

Laboratorio-2.R

Usuario

2025-09-03

```
# Laboratorio 2 -----
--
# Eusebio Alberto Espinosa Saucedo
# 1854858

# 03/09/25

# Base de datos Iris -----
--
data("iris")
View(iris)
head(iris)

##   Sepal.Length Sepal.Width Petal.Length Petal.Width Species
## 1           5.1          3.5          1.4          0.2   setosa
## 2           4.9          3.0          1.4          0.2   setosa
## 3           4.7          3.2          1.3          0.2   setosa
## 4           4.6          3.1          1.5          0.2   setosa
## 5           5.0          3.6          1.4          0.2   setosa
## 6           5.4          3.9          1.7          0.4   setosa

# Resumen general de La base
summary(iris)

##   Sepal.Length   Sepal.Width   Petal.Length   Petal.Width
## Min.   :4.300   Min.   :2.000   Min.   :1.000   Min.   :0.100
## 1st Qu.:5.100   1st Qu.:2.800   1st Qu.:1.600   1st Qu.:0.300
## Median :5.800   Median :3.000   Median :4.350   Median :1.300
## Mean   :5.843   Mean   :3.057   Mean   :3.758   Mean   :1.199
## 3rd Qu.:6.400   3rd Qu.:3.300   3rd Qu.:5.100   3rd Qu.:1.800
## Max.   :7.900   Max.   :4.400   Max.   :6.900   Max.   :2.500
##      Species
## setosa    :50
## versicolor:50
## virginica :50
##
##
##

names(iris)

## [1] "Sepal.Length" "Sepal.Width"  "Petal.Length" "Petal.Width"
## "Species"
```

```

# Seleccionar especies -----
--

# Filtrar solo versicolor y virginica
data_sub <- subset(iris, Species %in% c("versicolor", "virginica"))

# Comprobar el subconjunto
table(data_sub$Species)

##
##      setosa versicolor  virginica
##          0         50         50

head(data_sub)

##      Sepal.Length Sepal.Width Petal.Length Petal.Width   Species
## 51           7.0         3.2         4.7         1.4 versicolor
## 52           6.4         3.2         4.5         1.5 versicolor
## 53           6.9         3.1         4.9         1.5 versicolor
## 54           5.5         2.3         4.0         1.3 versicolor
## 55           6.5         2.8         4.6         1.5 versicolor
## 56           5.7         2.8         4.5         1.3 versicolor

# Estadística descriptiva -----
--

# Calcular descriptivos
tapply(data_sub$Petal.Length, data_sub$Species, summary)

## $setosa
## NULL
##
## $versicolor
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      3.00   4.00   4.35   4.26   4.60   5.10
##
## $virginica
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      4.500  5.100  5.550  5.552  5.875  6.900

aggregate(Petal.Length ~ Species, data=data_sub, mean)

##      Species Petal.Length
## 1 versicolor      4.260
## 2  virginica      5.552

aggregate(Petal.Length ~ Species, data=data_sub, sd)

##      Species Petal.Length
## 1 versicolor    0.4699110
## 2  virginica    0.5518947

```

```

# Plantamiento de hipótesis -----
--

# Hipótesis:

#  $H_0: \mu_1 = \mu_2$  (No hay diferencia en la media de Petal.Length entre
versicolor
# y virginica )

#  $H_1$ : sí hay diferencia en las medias de Petal.Length

# Pureba de t -----

# Prueba de normalidad

shapiro.test(data_sub$Petal.Length)

##
## Shapiro-Wilk normality test
##
## data: data_sub$Petal.Length
## W = 0.99099, p-value = 0.7445

# Prueba de homogeneidad de varianzas (F-test)
var.test(Petal.Length ~ Species, data=data_sub)

##
## F test to compare two variances
##
## data: Petal.Length by Species
## F = 0.72497, num df = 49, denom df = 49, p-value = 0.2637
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
## 0.411402 1.277530
## sample estimates:
## ratio of variances
## 0.7249678

# Por defecto usamos Welch
t.test(Petal.Length ~ Species, data=data_sub, var.equal = FALSE)

##
## Welch Two Sample t-test
##
## data: Petal.Length by Species
## t = -12.604, df = 95.57, p-value < 2.2e-16
## alternative hypothesis: true difference in means between group
versicolor and group virginica is not equal to 0
## 95 percent confidence interval:

```

```
## -1.49549 -1.08851
## sample estimates:
## mean in group versicolor mean in group virginica
##                4.260                5.552

# Efecto -----
--

# Datos por especie

versicolor <- subset(data_sub, Species=="versicolor")$Petal.Length
virginica  <- subset(data_sub, Species=="virginica")$Petal.Length

# Estadística
n1 <- length(versicolor)
n2 <- length(virginica)

m1 <- mean(versicolor)
m2 <- mean(virginica)

s1 <- sd(versicolor)
s2 <- sd(virginica)

s_pooled <- sqrt(((n1-1)*s1^2 + (n2-1)*s2^2) / (n1+n2-2))

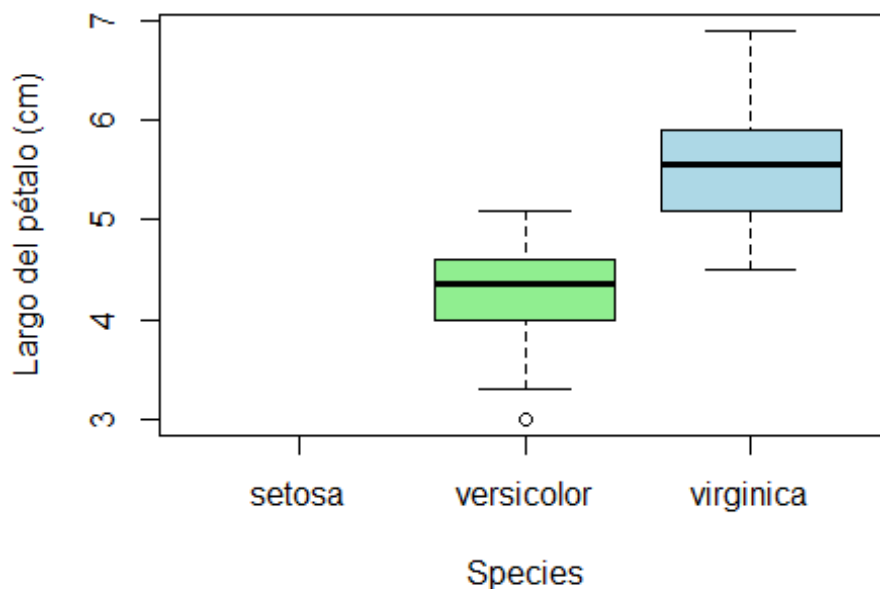
# Cohen's d
d <- (m1 - m2) / s_pooled
d

## [1] -2.520756

# Gráfico comparativo -----
--

boxplot(Petal.Length ~ Species, data=data_sub,
        col=c("lightblue", "lightgreen"),
        main="Comparación del largo de pétalo vs especies",
        ylab="Largo del pétalo (cm)")
```

Comparación del largo de pétalo vs especies



```
# Interpretación de resultados -----  
--  
  
# El análisis realizado sobre la base de datos iris permitió evaluar si  
# existían diferencias en la longitud de los pétalos (Petal.Length) entre  
# las especies  
# Iris versicolor e Iris virginica. Los resultados descriptivos iniciales  
# mostraron que la media de la longitud de los pétalos en versicolor fue  
# de  
# aproximadamente 4.26 cm, mientras que en virginica alcanzó alrededor de  
# 5.55 cm. Esta diferencia preliminar fue confirmada con la aplicación de  
# una prueba t de dos muestras independientes.  
  
# El resultado de la prueba t de Welch indicó un valor de t cercano a -  
# 12.6,  
# con un valor-p muy inferior a 0.001. Esto permitió rechazar la  
# hipótesis nula  
# de igualdad de medias y concluir que sí existe una diferencia  
# significativa  
# en el largo de los pétalos entre ambas especies. La magnitud de esta  
# diferencia se cuantificó mediante el tamaño del efecto (Cohen's d ≈  
# 2.5),  
# considerado extremadamente grande, lo cual significa que la separación  
# entre  
# las distribuciones de versicolor y virginica es muy marcada.
```

La visualización mediante diagramas de caja reforzó esta conclusión,
mostrando claramente que la distribución de virginica se encuentra
desplazada
hacia valores más altos en comparación con versicolor. En conjunto,
los resultados no solo evidencian una diferencia estadísticamente
significativa, sino que además demuestran una separación biológicamente
relevante, ya que el largo de los pétalos constituye un rasgo
morfológico
sólido y confiable para la diferenciación de estas dos especies
dentro del género Iris.