

LAB.-4-Pruebas-de-t-27-08-25.R

angel

2025-11-26

```
#####  
# LAB. 4 Prueba de t Caso de muestras independientes  
# FLOR ANGELI CRUZ ROSALES  
# DR. MARCO A. GONZALEZ TAGLE  
# 27/08/25  
#####
```

```
# Verificar que todo esté actualizado  
packageVersion("rmarkdown")
```

```
## [1] '2.30'
```

```
packageVersion("knitr")
```

```
## [1] '1.50'
```

```
packageVersion("tinytex")
```

```
## [1] '0.58'
```

```
library(readr)
```

```
## Warning: package 'readr' was built under R version 4.5.2
```

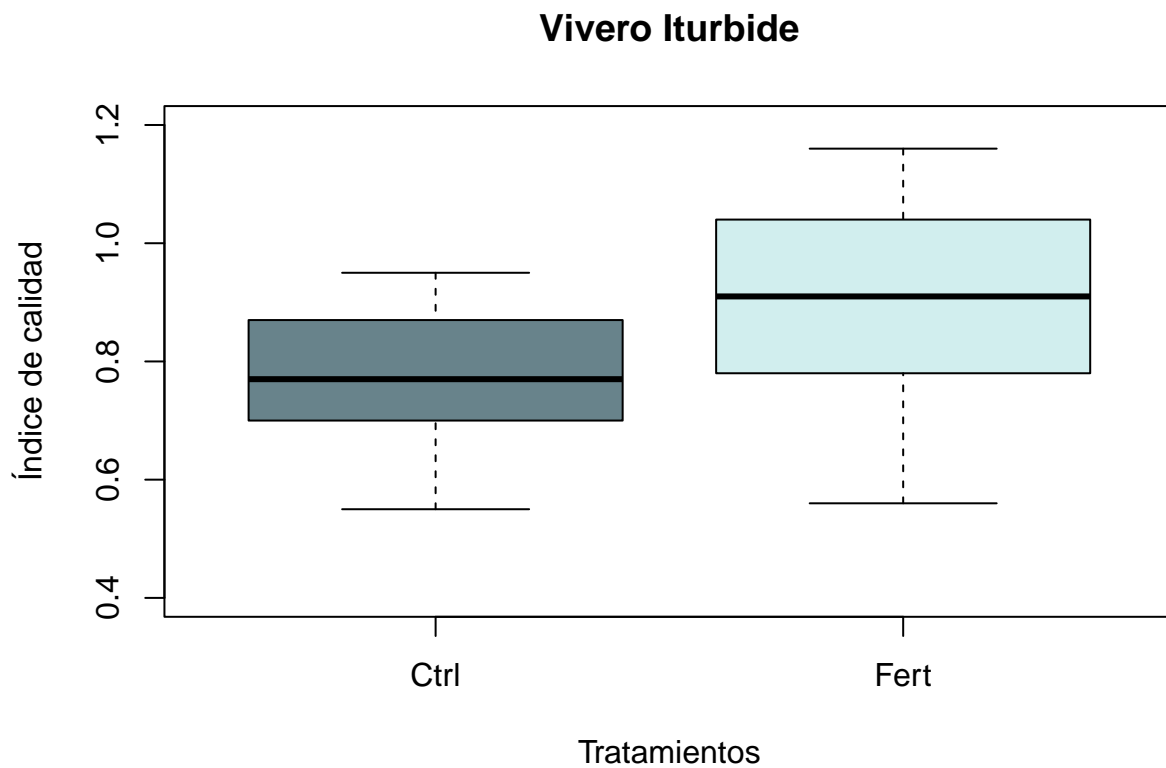
```
calidad_plantas <- read_csv("calidad_plantas.csv")
```

```
## Rows: 42 Columns: 3  
## -- Column specification -----  
## Delimiter: ","  
## chr (1): Tratamiento  
## dbl (2): planta, IE  
##  
## i Use 'spec()' to retrieve the full column specification for this data.  
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
View(calidad_plantas)

# Función as.factor para que lo tome como un factor, para que lo tome
# como dos grupos --> Ctrl y Fert
calidad_plantas$Tratamiento <- as.factor(calidad_plantas$Tratamiento)

colores <- c("#68838B", "#D1EEEE")
boxplot(calidad_plantas$IE ~ calidad_plantas$Tratamiento,
        col= colores,
        xlab= "Tratamientos",
        ylab= "Índice de calidad",
        ylim= c(0.4,1.2),
        main="Vivero Iturbide")
```



```
# Estadística descriptiva: tapply sirve para obtener un valor cuando tenemos
# varios grupos

# Función tapply, mean= promedio
tapply(calidad_plantas$IE,calidad_plantas$Tratamiento,mean)
```

```
##      Ctrl      Fert
## 0.7676190 0.9066667
```

```
# Ctrl 0.7676190 Fert 0.9066667

# Función tapply, var= varianza
tapply(calidad_plantas$IE, calidad_plantas$Tratamiento, var)
```

```
##          Ctrl          Fert
## 0.01329905 0.03238333
```

```
# Ctrl 0.01329905 Fert 0.03238333

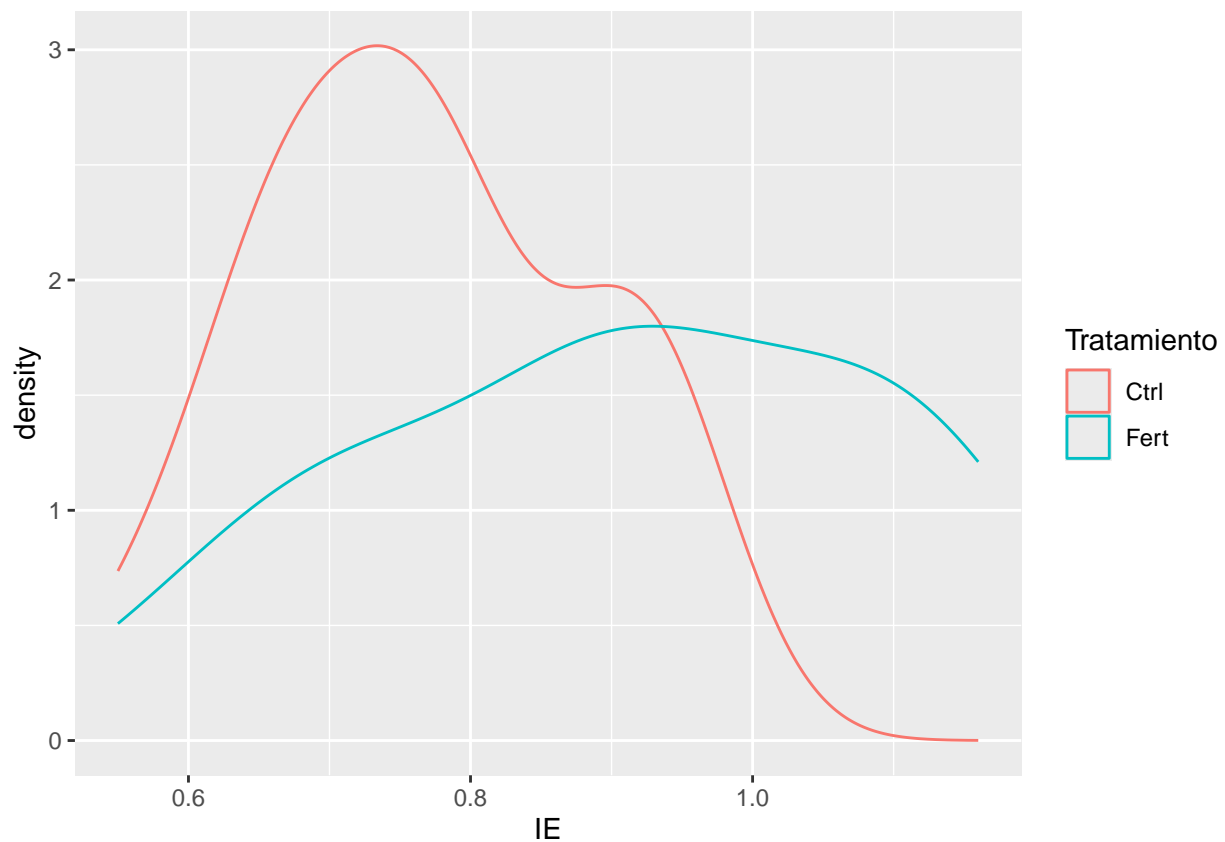
# varianza del grupo fert es 3 veces más
# grande que el grupo Control (Ctrl)

# Revisar el comportamiento de los datos: Primero se debe correr
# el paquete que instalamos

library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.5.2
```

```
ggplot(calidad_plantas, aes(x=IE, color=Tratamiento))+
  geom_density()
```



```

# Línea roja son los datos de ctrl, azul es la de fertilizante

# Función tapply, sd= desviación estándar

tapply(calidad_plantas$IE, calidad_plantas$Tratamiento, sd)

##      Ctrl      Fert
## 0.1153215 0.1799537

# Ctrl 0.1153215 Fert 0.1799537

# Como separar los datos por tratamiento usando subset

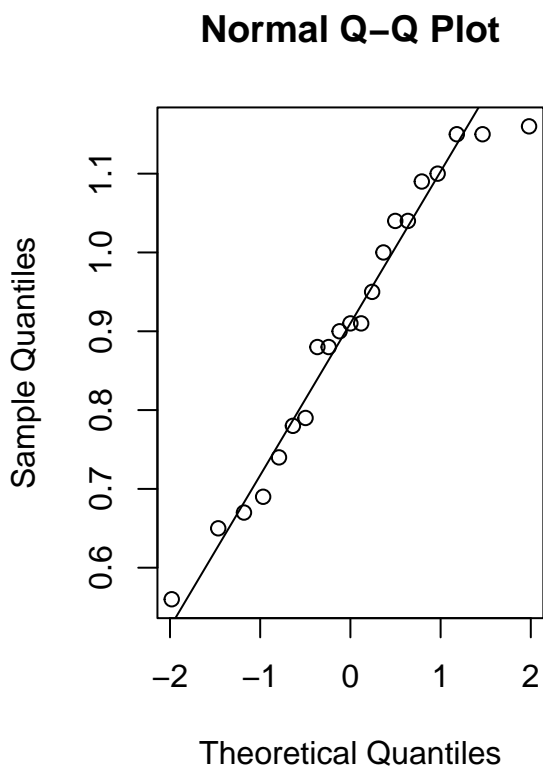
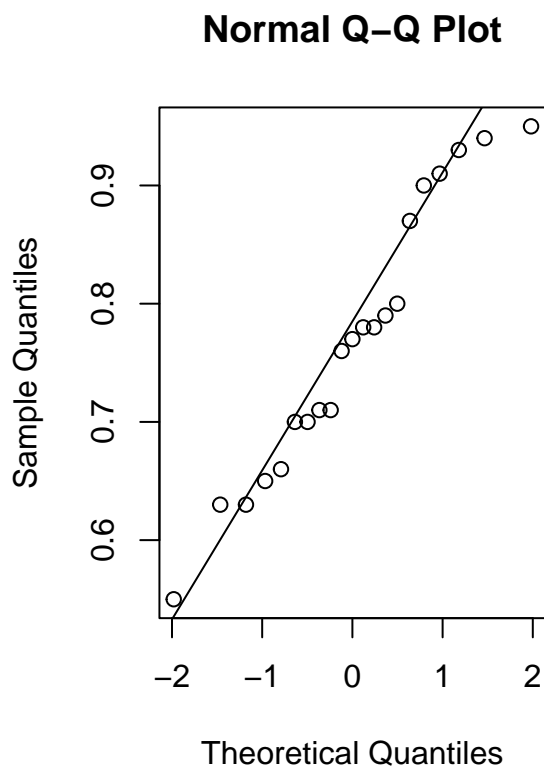
df_ctrl <- subset(calidad_plantas, Tratamiento == "Ctrl")
df_fert <- subset(calidad_plantas, Tratamiento == "Fert")

# qqnorm es un gráfico que nos ayudará a revisar la normalidad de los datos

# Función para que la ventana de gráficos permita que nos
# aparezca dos gráficos par(mfrow)
# Una fila con dos columnas (1,2)

par(mfrow=c(1,2))
qqnorm(df_ctrl$IE); qqline(df_ctrl$IE)
qqnorm(df_fert$IE); qqline(df_fert$IE)

```



```
par(mfrow=c(1,1)) #Mostrarse un solo gráfico
```

```
# Normalidad de los datos (prueba Shapiro)
```

```
shapiro.test(df_ctlr$IE)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: df_ctlr$IE  
## W = 0.9532, p-value = 0.3908
```

```
# data:df_ctlr$IE  
# W = 0.9532  
# p-value = 0.3908
```

```
shapiro.test(df_fert$IE)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: df_fert$IE  
## W = 0.95339, p-value = 0.3941
```

```
# data:df_Fert$IE  
# W = 0.95339,  
# p-value = 0.3941  
# Normalidad, p-value es mayor a 0.05
```

```
# Revisar homogeneidad de varianzas (criterio)
```

```
var.test(calidad_plantas$IE ~ calidad_plantas$Tratamiento)
```

```
##  
## F test to compare two variances  
##  
## data: calidad_plantas$IE by calidad_plantas$Tratamiento  
## F = 0.41068, num df = 20, denom df = 20, p-value = 0.05304  
## alternative hypothesis: true ratio of variances is not equal to 1  
## 95 percent confidence interval:  
## 0.1666376 1.0121038  
## sample estimates:  
## ratio of variances  
## 0.4106757
```

```
# p-value = 0.05304  
# Normalidad
```

```
# Criterios para una prueba de t student:  
# 1- Distribución normal de los datos
```

```

# 2- Homogeneidad de varianzas
# 3- Más de 30 datos

# Aplicar la prueba de t, varianzas iguales *dos colas = two sided *

t.test(calidad_plantas$IE ~ calidad_plantas$Tratamiento,
       var.equal=T,
       alternative="two.sided")

##
## Two Sample t-test
##
## data:  calidad_plantas$IE by calidad_plantas$Tratamiento
## t = -2.9813, df = 40, p-value = 0.004868
## alternative hypothesis: true difference in means between group Ctrl and group Fert is not equal to 0
## 95 percent confidence interval:
##  -0.23331192 -0.04478332
## sample estimates:
## mean in group Ctrl mean in group Fert
##      0.7676190      0.9066667

# Mean in group 0.7676190 Ctrl
# Mean in group 0.9066667 Fert
# 95% confidence interval: [-0.2333, -0.0448]

# Existe diferencia significativa entre grupos.
# El fertilizante AUMENTA el índice de eficiencia.
# La diferencia real está entre -0.2333 y -0.0448,
# y nuestra diferencia observada (-0.139) cae dentro de este rango

cohens_efecto <- function(x,y) {
  n1 <- length(x); n2 <- length(y)
  s1 <- sd(x); s2 <- sd(y)
  sp <- sqrt(((n1-1)* s1^2 + (n2-1)* s2^2)/(n1+n2 - 2))
  (mean(x)- mean(y))/sp
}

d_cal <- cohens_efecto(df_ctrl$IE, df_fert$IE)

cohens_efecto(df_ctrl$IE, df_fert$IE)

## [1] -0.9200347

# Cohen's d = -0.9200347
# El tamaño del efecto es grande por lo que se puede concluir que
# el efecto del fertilizante tiene un gran impacto en las plantulas

```