Análisis y Diseño de Experimentos

Segundo Trabajo Encargado

Se realizó un experimento para observar el rendimiento en kilogramos por parcela de 5 variedades de garbanzo (A, B, C, D, E) en el cual se tuvo que utilizar el diseño Cuadrado Latino. Las filas fueron definidas como niveles de riego y las columnas como fertilidad del suelo. Los datos se presentan a continuación

Niveles	Fertilidad del suelo				
de riego	1	2	3	4	5
1	B = 65	C = 80	A = 55	E = 83	D = 80
2	C = 95	A = 60	E = 94	D = 95	B = 62
3	A = 63	E = 98	D = 79	B = 69	C = 100
4	E = 97	D =94	B = 46	C = 71	A = 42
5	D = 76	B = 54	C = 106	A = 36	E = 96

a. Presente el modelo aditivo lineal e interprete cada uno de sus componentes en términos del enunciado.

El modelo aditivo lineal para un Diseño Cuadrado Latino es el siguiente:

$$Y_{(i)jk} = \mu + \tau_{(i)} + \beta_j + \gamma_k + \epsilon_{(i)jk}$$

Para todo : i, j, k = 1,…,t

Donde:

- $Y_{(i)jk}$ es el valor o rendimiento observado en el i-ésimo variedad de Garbanzo, j-ésima nivel de riesgo, k-ésima fertilidad del suelo.
- μ es el efecto de la media general.
- $\tau_{(i)}$ es el efecto del i-ésima variedad de garbanzo
- β_j es el efecto de la j-ésima nivel de riesgo
- γ_k es el efecto de la k-ésima fertilidad del suelo
- $\epsilon_{(i)jk}$ es el efecto del error experimental en el i-ésimo variedad de Garbanzo, j-ésima nivel de riesgo, k-ésima fertilidad del suelo.
- t es el número de tratamientos que es igual al número de niveles de riego y al número de fertilidades de suelo.

b. Realice el diagnóstico del modelo utilizando gráfico y pruebas de hipótesis

```
library(agricolae)
garbanzo <- read.table("garbanzo.txt", header = TRUE)
head(garbanzo)
```

```
##   rendimiento fila columna tratamiento
## 1          65    1       1           B
## 2          95    1       2           C
## 3          63    1       3           A
## 4          97    1       4           E
## 5          76    1       5           D
## 6          80    2       1           C
```

```
str(garbanzo)
```

```
## 'data.frame':    25 obs. of  4 variables:
##  $ rendimiento: int  65 95 63 97 76 80 60 98 94 54 ...
```

```
##  $ fila       : int  1 1 1 1 1 2 2 2 2 2 ...
##  $ columna    : int  1 2 3 4 5 1 2 3 4 5 ...
##  $ tratamiento: Factor w/ 5 levels "A","B","C","D",..: 2 3 1 5 4 3 1 5 4 2 ...
```

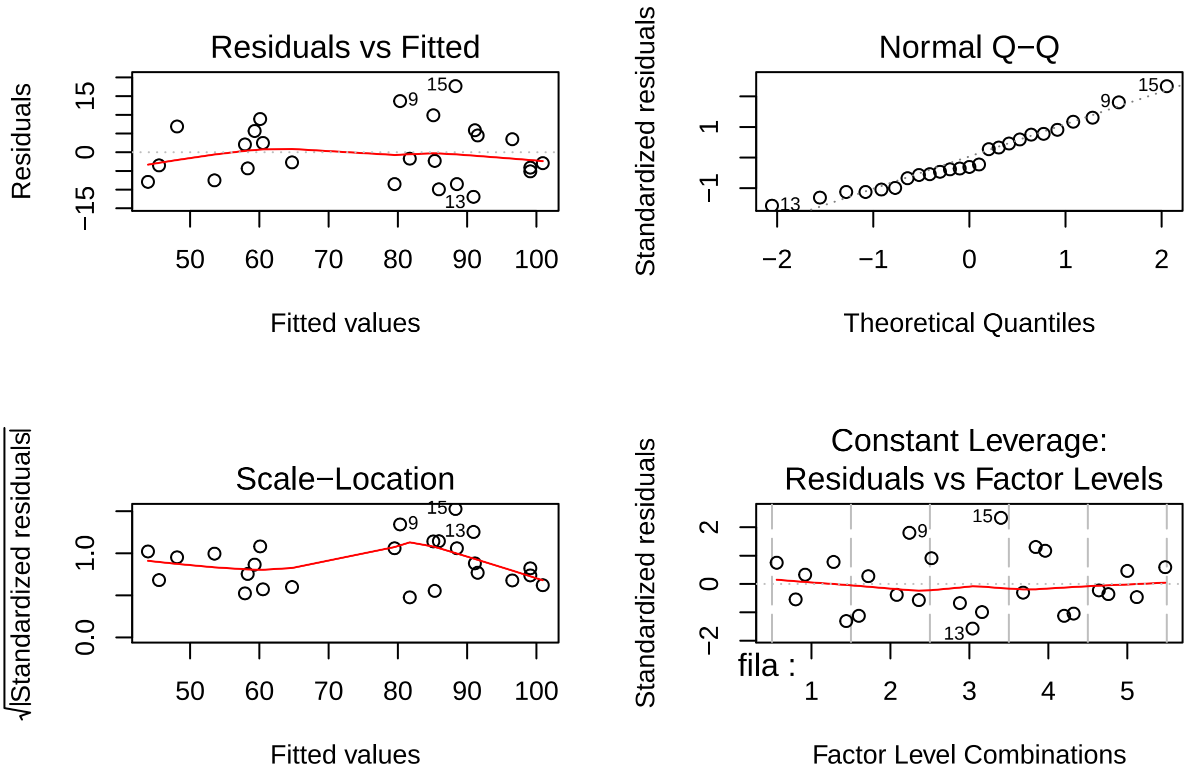
```r
garbanzo$fila<-factor(garbanzo$fila)
garbanzo$columna<-factor(garbanzo$columna)
garbanzo$tratamiento<-factor(garbanzo$tratamiento)
str(garbanzo)
```

```
## 'data.frame':    25 obs. of  4 variables:
##  $ rendimiento: int  65 95 63 97 76 80 60 98 94 54 ...
##  $ fila       : Factor w/ 5 levels "1","2","3","4",..: 1 1 1 1 1 2 2 2 2 2 ...
##  $ columna    : Factor w/ 5 levels "1","2","3","4",..: 1 2 3 4 5 1 2 3 4 5 ...
##  $ tratamiento: Factor w/ 5 levels "A","B","C","D",..: 2 3 1 5 4 3 1 5 4 2 ...
```

```r
mod<-lm(rendimiento~.,data=garbanzo)
anva<-anova(mod)
anva
```

```
## Analysis of Variance Table
##
## Response: rendimiento
##             Df Sum Sq Mean Sq F value   Pr(>F)
## fila         4  193.0   48.24  0.4030 0.802966
## columna      4  569.4  142.34  1.1891 0.364656
## tratamiento  4 7458.6 1864.64 15.5767 0.000107 ***
## Residuals   12 1436.5  119.71
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
par(mfrow=c(2,2))
plot(mod)
```




```
ri<-rstandard(mod)
shapiro.test(ri)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  ri
## W = 0.95981, p-value = 0.4108
```

```
library(car)
```

```
## Loading required package: carData
```

```
ncvTest(mod)
```

```
## Non-constant Variance Score Test
## Variance formula: ~ fitted.values
## Chisquare = 0.4684703, Df = 1, p = 0.49369
```

```
library(multcomp)
```

```
## Loading required package: mvtnorm
```

```
## Loading required package: survival
```

```
## Loading required package: TH.data
```

```
## Loading required package: MASS
```

```
##
## Attaching package: 'TH.data'
```

```
## The following object is masked from 'package:MASS':
##
```

```
##     geyser
```

```
amod<-aov(rendimiento~.,data=garbanzo)
comptrat<-glht(amod,linfct=mcp(tratamiento="Tukey"))
summary(comptrat)
```

```
##
##   Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: aov(formula = rendimiento ~ ., data = garbanzo)
##
## Linear Hypotheses:
##            Estimate Std. Error t value Pr(>|t|)
## B - A == 0     8.00       6.92   1.156 0.774818
## C - A == 0    39.20       6.92   5.665 0.000810 ***
## D - A == 0    33.60       6.92   4.856 0.002975 **
## E - A == 0    42.40       6.92   6.127 0.000373 ***
## C - B == 0    31.20       6.92   4.509 0.005298 **
## D - B == 0    25.60       6.92   3.700 0.020808 *
## E - B == 0    34.40       6.92   4.971 0.002509 **
## D - C == 0    -5.60       6.92  -0.809 0.922787
## E - C == 0     3.20       6.92   0.462 0.989389
## E - D == 0     8.80       6.92   1.272 0.712140
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)
```

```
library(agricolae)
anva<-aov(mod)
compd<-duncan.test(anva,"tratamiento")
compd
```

```
## $statistics
##    MSerror Df  Mean      CV
##   119.7067 12 75.84 14.4265
##
## $parameters
##     test      name.t ntr alpha
##   Duncan tratamiento   5  0.05
##
## $duncan
##      Table CriticalRange
## 2 3.081307      15.07680
## 3 3.225244      15.78108
## 4 3.312453      16.20779
## 5 3.370172      16.49021
##
## $means
##   rendimiento       std r Min Max Q25 Q50 Q75
## A        51.2 11.691878 5  36  63  42  55  60
## B        59.2  9.203260 5  46  69  54  62  65
## C        90.4 14.501724 5  71 106  80  95 100
```

```
## D        84.8  8.983318 5  76  95  79  80  94
## E        93.6  6.107373 5  83  98  94  96  97
##
## $comparison
## NULL
##
## $groups
##   rendimiento groups
## E        93.6      a
## C        90.4      a
## D        84.8      a
## B        59.2      b
## A        51.2      b
##
## attr(,"class")
## [1] "group"
```

g. Se desea comparar la media de los rendimientos obtenidos con la variedades A y B versus la media de los rendimientos obtenidos con las variedades C, D y E. Obtenga los contrastes ortogonales y utilice la prueba de F para probar el contraste dado

$$HO : \mu_A + \mu_B = \mu_C + \mu_D + \mu_E$$
$$H1 : \mu_A + \mu_B \neq \mu_C + \mu_D + \mu_E$$