

HW02_-JIGA.R

ji_ti

2025-08-31

```
# Asigancion 3: Contraste de medias
# 30/08/2025
# Jorge Ivan Garza Aldape
# 1677695
```

```
# Base de datos Iris -----
```

```
data(iris)
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
##
##
##
```

```
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
```

```
View(iris)
str(iris)
```

```
## 'data.frame':   150 obs. of  5 variables:
##  $ Sepal.Length: num  5.1 4.9 4.7 4.6 5 5.4 4.6 5 4.4 4.9 ...
```

```
## $ Sepal.Width : num 3.5 3 3.2 3.1 3.6 3.9 3.4 3.4 2.9 3.1 ...
## $ Petal.Length: num 1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...
## $ Petal.Width : num 0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...
## $ Species      : Factor w/ 3 levels "setosa","versicolor",...: 1 1 1 1 1 1 1 1 1 1 ...
```

```
# Operador %in%
```

```
iris_sub <- subset(iris, Species %in% c("versicolor", "virginica"))
iris_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
## 57	6.3	3.3	4.7	1.6	versicolor
## 58	4.9	2.4	3.3	1.0	versicolor
## 59	6.6	2.9	4.6	1.3	versicolor
## 60	5.2	2.7	3.9	1.4	versicolor
## 61	5.0	2.0	3.5	1.0	versicolor
## 62	5.9	3.0	4.2	1.5	versicolor
## 63	6.0	2.2	4.0	1.0	versicolor
## 64	6.1	2.9	4.7	1.4	versicolor
## 65	5.6	2.9	3.6	1.3	versicolor
## 66	6.7	3.1	4.4	1.4	versicolor
## 67	5.6	3.0	4.5	1.5	versicolor
## 68	5.8	2.7	4.1	1.0	versicolor
## 69	6.2	2.2	4.5	1.5	versicolor
## 70	5.6	2.5	3.9	1.1	versicolor
## 71	5.9	3.2	4.8	1.8	versicolor
## 72	6.1	2.8	4.0	1.3	versicolor
## 73	6.3	2.5	4.9	1.5	versicolor
## 74	6.1	2.8	4.7	1.2	versicolor
## 75	6.4	2.9	4.3	1.3	versicolor
## 76	6.6	3.0	4.4	1.4	versicolor
## 77	6.8	2.8	4.8	1.4	versicolor
## 78	6.7	3.0	5.0	1.7	versicolor
## 79	6.0	2.9	4.5	1.5	versicolor
## 80	5.7	2.6	3.5	1.0	versicolor
## 81	5.5	2.4	3.8	1.1	versicolor
## 82	5.5	2.4	3.7	1.0	versicolor
## 83	5.8	2.7	3.9	1.2	versicolor
## 84	6.0	2.7	5.1	1.6	versicolor
## 85	5.4	3.0	4.5	1.5	versicolor
## 86	6.0	3.4	4.5	1.6	versicolor
## 87	6.7	3.1	4.7	1.5	versicolor
## 88	6.3	2.3	4.4	1.3	versicolor
## 89	5.6	3.0	4.1	1.3	versicolor
## 90	5.5	2.5	4.0	1.3	versicolor
## 91	5.5	2.6	4.4	1.2	versicolor
## 92	6.1	3.0	4.6	1.4	versicolor
## 93	5.8	2.6	4.0	1.2	versicolor

## 94	5.0	2.3	3.3	1.0	versicolor
## 95	5.6	2.7	4.2	1.3	versicolor
## 96	5.7	3.0	4.2	1.2	versicolor
## 97	5.7	2.9	4.2	1.3	versicolor
## 98	6.2	2.9	4.3	1.3	versicolor
## 99	5.1	2.5	3.0	1.1	versicolor
## 100	5.7	2.8	4.1	1.3	versicolor
## 101	6.3	3.3	6.0	2.5	virginica
## 102	5.8	2.7	5.1	1.9	virginica
## 103	7.1	3.0	5.9	2.1	virginica
## 104	6.3	2.9	5.6	1.8	virginica
## 105	6.5	3.0	5.8	2.2	virginica
## 106	7.6	3.0	6.6	2.1	virginica
## 107	4.9	2.5	4.5	1.7	virginica
## 108	7.3	2.9	6.3	1.8	virginica
## 109	6.7	2.5	5.8	1.8	virginica
## 110	7.2	3.6	6.1	2.5	virginica
## 111	6.5	3.2	5.1	2.0	virginica
## 112	6.4	2.7	5.3	1.9	virginica
## 113	6.8	3.0	5.5	2.1	virginica
## 114	5.7	2.5	5.0	2.0	virginica
## 115	5.8	2.8	5.1	2.4	virginica
## 116	6.4	3.2	5.3	2.3	virginica
## 117	6.5	3.0	5.5	1.8	virginica
## 118	7.7	3.8	6.7	2.2	virginica
## 119	7.7	2.6	6.9	2.3	virginica
## 120	6.0	2.2	5.0	1.5	virginica
## 121	6.9	3.2	5.7	2.3	virginica
## 122	5.6	2.8	4.9	2.0	virginica
## 123	7.7	2.8	6.7	2.0	virginica
## 124	6.3	2.7	4.9	1.8	virginica
## 125	6.7	3.3	5.7	2.1	virginica
## 126	7.2	3.2	6.0	1.8	virginica
## 127	6.2	2.8	4.8	1.8	virginica
## 128	6.1	3.0	4.9	1.8	virginica
## 129	6.4	2.8	5.6	2.1	virginica
## 130	7.2	3.0	5.8	1.6	virginica
## 131	7.4	2.8	6.1	1.9	virginica
## 132	7.9	3.8	6.4	2.0	virginica
## 133	6.4	2.8	5.6	2.2	virginica
## 134	6.3	2.8	5.1	1.5	virginica
## 135	6.1	2.6	5.6	1.4	virginica
## 136	7.7	3.0	6.1	2.3	virginica
## 137	6.3	3.4	5.6	2.4	virginica
## 138	6.4	3.1	5.5	1.8	virginica
## 139	6.0	3.0	4.8	1.8	virginica
## 140	6.9	3.1	5.4	2.1	virginica
## 141	6.7	3.1	5.6	2.4	virginica
## 142	6.9	3.1	5.1	2.3	virginica
## 143	5.8	2.7	5.1	1.9	virginica
## 144	6.8	3.2	5.9	2.3	virginica
## 145	6.7	3.3	5.7	2.5	virginica
## 146	6.7	3.0	5.2	2.3	virginica
## 147	6.3	2.5	5.0	1.9	virginica

```
## 148      6.5      3.0      5.2      2.0 virginica
## 149      6.2      3.4      5.4      2.3 virginica
## 150      5.9      3.0      5.1      1.8 virginica
```

```
# Estadística descriptiva -----
```

```
versicolor <- iris$Petal.Length[iris$Species == "versicolor"]
virginica <- iris$Petal.Length[iris$Species == "virginica"]
```

```
summary(versicolor)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      3.00   4.00   4.35   4.26   4.60   5.10
```

```
summary(virginica)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      4.500  5.100  5.550  5.552  5.875  6.900
```

```
mean(versicolor)
```

```
## [1] 4.26
```

```
sd(versicolor)
```

```
## [1] 0.469911
```

```
mean(virginica)
```

```
## [1] 5.552
```

```
sd(virginica)
```

```
## [1] 0.5518947
```

```
# Prueba estadística -----
```

```
# Pregunta de investigación:
```

```
# ¿Existe una diferencia significativa en la longitud de los pétalos (Petal.Length)  
# entre Iris versicolor e Iris virginica?
```

```
# Hipótesis estadísticas para una prueba t de dos muestras "two.sided":
```

```
# - H0 (nula): No existen diferencias significativas entre la longitud de los pétalos de  
# Iris versicolor e Iris virginica.
```

```
# - H1 (alternativa): Existen diferencias significativas entre la longitud de los pétalos  
# de Iris versicolor e Iris virginica.
```

```
# Ejecutar la prueba en R
```

```
# Homogeneidad de varianzas
```

```
var.test(versicolor, virginica)
```

```
##
## F test to compare two variances
##
## data: versicolor and virginica
## 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

# Prueba t de dos muestras independientes (Clasica)

# p-value = 0.2637 > 0.05. Se utilizo prueba de t clasica dado que no hay evidencia
# estadística suficiente para decir que las varianzas son distintas (var.equal = T).

t.test(versicolor, virginica, alternative = "two.sided", var.equal = T)

##
## Two Sample t-test
##
## data: versicolor and virginica
## t = -12.604, df = 98, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -1.495426 -1.088574
## sample estimates:
## mean of x mean of y
##      4.260      5.552

# Efecto de Cohen

library(effsize)
cohen.d(versicolor, virginica)

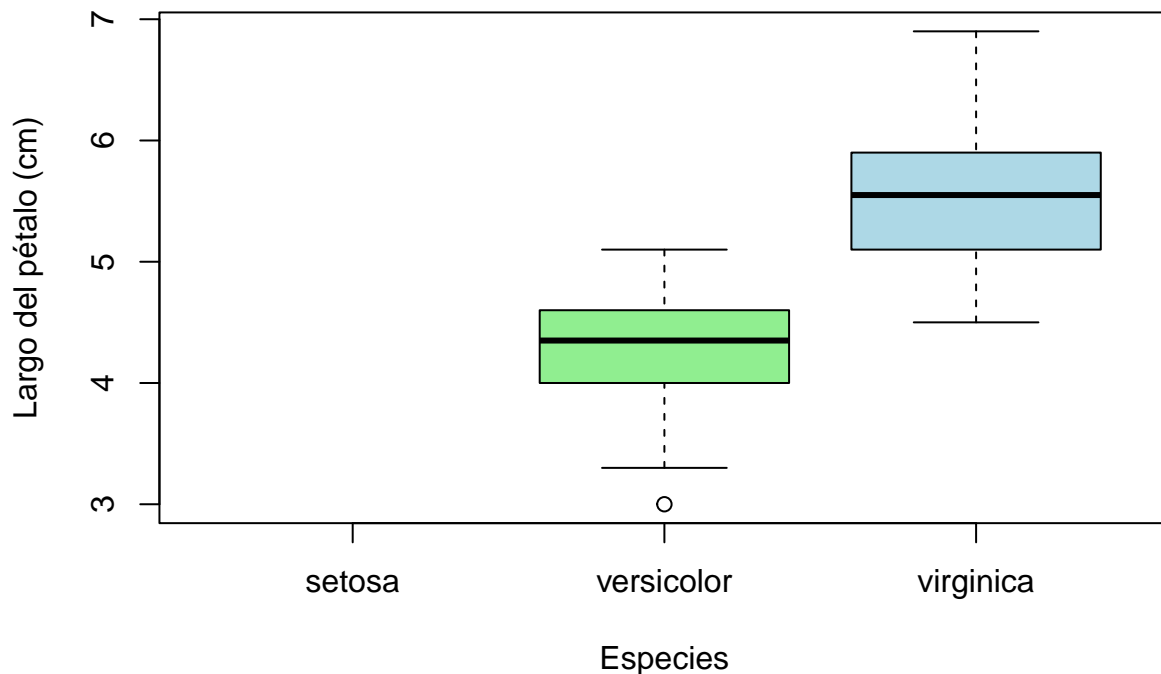
##
## Cohen's d
##
## d estimate: -2.520756 (large)
## 95 percent confidence interval:
##      lower      upper
## -3.052397 -1.989115

# El efecto de Cohen es de ~ -2.5, lo que corresponde a un efecto muy grande.

# Visualizacion -----

boxplot(Petal.Length ~ Species, data = iris_sub,
        col = c("lightblue", "lightgreen"),
        main = "Comparación de Petal.Length entre Iris versicolor e Iris virginica",
        ylab = "Largo del pétalo (cm)",
        xlab = "Especies")
```

Comparación de Petal.Length entre Iris versicolor e Iris virginica



Informe escrito -----

El presente análisis tuvo como objetivo determinar si existe una diferencia significativa en la longitud de los pétalos (Petal.Length) entre las especies Iris versicolor e Iris virginica, utilizando la base de datos iris en R. La pregunta central fue: ¿las dos especies presentan diferencias en la media de la longitud de los pétalos? Para responderla se plantearon las hipótesis: la hipótesis nula (H_0) establece que las medias son iguales, mientras que la hipótesis alternativa (H_1) sostiene que existen diferencias entre ambas.

Como primer paso se verificó la homogeneidad de varianzas mediante una prueba F, obteniendo un estadístico $F = 0.72497$ y un valor $p = 0.2637$. Dado que el valor p es mayor que 0.05, no se rechaza la hipótesis nula de igualdad de varianzas, lo que permite aplicar la prueba t de Student clásica para comparar medias bajo el supuesto de varianzas iguales.

La prueba t arrojó un valor $t = -12.604$ con un $p\text{-value} < 2.2e-16$, lo que indica que se rechaza la hipótesis nula. Esto significa que existen diferencias estadísticamente significativas en la longitud de los pétalos entre versicolor y virginica. La diferencia de medias fue aproximadamente -1.29, es decir, los pétalos de versicolor resultaron ser más cortos que los de virginica. Además, el tamaño del efecto medido con Cohen's d fue cercano a -2.5, lo cual representa una diferencia extremadamente grande en magnitud.

Los resultados gráficos a través de un boxplot comparativo, muestran visualmente que la distribución de la longitud de los pétalos de virginica se ubica en valores

considerablemente más altos que los de versicolor, reforzando la evidencia estadística.

Desde el punto de vista estadístico, el análisis demuestra que las medias difieren
de manera muy significativa y que esta diferencia es de gran magnitud. Biológicamente,
la longitud de los pétalos constituye un rasgo morfológico fundamental para distinguir
entre ambas especies. Virginica presenta pétalos mucho más largos que versicolor,
lo cual no solo valida su diferenciación taxonómica, sino que también resalta la
importancia de este carácter floral en estudios de clasificación y ecología de plantas.