

Laboratorio 2

Usuario: Valeria Rubí Luna Bazaldúa

2025-11-19

```
#####
#####  
#Base de datos iris  
#Tarea: Analisis de datos iris  
# Laboatorio 2  
# Curso: 2025- Metodos estadisticos  
# Valeria Rubí Luna Bazaldúa  
#####  
#####  
# EXPORTAR DATOS IRIS  
# Ver las primeras filas de la base de datos 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 de todas las variables  
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"
```

```

#2. SELECCIONDEESPECIES
#Filtrarversicolorvirginica
data_sub<-subset(iris,Species%in%c("versicolor","virginica"))
#Conteo de especies
table(data_sub$Species)

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

#####
#####
#3. ESTADISTICAS DESCRIPTIVAS
#Calcular medias y desviación de Petal.Length por especie
aggregate(Petal.Length~Species,data=data_sub,
          FUN=function(x)c(media=mean(x),sd=sd(x),
                           min=min(x),max=max(x)))

##      Species Petal.Length.media Petal.Length.sd Petal.Length.min
## 1 versicolor      4.2600000    0.4699110     3.0000000
## 2 virginica       5.5520000    0.5518947     4.5000000
##   Petal.Length.max
## 1      5.100000
## 2      6.900000

#####
#####

#####
#####

#PRUEBA DE HIPÓTESIS
#Pregunta de investigación:
#¿Existen diferencias en el largo del pétalo (Petal.Length)
#entre versicolor y virginica?
#Hipótesis
#H0: mu_versicolor = mu_virginica
#H1: mu_versicolor != mu_virginica
#Probar igualdad de varianzas
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

```

```

#PruebatconWelchpordefectosivarianzasdistintas
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
#####
######TAMAÑODELEFFECTOCOHEN'SD

#CreamosvectoresconlavariableLengthparacadaespecie
#paratrabajarconlosdatosdecadagrupoefacilmentedemaneradirecta
versicolor<-data_sub$Petal.Length[data_sub$Species=="versicolor"]
virginica<-data_sub$Sepal.Length[data_sub$Species=="virginica"]

#Calculamosladiferenciaentrelasmediasdelasdosespecies
#parateneraideadecuantomayoresellargodelpetalovirginica
mean_diff<-mean(virginica)-mean(versicolor)

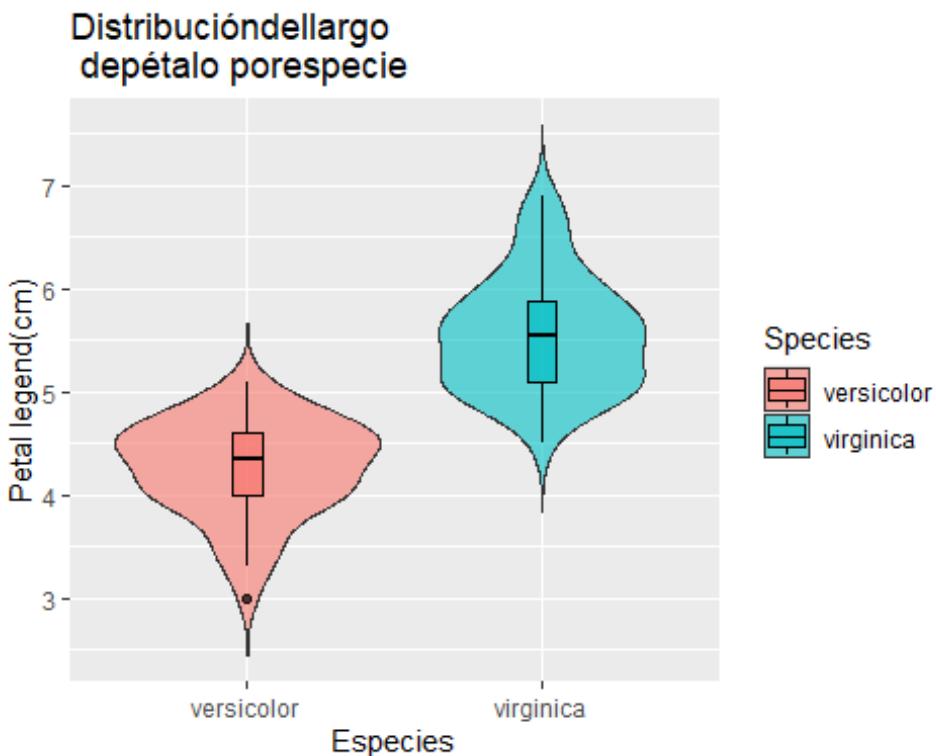
#Calculamosladesviacionestandardcombinadadeambosgrupospara
#estandarizarladiferenciademediasCohen'sd
pooled_sd<-sqrt(((sd(virginica))^2+(sd(versicolor))^2)/2)

#Calculodeltamañodelefectocohen'sdqueindicalamagnitudde
#diferenciasentrelaspecies
#Valores:0.2(pequeño),0.5(mediano),0.8(grande)
cohen.d<-mean_diff/pooled_sd
cohen.d

## [1] 4.163923
#####
######GRAFICOGGPLOT
library(ggplot2)
ggplot(data_sub,aes(x=Species,y=Petal.Length,fill=Species))+geom_violin(trim=FALSE,alpha=0.6)+geom_boxplot(width=0.1,
color=

```

```
"black",alpha=0.7)+labs(title="Distribución dellargo  
depétalo porespecie",y="Petal legend(cm)",x="Especies")
```



```
#####
#####
```