

# Bootstrap aplicado

Wladimir Aldo Carlosviza Amanqui

September 16, 2025

## Bootstrap aplicado a la base *iris*

```
# =====
# Bootstrap aplicado a iris
# =====

# Usamos la base de datos iris
data(iris)

set.seed(123) # Reproducibilidad

# Número de repeticiones bootstrap
B <- 1000

# Variables numéricas de iris
vars <- names(iris)[1:4]

# Lista para guardar resultados
resultados <- list()

for (var in vars) {
  n <- nrow(iris)
  valores_boot <- numeric(B)

  for (b in 1:B) {
    # Muestreo con reemplazo
    muestra <- iris[sample(1:n, size = n, replace = TRUE), ]

    # Guardamos la media bootstrap de la variable actual
    valores_boot[b] <- mean(muestra[[var]])
  }

  # Guardamos resultados
  resultados[[var]] <- list(
    media_original = mean(iris[[var]]),
    media_bootstrap = mean(valores_boot),
    IC_95 = quantile(valores_boot, c(0.025, 0.975)),
    distribucion = valores_boot
  )
}

# =====
# Mostrar resultados
```

```

# =====
for (var in vars) {
  cat("\n-----\n")
  cat("Variable:", var, "\n")
  cat("Media_original:", resultados[[var]]$media_original, "\n")
  cat("Media_bootstrap:", resultados[[var]]$media_bootstrap, "\n")
  cat("IC_95%:", resultados[[var]]$IC_95, "\n")
}
## -----
## Variable: Sepal.Length
## Media original: 5.843333
## Media bootstrap: 5.841857
## IC 95%: 5.714 5.978717
##
## -----
## Variable: Sepal.Width
## Media original: 3.057333
## Media bootstrap: 3.056507
## IC 95%: 2.986667 3.12735
##
## -----
## Variable: Petal.Length
## Media original: 3.758
## Media bootstrap: 3.766603
## IC 95%: 3.48785 4.053383
##
## -----
## Variable: Petal.Width
## Media original: 1.199333
## Media bootstrap: 1.197323
## IC 95%: 1.075317 1.320683
# =====
# Graficos de densidad
# =====
par(mfrow = c(2,2)) # Panel 2x2
for (var in vars) {
  densidad <- density(resultados[[var]]$distribucion)
  plot(densidad, main = paste("Bootstrap_de", var),
    xlab = "Media_bootstrap", col = "blue", lwd = 2)
  abline(v = resultados[[var]]$media_original, col = "red", lwd = 2, lty = 2)
  legend("topright", legend = c("Distribucion_bootstrap", "Media_original"),
    col = c("blue", "red"), lwd = 2, lty = c(1,2), bty = "n")
}

```

## Salida

```

-----
Variable: Sepal.Length
Media original: 5.843333
Media bootstrap: 5.841857
IC 95%: 5.714 5.978717

-----
Variable: Sepal.Width

```

```

Media original: 3.057333
Media bootstrap: 3.056507
IC 95%: 2.986667 3.12735

```

```

-----
Variable: Petal.Length
Media original: 3.758
Media bootstrap: 3.766603
IC 95%: 3.48785 4.053383

```

```

-----
Variable: Petal.Width
Media original: 1.199333
Media bootstrap: 1.197323
IC 95%: 1.075317 1.320683

```

## Tabla resumen de resultados

Variable	Media original	Media bootstrap	IC 95%
Sepal.Length	5.843333	5.841857	[5.714, 5.978717]
Sepal.Width	3.057333	3.056507	[2.986667, 3.12735]
Petal.Length	3.758000	3.766603	[3.48785, 4.053383]
Petal.Width	1.199333	1.197323	[1.075317, 1.320683]

Table 1: Medias originales, medias bootstrap y intervalos de confianza al 95% obtenidos mediante  $B = 1000$  réplicas.

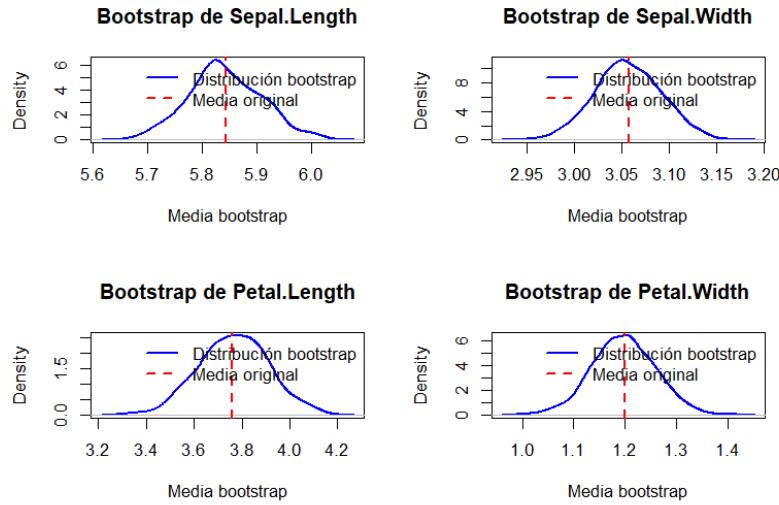


Figure 1: Distribución bootstrap de la media de Sepal.Length

## **Interpretación**

El bootstrap con  $B = 1000$  réplicas confirma que las medias de las cuatro variables numéricas de *iris* son estables y que las estimaciones bootstrap están muy próximas a las medias originales. Los intervalos de confianza al 95% son relativamente estrechos para las medidas del sépalo (Sepal.Length y Sepal.Width), y algo más amplios para las medidas del pétalo (Petal.Length y Petal.Width), lo que sugiere mayor variabilidad en las últimas. En conjunto, estos resultados apoyan la idea de que el método bootstrap proporciona estimaciones robustas de la incertidumbre de las medias sin requerir supuestos fuertes de normalidad.