

caso 2. Diseño factorial

Arturo Zuñiga

2025-09-19

Caso 2:

Se desea evaluar el efecto de 3 recipientes (R1, R2, y R3) para la producción de plántulas y dos especies de eucaliptos (E1 y E2). La variable respuesta medida fue la altura de las plántulas, en cm, a los 80 días de edad. Los recipientes y las especies fueron:

R1: bolsa plástica pequeña

R2: bolsa plástica grande

R3: laminado

E1: Eucalyptus citriodora

E2: Eucalyptus grandis

Variable respuesta: ...

Factores en estudio: ...

Tratamientos: ...

Lectura y presentación de datos

```
datos.2 = read.table("datos2.txt",T) # Lectura de datos
datos.2 # Presentación de los datos
```

##	Altura	Envase	Especie
## 1	26.2	R1	E1
## 2	26.0	R1	E1
## 3	25.0	R1	E1
## 4	25.4	R1	E1
## 5	25.7	R2	E1
## 6	26.3	R2	E1
## 7	25.1	R2	E1
## 8	26.4	R2	E1
## 9	22.8	R3	E1
## 10	19.4	R3	E1
## 11	18.8	R3	E1
## 12	19.2	R3	E1
## 13	26.8	R1	E2
## 14	26.6	R1	E2
## 15	26.7	R1	E2
## 16	26.2	R1	E2
## 17	25.6	R2	E2
## 18	27.1	R2	E2
## 19	25.0	R2	E2

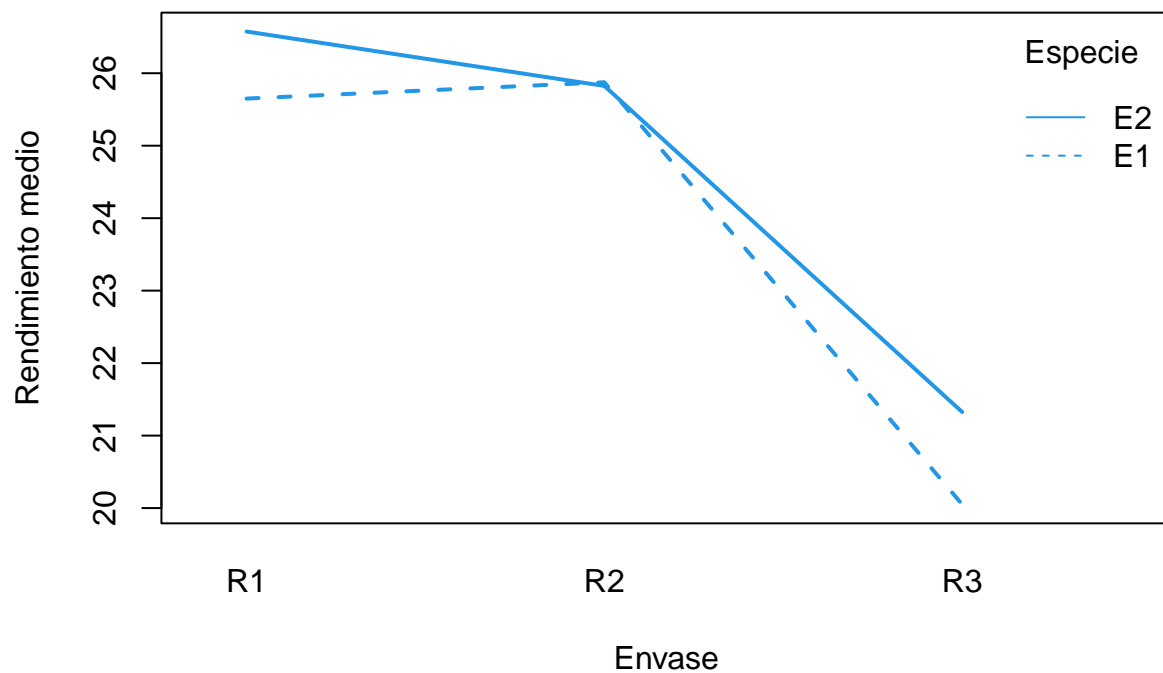
```
## 20 25.6 R2 E2
## 21 19.8 R3 E2
## 22 21.4 R3 E2
## 23 22.8 R3 E2
## 24 21.3 R3 E2
```

```
attach(datos.2) # Asignaci3n de variables
xtabs(Altura ~ Envase + Especie) # Presentaci3n de la tabla de datos
```

```
##      Especie
## Envase  E1  E2
##    R1 102.6 106.3
##    R2 103.5 103.3
##    R3  80.2  85.3
```

Gráfico de interacci3n

```
interaction.plot(Envase, Especie, Altura, lwd = 2, col = 4, ylab = "Rendimiento medio")
```



```
library(agricolae)
mod.ef2 = lm(Altura ~ Envase * Especie) # Construcci3n del modelo
summary(aov(mod.ef2)) # Análisis de varianza
```

```
##      Df Sum Sq Mean Sq F value    Pr(>F)
## Envase      2 149.74    74.87  69.501 3.42e-09 ***
## Especie      1   3.08     3.08   2.861  0.108
## Envase:Especie 2   1.89     0.94   0.875  0.434
## Residuals    18  19.39     1.08
```

```

## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

cv.model(mod.ef2)      # Coeficiente de variabilidad

## [1] 4.285863

Verificaci?n de supuestos

library(nortest)
library(car)

## Cargando paquete requerido: carData
shapiro.test(residuals(mod.ef2)) # Prueba de normalidad de errores

##
##  Shapiro-Wilk normality test
##
## data:  residuals(mod.ef2)
## W = 0.92659, p-value = 0.08183
ad.test(residuals(mod.ef2)) # Prueba de normalidad de errores

##
##  Anderson-Darling normality test
##
## data:  residuals(mod.ef2)
## A = 0.55617, p-value = 0.1349
library(car)
ncvTest(mod.ef2) # Prueba de homocedasticidad (homogeneidad de varianzas)

## Non-constant Variance Score Test
## Variance formula: ~ fitted.values
## Chisquare = 11.57844, Df = 1, p = 0.00066721
library(lmtest)

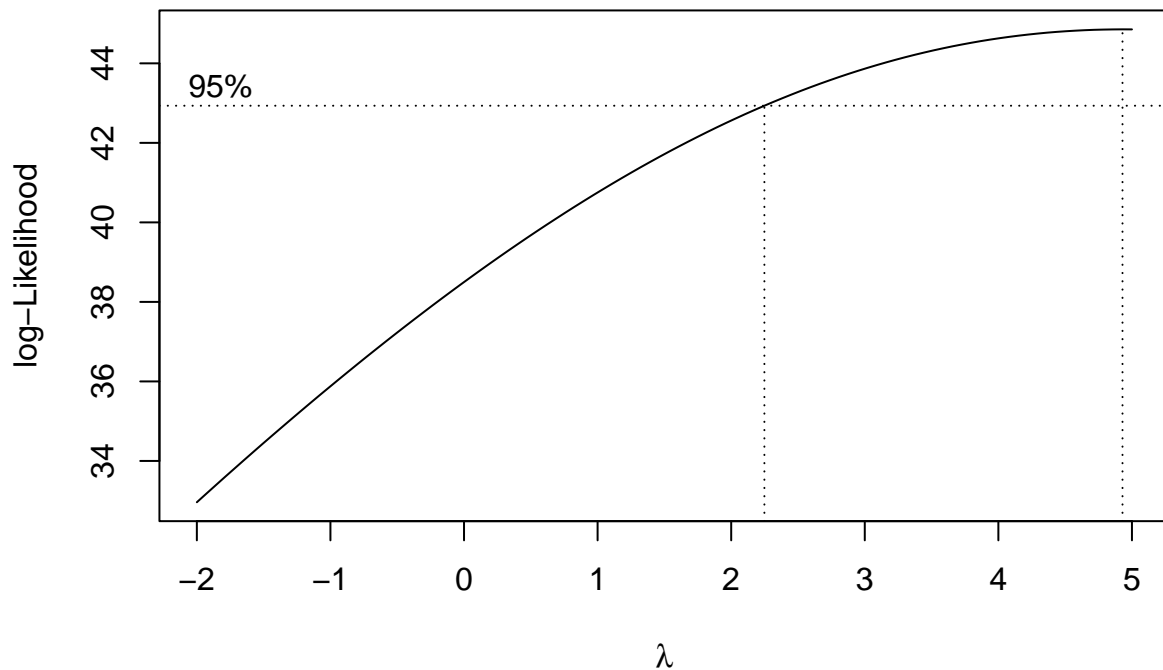
## Cargando paquete requerido: zoo
##
## Adjuntando el paquete: 'zoo'
## The following objects are masked from 'package:base':
##
##   as.Date, as.Date.numeric
dwtest(mod.ef2,alternative = c("two.sided")) # Prueba de independencia de errores

##
##  Durbin-Watson test
##
## data:  mod.ef2
## DW = 1.9951, p-value = 0.2449
## alternative hypothesis: true autocorrelation is not 0

Transformaci?n de datos

library(MASS)
boxcox(mod.ef2, plotit=TRUE, lambda = seq(-2, 5, 1/10)) # Gr?fico de Box Cox para elegir lambda

```



```
boxcox(mod.ef2, plotit=FALSE, lambda = seq(-2, 5, 1/10)) # Elegir lambda que maximice
```

```
## $x
## [1] -2.0 -1.9 -1.8 -1.7 -1.6 -1.5 -1.4 -1.3 -1.2 -1.1 -1.0 -0.9 -0.8 -0.7 -0.6
## [16] -0.5 -0.4 -0.3 -0.2 -0.1 0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9
## [31] 1.0 1.1 1.2 1.3 1.4 1.5 1.6 1.7 1.8 1.9 2.0 2.1 2.2 2.3 2.4
## [46] 2.5 2.6 2.7 2.8 2.9 3.0 3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8 3.9
## [61] 4.0 4.1 4.2 4.3 4.4 4.5 4.6 4.7 4.8 4.9 5.0
##
## $y
## [1] 32.96094 33.26406 33.56479 33.86308 34.15887 34.45211 34.74272 35.03064
## [9] 35.31582 35.59818 35.87767 36.15420 36.42772 36.69815 36.96542 37.22945
## [17] 37.49019 37.74755 38.00145 38.25183 38.49861 38.74170 38.98104 39.21655
## [25] 39.44815 39.67576 39.89931 40.11872 40.33391 40.54482 40.75135 40.95346
## [33] 41.15105 41.34406 41.53242 41.71606 41.89491 42.06892 42.23801 42.40214
## [41] 42.56124 42.71525 42.86414 43.00784 43.14631 43.27952 43.40742 43.52997
## [49] 43.64716 43.75894 43.86530 43.96621 44.06167 44.15166 44.23617 44.31520
## [57] 44.38875 44.45684 44.51945 44.57663 44.62837 44.67470 44.71566 44.75126
## [65] 44.78156 44.80657 44.82635 44.84094 44.85039 44.85475 44.85408
```

```
library(forecast)
```

```
## Registered S3 method overwritten by 'quantmod':
## method from
## as.zoo.data.frame zoo
```

```
Altura2 = BoxCox(Altura,4.6)
```

```
datos.2$Altura2 = Altura2
mod.ef2b = lm(Altura2 ~ Envase * Especie) # Construcci?n del modelo
summary(aov(mod.ef2b))
```

```
##              Df      Sum Sq   Mean Sq F value    Pr(>F)
## Envase        2 1.069e+12 5.343e+11  79.121 1.21e-09 ***
## Especie       1 2.050e+10 2.050e+10   3.035  0.0985 .
## Envase:Especie 2 1.410e+10 7.050e+09   1.044  0.3725
## Residuals    18 1.216e+11 6.754e+09
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
cv.model(mod.ef2b)
```

```
## [1] 14.83578
```

Verificando los supuestos

```
shapiro.test(residuals(mod.ef2b)) # Prueba de normalidad de errores
```

```
##
##  Shapiro-Wilk normality test
##
## data:  residuals(mod.ef2b)
## W = 0.93593, p-value = 0.1323
```

```
ncvTest(mod.ef2b) # Prueba de homocedasticidad (homogeneidad de varianzas)
```

```
## Non-constant Variance Score Test
## Variance formula: ~ fitted.values
## Chisquare = 0.7083212, Df = 1, p = 0.4
```

```
dwtest(mod.ef2b,alternative = c("two.sided")) # Prueba de independencia de errores
```

```
##
##  Durbin-Watson test
##
## data:  mod.ef2b
## DW = 2.2325, p-value = 0.5988
## alternative hypothesis: true autocorrelation is not 0
```

Podemos ver que la interacci?n resulta no significativa:

$$H_0 : (\alpha\beta)_{ij} = 0$$

$$H_1 : (\alpha\beta)_{ij} \neq 0 \text{ para al menos un } i, j.$$

$$\alpha = 0.10$$

$$p - \text{valor} = 0.3725$$

As?, pasamos a analizar los efectos principales. Para el envase:

$$H_0 : \alpha_i = 0$$

$$H_1 : \alpha_i \neq 0 \text{ para al menos un } i.$$

$$\alpha = 0.10$$

$$p - \text{valor} = 1.21 \times 10^{-9}$$

Entonces se rechaza la hip?tesis nula, lo que significa que existen diferencias significativas en la altura media para al menos uno de los envases en comparaci?n.

Luego, para la especie:

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0 \text{ para al menos un } j.$$

$$\alpha = 0.10$$

$$p\text{-valor} = 0.0985$$

Entonces se rechaza la hipótesis nula, lo que significa que existen diferencias significativas en la altura media para al menos uno de las especies en comparación, sin embargo estas diferencias son menores si son comparadas con las diferencias entre envases.

Sabemos que existen diferencias en las alturas medias cuando se usan distintos envases, pero ¿cuál(es) maximiza(n) la altura? ¿cuál(es) la minimiza(n)? Es necesario realizar comparaciones múltiples para responder a esta pregunta.

```
# usa el nombre real de tu data.frame, aquí supongo df
```

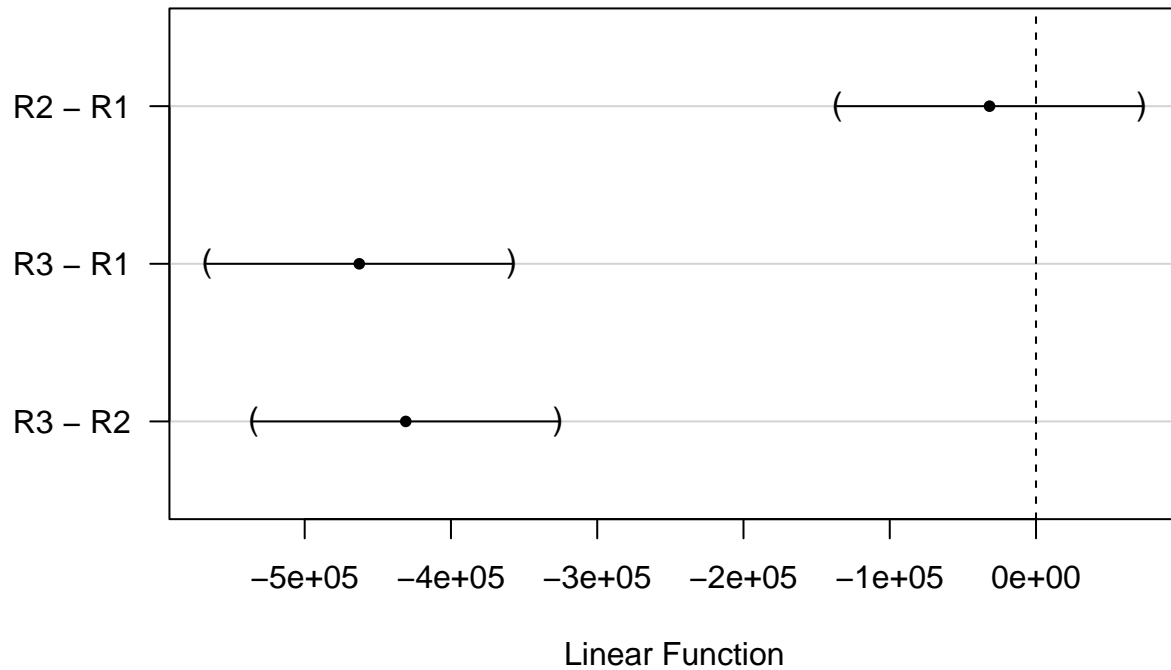
```
datos.2$Envase <- factor(datos.2$Envase)
datos.2$Especie <- factor(datos.2$Especie)
```

```
# ajusta el modelo indicando 'data' explícitamente
mod <- aov(Altura2 ~ Envase + Especie, data = datos.2)
```

```
library(multcomp)
comp.2a <- glht(mod, linfct = mcp(Envase = "Tukey"))
summary(comp.2a); plot(comp.2a)
```

```
##
## Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: aov(formula = Altura2 ~ Envase + Especie, data = datos.2)
##
## Linear Hypotheses:
##           Estimate Std. Error t value Pr(>|t|)
## R2 - R1 == 0   -31792      41180  -0.772    0.724
## R3 - R1 == 0  -462686      41180 -11.236 <1e-04 ***
## R3 - R2 == 0  -430895      41180 -10.464 <1e-04 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)
```

95% family-wise confidence level



```
comp.2b <- glht(mod, linfct = mcp(Especie = "Tukey"))
summary(comp.2b); plot(comp.2b)
```

```
##
## Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: aov(formula = Altura2 ~ Envase + Especie, data = datos.2)
##
## Linear Hypotheses:
##           Estimate Std. Error t value Pr(>|t|)
## E2 - E1 == 0    58450      33623   1.738  0.0975 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)
```

95% family-wise confidence level

