Diseño de experimentos

Exámen de semana 02

Kevin Heberth Haquehua Apaza

18 de setiembre del 2025

Tabla de Contenidos

[DISEÑO BLOQUE COMPLETO AL AZAR 1](#_Toc209121789)

[EJEMPLO: RENDIMIENTO DE PAPA PERUANA CON DIFERENTES FÓRMULAS DE FERTILIZANTE 1](#_Toc209121790)

[Preguntas 2](#_Toc209121791)

[Solución 2](#_Toc209121792)

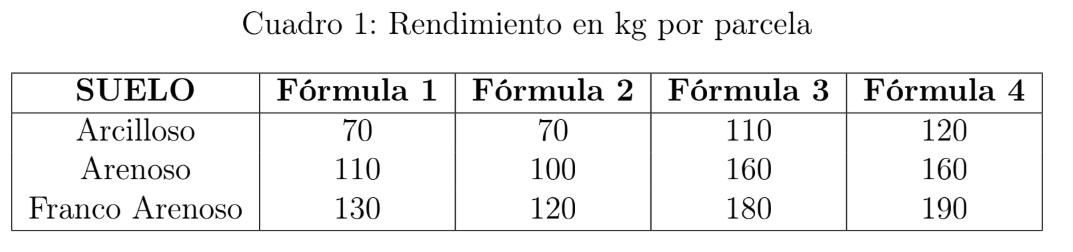
# DISEÑO BLOQUE COMPLETO AL AZAR

## EJEMPLO: RENDIMIENTO DE PAPA PERUANA CON DIFERENTES FÓRMULAS DE FERTILIZANTE

Se realizó un experimento para determinar si existe alguna diferencia en el rendimiento de papa variedad peruana con 4 fórmulas diferentes de fertilizante. Las fórmulas empleadas fueron las siguientes:

* **Fórmula 1 (Testigo):** Sin fósforo (P) y sin nitrógeno (N).
* **Fórmula 2:** Solo fósforo.
* **Fórmula 3:** Solo nitrógeno.
* **Fórmula 4:** Con fósforo y nitrógeno.

Una variante en particular en la conducción del experimento fue el tipo de suelo, ya que no fue el mismo para todas las parcelas en estudio. Los rendimientos obtenidos en Kg por parcela se presentan a continuación:



### Preguntas

1. Presente el modelo aditivo lineal y explique sus componentes según el enunciado de la pregunta.
2. Asumiendo el cumplimiento de supuestos, pruebe si al menos una **fórmula** presenta un rendimiento medio de papa distinto a las demás.
3. Se planeó evaluar si con la **fórmula 4** se obtienen mejores rendimientos que con la **fórmula 2**, ¿es cierta la información que se planteó?
4. Un especialista afirma que el rendimiento medio de papa con la **fórmula 2** es diferente al obtenido con la **fórmula 1**. ¿Es cierta la afirmación del especialista?. Use el reporte y la prueba estadística adecuada.
5. El especialista desea comparar la **fórmula que no contiene fósforo ni nitrógeno** con el resto de fórmulas; realice la prueba estadística más adecuada.
6. El especialista desea realizar todas las comparaciones posibles entre las fórmulas empleadas; realice la prueba estadística más adecuada.
7. El especialista afirma que el rendimiento medio de papa obtenido al aplicar la **fórmula 1** es inferior al rendimiento medio de papa cuando se aplica conjuntamente la **fórmula 3 y 4**. ¿Es cierta la afirmación del especialista? Realice la prueba estadística más adecuada.

**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 (14-09-2025)

### Solución

1. *Presente el modelo aditivo lineal y explique sus componentes según el enunciado de la pregunta.*

Al ser un diseño de bloques completo al azar con una observación por unidad experimental, la observación que representa el rendimiento por kg de las papas puede representarse por el modelo siguiente:

Donde:

* Es el rendimiento por kg de papa obtenida por el -ésimo bloque (tipo de suelo) sujeta al tratamiento (fórmula).
* El efecto de la media común.
* Es el efecto de la fórmula , ().
* Es el efecto del tipo de suelo , ().
* Es una variable aleatoria no observable llamado error.

1. *Asumiendo el cumplimiento de supuestos, pruebe si al menos una* ***fórmula*** *presenta un rendimiento medio de papa distinto a las demás.*

Nos planteamos las hipótesis

Abrimos la bd creada

library(readxl)  
library(here)

## Warning: package 'here' was built under R version 4.5.1

## here() starts at C:/Users/Kev/Desktop/Trabajo\_Kevin/Maestria\_UNSAAC

data <- read\_excel(here("11 Diseno Experimentos/Examen\_2/data\_papa.xlsx"))  
head(data)

## # A tibble: 6 × 3  
## formula suelo rendimiento  
## <chr> <chr> <dbl>  
## 1 f1 arcilloso 70  
## 2 f2 arcilloso 70  
## 3 f3 arcilloso 110  
## 4 f4 arcilloso 120  
## 5 f1 arenoso 110  
## 6 f2 arenoso 100

attach(data)  
modeg<-lm(rendimiento~suelo+formula)  
anva<-anova(modeg)  
anva

## Analysis of Variance Table  
##   
## Response: rendimiento  
## Df Sum Sq Mean Sq F value Pr(>F)   
## suelo 2 8016.7 4008.3 131.182 1.118e-05 \*\*\*  
## formula 3 8666.7 2888.9 94.546 1.929e-05 \*\*\*  
## Residuals 6 183.3 30.6   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Se obtiene un pvalor menor a 0.05 por lo tanto no aceptamos por lo que podemos decir que una fórmula presenta un rendimiento de papa distinto a los demás.

1. *Se planeó evaluar si con la* ***fórmula 4*** *se obtienen mejores rendimientos que con la* ***fórmula 2****, ¿es cierta la información que se planteó?*

Nos planteamos las hipótesis

modeg

##   
## Call:  
## lm(formula = rendimiento ~ suelo + formula)  
##   
## Coefficients:  
## (Intercept) sueloarenoso suelofranco arenoso   
## 69.167 40.000 62.500   
## formulaf2 formulaf3 formulaf4   
## -6.667 46.667 53.333

cm<-anva$Mean #Valores de la tabla anova

efect<-modeg$coefficients  
dmedia<-efect-efect[6]  
dmedia<-dmedia[4]  
dmedia

## formulaf2   
## -60

tc<-dmedia/sqrt(cm[3]\*(2/5))  
tc

## formulaf2   
## -17.16233

pvalue<-2\*pt(tc,df.residual(modeg))  
pvalue

## formulaf2   
## 2.505253e-06

Se observa un pvalor menor a 0.05, por lo que no aceptamos indicando que existen evidencia existencia de que la diferencia de la fórmula 4 con la fórmula 2 es negativo. Indicando que la fórmula 4 da mejores resultados que la fórmula 2.

1. *Un especialista afirma que el rendimiento medio de papa con la* ***fórmula 2*** *es diferente al obtenido con la* ***fórmula 1****. ¿Es cierta la afirmación del especialista?. Use el reporte y la prueba estadística adecuada.*

Realizamos la comprobación por pares, nos planteamos la hipótesis

modeg

##   
## Call:  
## lm(formula = rendimiento ~ suelo + formula)  
##   
## Coefficients:  
## (Intercept) sueloarenoso suelofranco arenoso   
## 69.167 40.000 62.500   
## formulaf2 formulaf3 formulaf4   
## -6.667 46.667 53.333

Se observa que la fórmula 2 con respecto a la fórmula 1 se tiene una diferencia de -6.667 indicando que la fórmula 2 es mayor su rendimiento con respecto a la fórmula 1. Para realizar la prueba de hipótesis adecuada utilizemos las comparaciones múltiples de Tukey-Cramer (Tukey HSD)

model\_formula <- aov(rendimiento ~ formula, data = data)  
TukeyHSD(model\_formula, "formula")

## Tukey multiple comparisons of means  
## 95% family-wise confidence level  
##   
## Fit: aov(formula = rendimiento ~ formula, data = data)  
##   
## $formula  
## diff lwr upr p adj  
## f2-f1 -6.666667 -90.37821 77.04488 0.9937160  
## f3-f1 46.666667 -37.04488 130.37821 0.3458669  
## f4-f1 53.333333 -30.37821 137.04488 0.2506036  
## f3-f2 53.333333 -30.37821 137.04488 0.2506036  
## f4-f2 60.000000 -23.71155 143.71155 0.1782406  
## f4-f3 6.666667 -77.04488 90.37821 0.9937160

Se tiene un p-valor mayor a 0.05 por lo que no rechazamos , es decir que no hay suficiente evidencia para indicar que la fórmula 2 es diferente a la fórmula 1.

1. *El especialista desea comparar la* ***fórmula que no contiene fósforo ni nitrógeno*** *con el resto de fórmulas; realice la prueba estadística más adecuada.*

Tomando en cuenta que la fórmula que no contiene fósforo ni nitrógeno (fórmula 1) es un grupo control se usa la prueba estadística de Dunnet

library(agricolae)

## Warning: package 'agricolae' was built under R version 4.5.1

library(multcomp)

## Warning: package 'multcomp' was built under R version 4.5.1

## Cargando paquete requerido: mvtnorm

## Warning: package 'mvtnorm' was built under R version 4.5.1

## Cargando paquete requerido: survival

## Cargando paquete requerido: TH.data

## Cargando paquete requerido: MASS

##   
## Adjuntando el paquete: 'TH.data'

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

data$formula <- factor(data$formula)  
data$suelo <- factor(data$suelo)  
modg = aov(rendimiento~formula+suelo, data = data)  
summary(glht(modg, linfct = mcp(formula = "Dunnett")))

##   
## Simultaneous Tests for General Linear Hypotheses  
##   
## Multiple Comparisons of Means: Dunnett Contrasts  
##   
##   
## Fit: aov(formula = rendimiento ~ formula + suelo, data = data)  
##   
## Linear Hypotheses:  
## Estimate Std. Error t value Pr(>|t|)   
## f2 - f1 == 0 -6.667 4.513 -1.477 0.389   
## f3 - f1 == 0 46.667 4.513 10.340 <0.001 \*\*\*  
## f4 - f1 == 0 53.333 4.513 11.817 <0.001 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
## (Adjusted p values reported -- single-step method)

Se observa que solo con la fórmula 2 no se tienen diferencias, mientras que con la fórmula 3 y 4 se tienen diferencias.

1. *El especialista desea realizar todas las comparaciones posibles entre las fórmulas empleadas; realice la prueba estadística más adecuada.*

Para realizar todas las combinaciones posibles utilizamos la prueba de Tukey

library(emmeans)

## Warning: package 'emmeans' was built under R version 4.5.1

## Welcome to emmeans.  
## Caution: You lose important information if you filter this package's results.  
## See '? untidy'

medias <- emmeans(modg, ~ formula)  
comparaciones\_tukey <- contrast(medias, method = "pairwise", adjust = "tukey")  
summary(comparaciones\_tukey)

## contrast estimate SE df t.ratio p.value  
## f1 - f2 6.67 4.51 6 1.477 0.5033  
## f1 - f3 -46.67 4.51 6 -10.340 0.0002  
## f1 - f4 -53.33 4.51 6 -11.817 0.0001  
## f2 - f3 -53.33 4.51 6 -11.817 0.0001  
## f2 - f4 -60.00 4.51 6 -13.294 <.0001  
## f3 - f4 -6.67 4.51 6 -1.477 0.5033  
##   
## Results are averaged over the levels of: suelo   
## P value adjustment: tukey method for comparing a family of 4 estimates

Se observa que respecto a la fórmula 1 y 2 no se tienen diferencias entre ellas, además de que entre la fórmula 3 y 4 tampoco se tienen diferencias, con respecto entre las otras combinaciones se encontraron diferencias significativas.

1. *El especialista afirma que el rendimiento medio de papa obtenido al aplicar la* ***fórmula 1*** *es inferior al rendimiento medio de papa cuando se aplica conjuntamente la* ***fórmula 3 y 4****. ¿Es cierta la afirmación del especialista? Realice la prueba estadística más adecuada.*

Nos planteamos la expresión

Nos planteamos las hipótesis - -

Y por medio de los contrastes ortogonales realizamos la comparación

library(gmodels)

## Warning: package 'gmodels' was built under R version 4.5.1

data$formula <- factor(data$formula, levels = c("f1","f2","f3","f4"))  
con = c(1, 0, -0.5, -0.5)  
fit.contrast(modg , "formula", con)

## Estimate Std. Error t value Pr(>|t|)  
## formula c=( 1 0 -0.5 -0.5 ) -50 3.90868 -12.79204 1.4015e-05

Nos indican diferencias significativas lo cual indica que el rendimiento medio de la fórmula 1 es inferior al rendimiento medio de papa cuando se aplica conjuntamente la formula 3 y 4.