Diseño de experimentos

Exámen de semana 01

Kevin Heberth Haquehua Apaza

08 de setiembre del 2025

Table of Contents

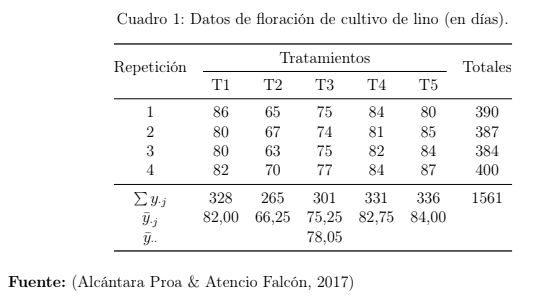
# DCA-Comparaciones múltiples-supuestos

## EJERCICIO

Un grupo de ingenieros agrónomos realizaron una investigación experimental en el fundo Huancayo en el distrito de Huariaca, provincia Pasco, región Pasco, en condiciones de campo con el fin de determinar el efecto de cinco fertilizantes complejos en días de floración del cultivo de lino *(Linum usitatissimum)*

**Descripción de los tratamientos (fertilizantes)**

* YaraVera AMIDAS
* YaraMila COMPLEX
* YaraTera KRISTA K
* N20, P20, K20
* N18, P15, K10



### Preguntas

1. Plantee y describa el modelo que considere adecuado para analizar este caso.
2. ¿Se puede afirmar que al menos uno de los días de floración difiere del resto al aplicar los distintos fertilizantes complejos?
3. Verifique si el modelo propuesto cumple con los supuestos.
4. Estime el coeficiente de variabilidad.
5. ¿Cuáles son los fertilizantes que tienen mejor efecto en la floración media en el cultivo de lino?
6. ¿El tiempo promedio de floración al aplicar el fertilizante es superior al aplicar el fertilizante en más de 2 días?
7. Al aplicar los fertilizantes y (en forma conjunta), ¿el tiempo medio en hacer efecto la floración es menor que al aplicar los fertilizantes y (en forma conjunta)?

**Indicaciones finales**

1. El documento debe utilizar un lenguaje académico.
2. Subir el informe final-documento en pdf.
3. Subir el Script y data comprimido (se descontaran puntos).
4. Subir tu informe antes de las 13:00 pm (07-09-2025)

### Solución

1. **Plantee y describa el modelo que considere adecuado para analizar este caso.**

Tomando en cuenta el contexto en el que solo se desea evaluar el efecto de cinco fertilizantes, el modelo adecuado a utilizar es un diseño completo al azar. En las cuales se tienen los tratamiento de los distintos fertilizantes, si evaluamos el estudio original indica que se utiliza el diseño bloque completo al azar, en donde las repeteciones vienen a ser los bloques.

El modelo aditivo lineal es el siguiente

Donde:

* Es la floración de *(Linum usitatissimum)* del i-ésimo tratamiento y j-ésima repetición.
* es el efecto de la media general en la floración de *(Linum usitatissimum)*.
* es el efecto del i-ésimo tratamiento.
* es el efecto del error experimental en el i-ésimo tratamiento y j-ésima repetición. para tratamientos y repeticiones. Además, el modelo asume que

1. **¿Se puede afirmar que al menos uno de los días de floración difiere del resto al aplicar los distintos fertilizantes complejos?**

Empezemos desarrollando los datos del archivo Excel que también se adjuntará en el trabajo, para posteriormente hacer la lectura del archivo.

library(readxl)  
library(here)

## here() starts at C:/Users/La Fuente Estadistic/Desktop/INLASER-LA FUENTE/General\_Analytics/Maestria\_UNSAAC

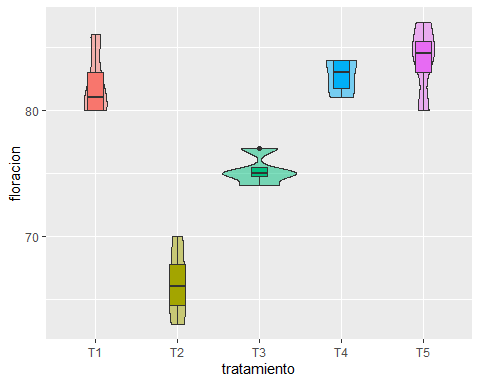
data <- read\_excel(here("11 Diseno Experimentos/Trabajo 1/data\_floracion.xlsx"))  
head(data)

## # A tibble: 6 × 2  
## tratamiento floracion  
## <chr> <dbl>  
## 1 T1 86  
## 2 T2 65  
## 3 T3 75  
## 4 T4 84  
## 5 T5 80  
## 6 T1 80

attach(data) # Para tomar los datos  
data$tratamiento = as.factor(tratamiento) # convierte tratamiento en factor

Empezemos viendo un gráfico de la dispersión y el efecto de los tratamientos en la floración del cultivo de *(Linum usitatissimum)*

library(ggplot2)  
ggplot(data, aes(x = tratamiento, y = floracion, fill = tratamiento)) +  
 geom\_violin(alpha = 0.5) +  
 theme(legend.position = "none")+  
 geom\_boxplot(width = 0.2)



**Interpretación:** Se observa que el YaraMila COMPLEX presenta la menor floración de *(Linum usitatissimum)*, seguido del YaraTera KRISTA K. Seguidamente los tratamientos: YaraVera AMIDAS, (N20, P20, K20), (N18, P15, K10) presentan la mayor floración de *(Linum usitatissimum)* sin mostrar diferencias entre ellos.

Veamos la media general

med\_general = mean(floracion); med\_general

## [1] 78.05

Ahora la media de los tratamientos

library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

med\_trat = aggregate(floracion~tratamiento, FUN = mean)  
med\_trat%>% rename(Media = floracion) %>% mutate(Efecto = Media-med\_general)

## tratamiento Media Efecto  
## 1 T1 82.00 3.95  
## 2 T2 66.25 -11.80  
## 3 T3 75.25 -2.80  
## 4 T4 82.75 4.70  
## 5 T5 84.00 5.95

Se observa que el tratamiento que tiene el menor efecto en la floración de *(Linum usitatissimum)* es YaraMila COMPLEX y el que presenta el mayor efecto es (N18, P15, K10)

Ahora definamos el modelo

mod.dca = lm(floracion ~ tratamiento) # modelo DCA

Y veamos el análisis de la varianza para afirmar que al menos uno de los días de floración difiere del resto al aplicar los distintos fertilizantes complejos.

summary(aov(mod.dca)) # Análisis de varianza

## Df Sum Sq Mean Sq F value Pr(>F)   
## tratamiento 4 880.7 220.18 37.42 1.23e-07 \*\*\*  
## Residuals 15 88.3 5.88   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Efectivamente, se puede afirmar que al menos uno de los días de floración de *(Linum usitatissimum)* difiere del resto al aplicar los distintos fertilizantes complejos. Ya que existe evidencia suficiente para rechazar la hipótesis nula que indica que los días de floración de *(Linum usitatissimum)* no difieran en los tratamientos complejos.

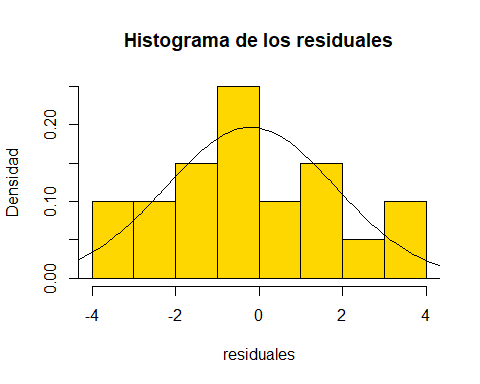
1. **Verifique si el modelo propuesto cumple con los supuestos.**

* **Normalidad de errores**

Nos planteamos las hipótesis Los errores se distribuyen normalmente Los errores no se distribuyen normalmente

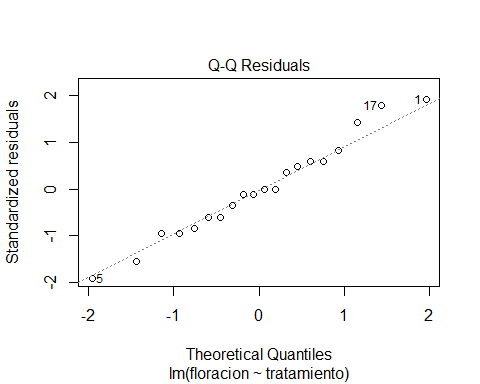
Veamos los residuales del modelo

library(agricolae)  
residuales = residuals(mod.dca) # extrae residuales  
histo = hist(residuales, col="gold", xlab="residuales", main="Histograma de los residuales", ylab="Densidad",prob=1)  
normal.freq(histo, frequency=3) # densidad



Se observa que los errores se distribuyen a una normal, veamos el QQ-plot

plot(mod.dca, which = 2)



A excepción de unos valores, la mayor parte se centra a la recta de residuales, veamos ahora la prueba de Shapiro-Wilk de normalidad de residuos

shapiro.test(residuales)

##   
## Shapiro-Wilk normality test  
##   
## data: residuales  
## W = 0.97749, p-value = 0.8977

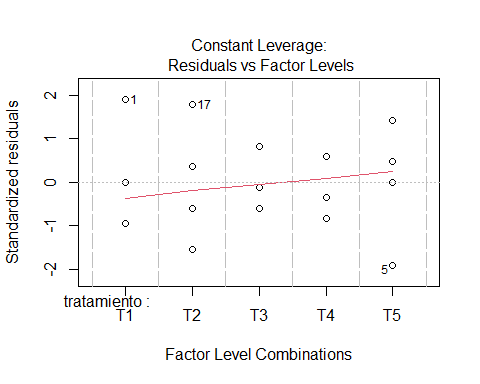
Como se tiene un pvalor no se rechaza lo cual indica que el modelo cumple el supuesto de normalidad de errores.

* **Homocedasticidad**

Nos planteamos las hipótesis: Existe homogeneidad de varianzas No existe homogeneidad de varianzas

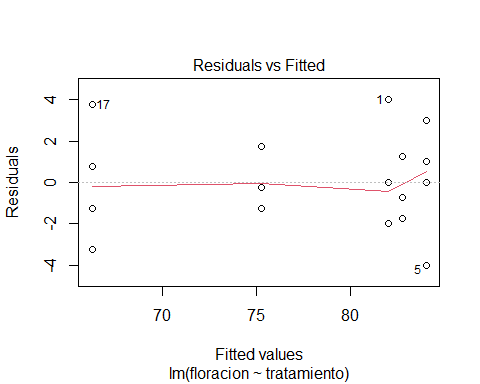
Veamos si la variabilidad es distinta en los factores

plot(mod.dca, which = 5)



Y si la variabilidad se incrementa con la media general

plot(mod.dca, which = 1)



Se observa que la varianza se mantiene constante en los residuos. Utilizemos la prueba de Breusch-Pagan para evaluar este supuesto

library(car)

## Loading required package: carData

##   
## Attaching package: 'car'

## The following object is masked from 'package:dplyr':  
##   
## recode

ncvTest(mod.dca)

## Non-constant Variance Score Test   
## Variance formula: ~ fitted.values   
## Chisquare = 0.06641474, Df = 1, p = 0.79663

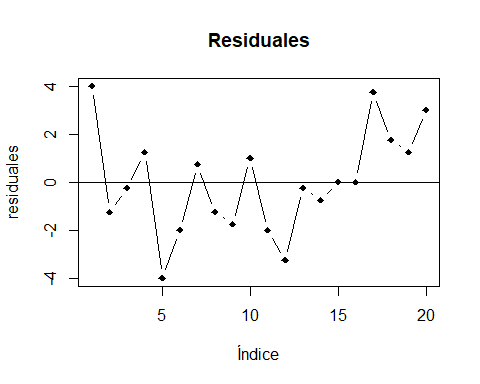
Como el pvalor no se rechaza en lo cual se indica que la varianza es constante en los residuos, por lo que se cumple el supuesto de homocedasticidad.

* **Independencia de errores**

Para ver la no autocorrelación de los errores nos planteamos las hipótesis Existe independencia de errores No existe independencia de errores

Veamos de forma gráfica la autocorrelación de los errores

plot(residuals(mod.dca), pch = 18,  
 type = "b", ylab = "residuales",  
 xlab = "Índice", main="Residuales")  
abline(h=0)



No se observa un patrón de dependencia en los errores, realizemos la prueba de Durbin-Watson

library(lmtest)

## Loading required package: zoo

##   
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':  
##   
## as.Date, as.Date.numeric

dwtest(mod.dca,alternative = c("two.sided"))

##   
## Durbin-Watson test  
##   
## data: mod.dca  
## DW = 1.3994, p-value = 0.3096  
## alternative hypothesis: true autocorrelation is not 0

Como el pvalor no se rechaza lo cual indica que existe independencia de errores o no están autocorrelacionados de primer orden, lo cual indica que se cumple el supuesto de independencia de errores.

1. **Estime el coeficiente de variabilidad.**
2. **¿Cuáles son los fertilizantes que tienen mejor efecto en la floración media en el cultivo de lino?**
3. **¿El tiempo promedio de floración al aplicar el fertilizante es superior al aplicar el fertilizante en más de 2 días?**
4. **Al aplicar los fertilizantes y (en forma conjunta), ¿el tiempo medio en hacer efecto la floración es menor que al aplicar los fertilizantes y (en forma conjunta)?**