

W = 0.9203, p-value = 0.41

H0: 3. parsele ait buğday ekim sonuçları ile normal dağılım arasında fark yoktur.

HS: 3. parsele ait buğday ekim sonuçları ile normal dağılım arasında fark vardır.

**Yorum:** Sig.=0,05 < p-value = 0.41 olduğundan H0 hipotezi kabul edilir.

19,5 metre kare büyüklüğündeki 3. parsele ait buğday ekim sonuçları ile normal dağılım arasında fark olmadığı %95 güven düzeyinde söylenebilir.

**##** parsel\_4 için

> shapiro.test(parsel\_4)

Shapiro-Wilk normality test

data: parsel\_4

W = 0.8811, p-value = 0.3434

H0: 4. parsele ait buğday ekim sonuçları ile normal dağılım arasında fark yoktur.

HS: 4. parsele ait buğday ekim sonuçları ile normal dağılım arasında fark vardır.

**Yorum:** Sig.=0,05 < p-value = 0.3434 olduğundan H0 hipotezi kabul edilir.

19,5 metre kare büyüklüğündeki 4. parsele ait buğday ekim sonuçları ile normal dağılım arasında fark olmadığı %95 güven düzeyinde söylenebilir.

Sonuç olarak, her bir parsele ait verilerin dağılımı Normal Dağılım göstermektedir.

**Normallik için Grafikler**

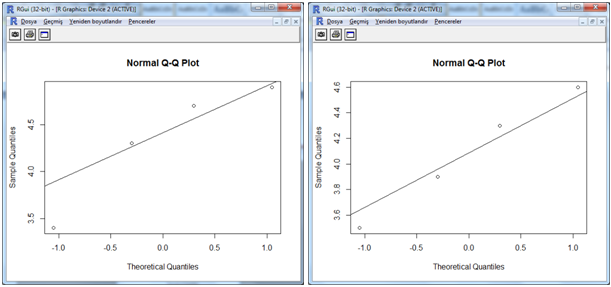
qqnorm(parsel\_1)

qqline(parsel\_1)

qqnorm(parsel\_2)

qqline(parsel\_2)

**parsel 1 parsel 2**



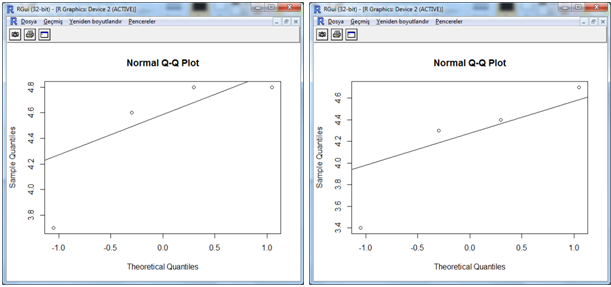
qqnorm(parsel\_3)

qqline(parsel\_3)

qqnorm(parsel\_4)

qqline(parsel\_4)

**parsel 3 parsel 4**



boxplot(kg~parsel,data=veri) // 4 ayrı parsele ait box plot çizdirilebilir.

**Varyanslasrın Homojenliği Testi**

library(lawstat) // lawstat paketini kütüphanemize ekleyelim.

> levene.test(kg,parsel, location="mean")

classical Levene's test based on the absolute deviations from the mean ( none not applied because the location is not set to median )

data: kg

Test Statistic = 0.09, p-value = 0.9642

H0: Her bir parsele ait varyanslar homojendir.

HS: Her bir parsele ait varyanslar homojen değildir.

**Yorum:** Sig.=0,05 <p-value = 0.9642 olduğundan H0 hipotezi kabul edilir.

Her bir parsele ait varyansların homojen olduğu %95 güven düzeyinde söylenebilir.

**##** Anova testini uygulayabilmek için gerekli olan varsayımlar sağlandı.

> anv=aov(kg~parsel+duzey, data=veri)

> summary(anv)

Df Sum Sq Mean Sq F value Pr(>F)

parsel 1 0.000 0.00013 0.000 0.984

duzey 2 0.352 0.17588 0.558 0.587

Residuals 12 3.783 0.31521

H0: Buğday verimi yönünden parsel türleri arasında fark yoktur.

HS: Buğday verimi yönünden parsel türleri arasında fark vardır.

**Yorum:** Pr(>F)=0.984 > Sig.=0.05 olduğundan H0 kabul edilir.

Parsel türleri arasında verim açısından fark olmadığı %0,05 anlamlılık düzeyinde söylenebilir.

H0: Buğday verimi yönünden deneme(düzey) türleri arasında fark yoktur.

HS: Buğday verimi yönünden deneme(düzey) türleri arasında fark vardır.

**Yorum:** Pr(>F)=0.587 > Sig.=0.05 olduğundan H0 kabul edilir.

Deneme(duzey) türleri arasında fark olmadığı %0,05 anlamlılık düzeyinde söylenebilir.

**Sonuç:** 19,5 metre kare büyüklüğündeki parsellere yapılan buğday ekim kg. sonuçlarına parsellerin(1,2,3,4) ve düzeylerin(0,Na,2Na,3Na) etkisi olmadığı %5 anlamlılık düzeyinde söylenebilir.

Burada sorulması gereken tablodaki düzey olarak almış olduğum 0, Na, 2Na, 3Na sütunların aslında sodyum olup olmadığıdır. Eğer sodyum iseler ve farklılığın bulunması durumunda bizim yapmamız gereken Post-Hoc testlerinden Dunnett testini uygulamalı ve kontrol grubu ile karşılaştırılmalıdır.

Son olarak; çok verimli ve keyifli geçen bu dönem için Serpil Hocama ve uygulamalarda ki özverisi ve her zamanki yardımseverliğinden dolayı Hande hocama da sonsuz teşekkürler...

Metin USLU

21076757

uslumetin@gmail.com