

Actividad 7: Intervalos de confianza

Daniela Jiménez Téllez

2024-08-21

Problema 1

Muestra que el nivel de confianza indica el porcentaje de intervalos de confianza extraídos de una misma población que contienen a la verdadera media a través de la simulación de intervalos:

- Haz la simulación de 150 muestras de tamaño 150 extraídas de una población normal con $\mu = 70$ y $\sigma = 9$

```
set.seed(85)

miu <- 70
sigma <- 9
n <- 150

muestras <- replicate(150, rnorm(n, mean = miu, sd = sigma))
```

- Calcula el intervalo con un nivel de confianza del 97% para cada una de esas medias. Obtendrás 150 intervalos de confianza.

```
alpha <- 0.03
n <- 150

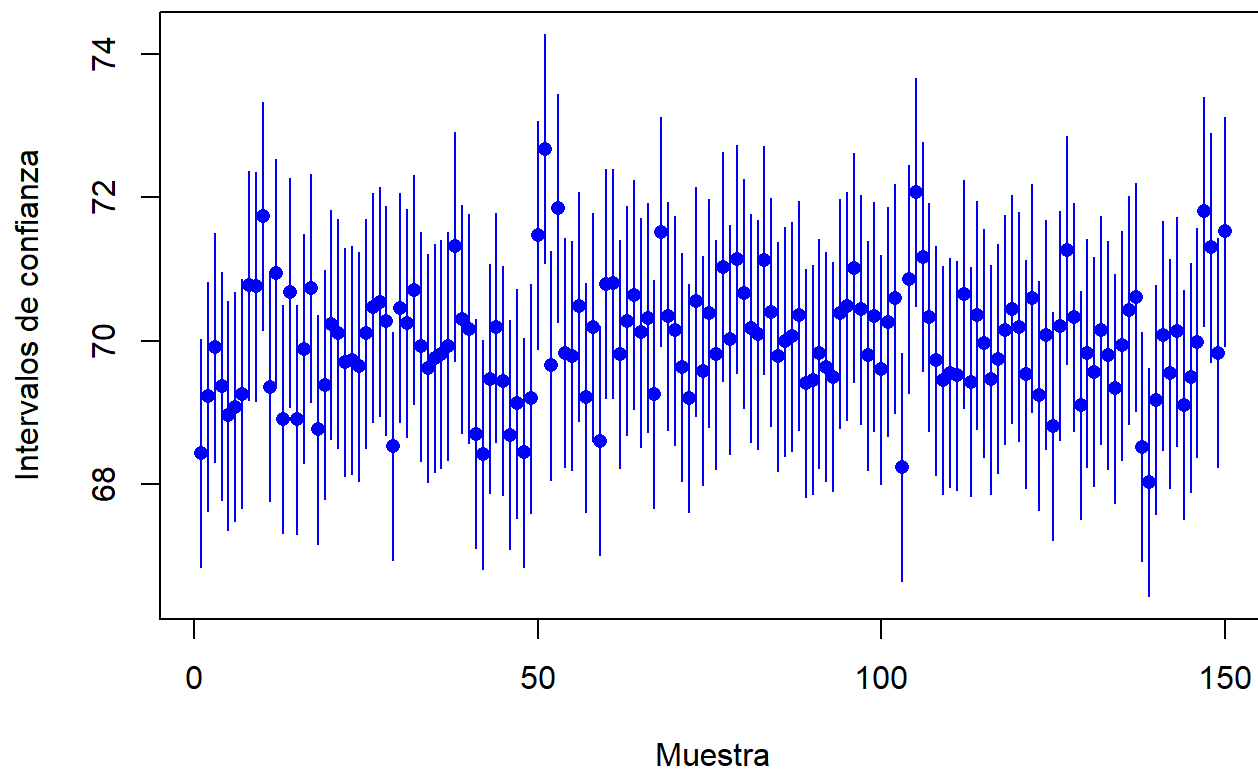
E1 <- abs(qnorm(alpha/2)) * sigma/sqrt(n)

medias_muest <- colMeans(muestras)

A1 <- medias_muest - E1
B1 <- medias_muest + E1
```

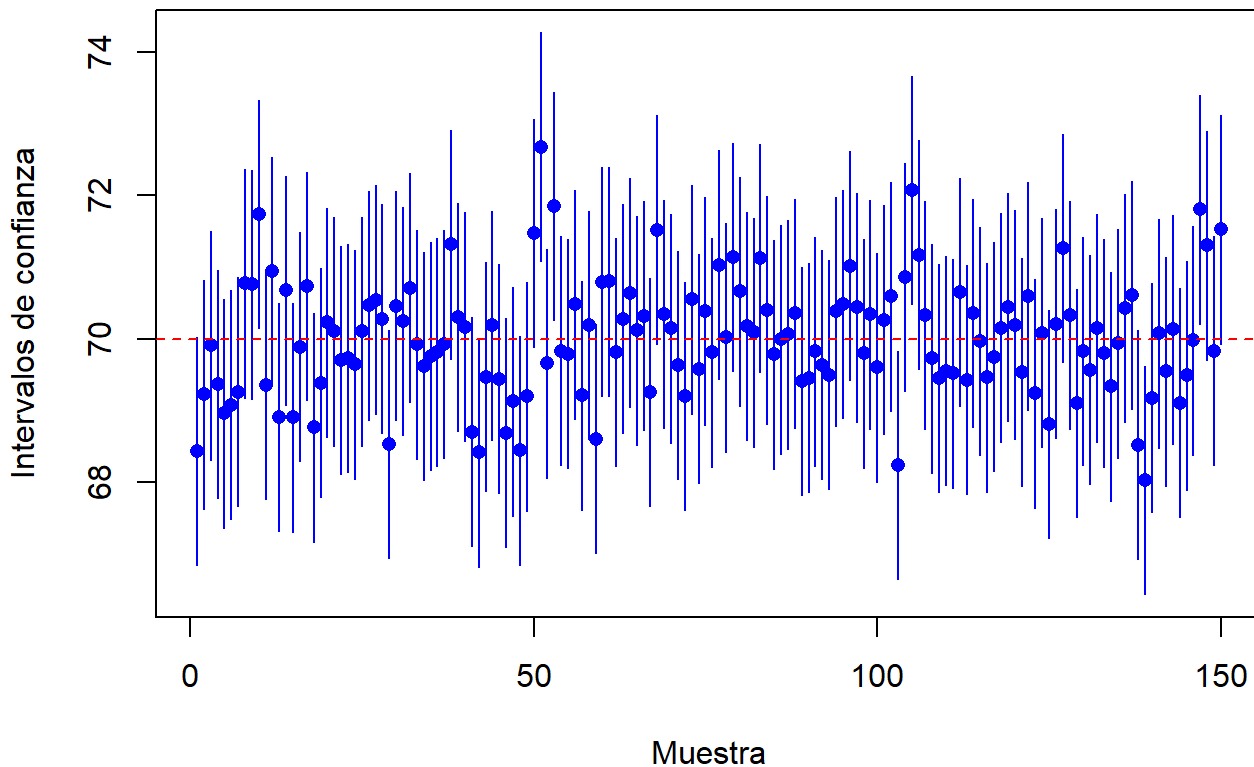
- Grafica los 150 intervalos de confianza

```
plot(1:150, medias_muest, ylim = c(min(A1), max(B1)), pch = 16, col = "blue", xlab = "Muestra",
ylab = "Intervalos de confianza")
segments(1:150, A1, 1:150, B1, col = "blue")
```



- Grafica la media poblacional ($\mu=70$) como una linea horizontal

```
plot(1:150, medias_muest, ylim = c(min(A1), max(B1)), pch = 16, col = "blue", xlab = "Muestra",  
ylab = "Intervalos de confianza")  
segments(1:150, A1, 1:150, B1, col = "blue")  
abline(h = miu, col = "red", lty = 2)
```



- Cuenta cuántos intervalos de confianza contienen a la verdadera media, ¿qué porcentaje representan?

```

tienen_media <- sum(A1 <= miu & B1 >= miu)
porcentaje <- (tienen_media / 150) * 100

```

```

cat("La cantidad de intervalos que tienen la verdadera media son: ", tienen_media, "\n")

```

```

## La cantidad de intervalos que tienen la verdadera media son: 143

```

```

cat("El porcentaje que representan es: ", porcentaje, "%")

```

```

## El porcentaje que representan es: 95.33333 %

```

Problema 2

Resuelve las dos partes del problema “El misterioso Helio”.

Primera parte. Suponga que la porosidad al helio (en porcentaje) de muestras de carbón, tomadas de cualquier veta en particular, está normalmente distribuida con una desviación estándar verdadera de 0.75. Se sabe que 10 años atrás la porosidad media de helio en la veta era de 5.3 y se tiene interés en saber si actualmente ha disminuido. Se toma una muestra al azar de 20 especímenes y su promedio resulta de 4.85.

- Haga una estimación por intervalo con una confianza del 97% para el promedio de porosidad para evaluar si ha disminuido.

```

sigma = 0.75
alpha = 0.03
xb1 = 4.85
n1 = 20

E1 = abs(qnorm(alpha/2)) * sigma/sqrt(n1)
A1 = xb1 - E1
B1 = xb1 + E1

cat("La verdadera media se encuentra entre:", A1, "y", B1)

```

```
## La verdadera media se encuentra entre: 4.486065 y 5.213935
```

- Se toma otra muestra de tamaño 16. El promedio de la muestra fue de 4.56. Calcule el intervalo de confianza al 97% de confianza

```

sigma = 0.75
alpha = 0.03
xb2 = 4.56
n2 = 16

E2 = abs(qnorm(alpha/2)) * sigma/sqrt(n2)
A2 = xb2 - E2
B2 = xb2 + E2

cat("La verdadera media se encuentra entre:", A2, "y", B2)

```

```
## La verdadera media se encuentra entre: 4.153108 y 4.966892
```

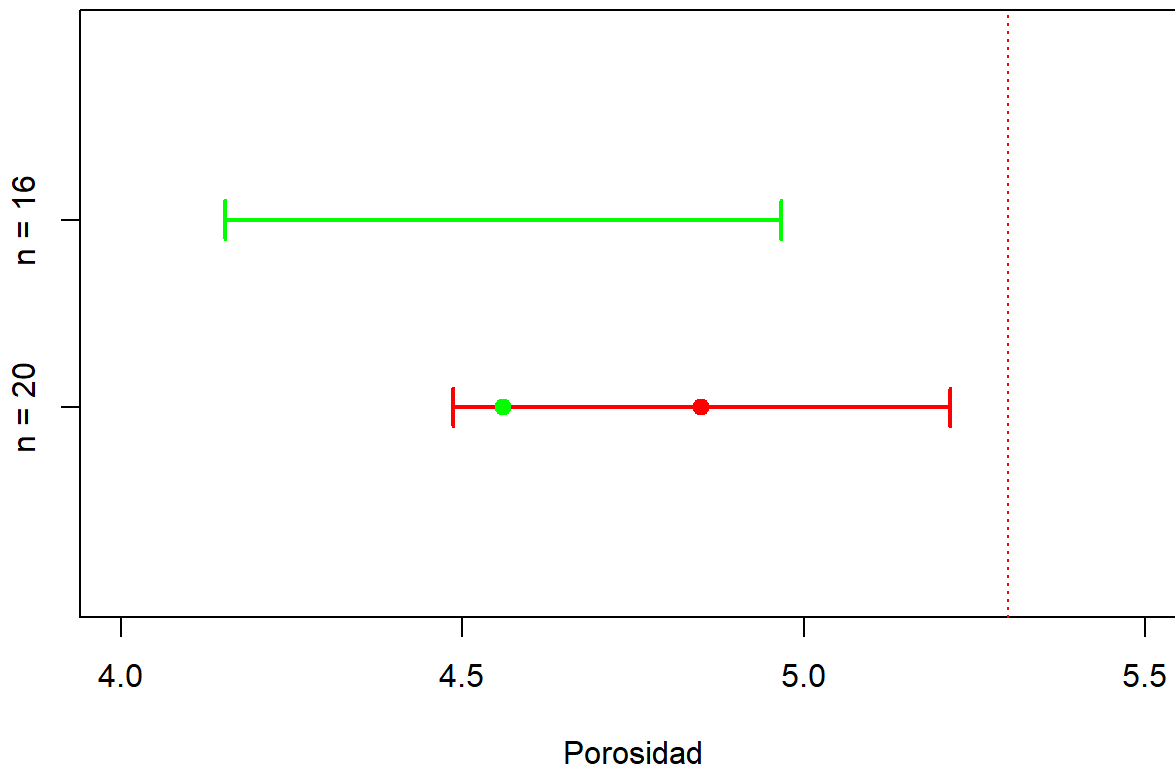
- ¿Podemos afirmar que la porosidad del helio ha disminuido?

```

plot(0, ylim = c(0, 2 + 1), xlim = c(4, 5.5), yaxt = "n", ylab = "", xlab = "Porosidad")
axis(2, at = c(1, 2), labels = c("n = 20", "n = 16"))

arrows(A1, 1, B1, 1, angle = 90, code = 3, length = 0.1, lwd = 2, col = "red")
arrows(A2, 2, B2, 2, angle = 90, code = 3, length = 0.1, lwd = 2, col = "green")
points(xb1, 1, pch = 19, cex = 1.1, col = "red")
points(xb2, 1, pch = 19, cex = 1.1, col = "green")
abline(v = 5.3, lty = 3, col = "red")

```



Con esta gráfica podemos observar que la porosidad del helio no ha disminuido. Al contrario, a mi parecer se puede ver que ha aumentado un poco por la manera en la que están los intervalos en la gráfica.

Segunda parte. Suponga que la porosidad al helio (en porcentaje) de muestras de carbón, tomadas de cualquier veta en particular, está normalmente distribuida con una desviación estándar verdadera de 0.75.

- ¿Qué tan grande tiene que ser el tamaño de la muestra si se desea que el ancho del intervalo con un 95% de confianza no sobrepase de 0.4?

```
sigma = 0.75
alpha = 0.05
W1 = 0.4

Z1 <- abs(qnorm(alpha / 2))
n1 <- (Z1 * sigma / (W1 / 2))^2
n1 <- ceiling(n1)

cat("El tamaño de muestra necesario es:", n1)
```

```
## El tamaño de muestra necesario es: 55
```

- ¿Qué tamaño de muestra necesita para estimar la porosidad promedio verdadera dentro de 0.2 unidades alrededor de la media muestral con una confianza de 99%?

```
sigma = 0.75
alpha = 0.01
E = 0.2

Z2 <- abs(qnorm(alpha / 2))
n2 <- (Z2 * sigma / (E))^2
n2 <- ceiling(n2)

cat("El tamaño de muestra necesario es:", n2)
```

```
## El tamaño de muestra necesario es: 94
```

Problema 3

- Con el archivo de datos de “El Marcapasos” haz los intervalos de confianza para la media de las siguientes variables:
 - Intensidad de pulsos con y sin Marcapasos (2 intervalos de confianza)
 - Periodo entre pulso con y sin Marcapasos (2 intervalos de confianza)

```
data <- read.csv("El marcapasos.csv")

con_mp <- subset(data, Marcapasos == "Con MP")
sin_mp <- subset(data, Marcapasos == "Sin MP")

intensidad_con <- con_mp$Intensidad.de.pulso
intensidad_sin <- sin_mp$Intensidad.de.pulso

periodo_con <- con_mp$Periodo.entre.pulsos
periodo_sin <- sin_mp$Periodo.entre.pulsos

alpha <- 0.05

# Intensidad

## Con

sigma1 <- sd(intensidad_con)
n1 <- length(intensidad_con)

E1 <- abs(qnorm(alpha/2)) * sigma1/sqrt(n1)

medias_muest1 <- mean(intensidad_con)

A1 <- medias_muest1 - E1
B1 <- medias_muest1 + E1

cat("La verdadera media de la intensidad con marcapasos se encuentra entre:", A1, "y", B1, "\n")
```

```
## La verdadera media de la intensidad con marcapasos se encuentra entre: 0.1645811 y 0.2273013
```

```
## Sin

sigma2 <- sd(intensidad_sin)
n2 <- length(intensidad_sin)

E2 <- abs(qnorm(alpha/2)) * sigma2/sqrt(n2)

medias_muest2 <- mean(intensidad_sin)

A2 <- medias_muest2 - E2
B2 <- medias_muest2 + E2

cat("La verdadera media de la intensidad sin marcapasos se encuentra entre:", A2, "y", B2, "\n")
```

```
## La verdadera media de la intensidad sin marcapasos se encuentra entre: 0.1708292 y 0.2433669
```

```
# Periodo

## Con

sigma3 <- sd(periodo_con)
n3 <- length(periodo_con)

E3 <- abs(qnorm(alpha/2)) * sigma3/sqrt(n3)

medias_muest3 <- mean(periodo_con)

A3 <- medias_muest3 - E3
B3 <- medias_muest3 + E3

cat("La verdadera media del periodo con marcapasos se encuentra entre:", A3, "y", B3, "\n")
```

```
## La verdadera media del periodo con marcapasos se encuentra entre: 0.8644566 y 0.9178964
```

```
## Sin

sigma4 <- sd(periodo_sin)
n4 <- length(periodo_sin)

E4 <- abs(qnorm(alpha/2)) * sigma4/sqrt(n4)

medias_muest4 <- mean(periodo_sin)

A4 <- medias_muest4 - E4
B4 <- medias_muest4 + E4

cat("La verdadera media del periodo sin marcapasos se encuentra entre:", A4, "y", B4, "\n")
```

```
## La verdadera media del periodo sin marcapasos se encuentra entre: 1.005521 y 1.218009
```

- Grafica los intervalos de confianza obtenidos en “El marcapasos”:
 - Grafica en un mismo eje coordenado la intensidad de pulso con y sin marcapasos
 - Grafica en un mismo eje coordenado el periodo entre pulso con y sin marcapasos

```
# Intensidad
```

```
plot(c(1, 2), c(medias_muest1, medias_muest2), ylim = c(min(A1, A2), max(B1, B2)),
     xaxt = "n", pch = 16, col = "blue",
     xlab = "Condición", ylab = "Intensidad de Pulso",
     main = "Intervalos de Confianza de la Intensidad de Pulso")
axis(1, at = c(1, 2), labels = c("Con MP", "Sin MP"))
arrows(1, A1, 1, B1, angle = 90, code = 3, length = 0.1, col = "blue")
arrows(2, A2, 2, B2, angle = 90, code = 3, length = 0.1, col = "blue")
```

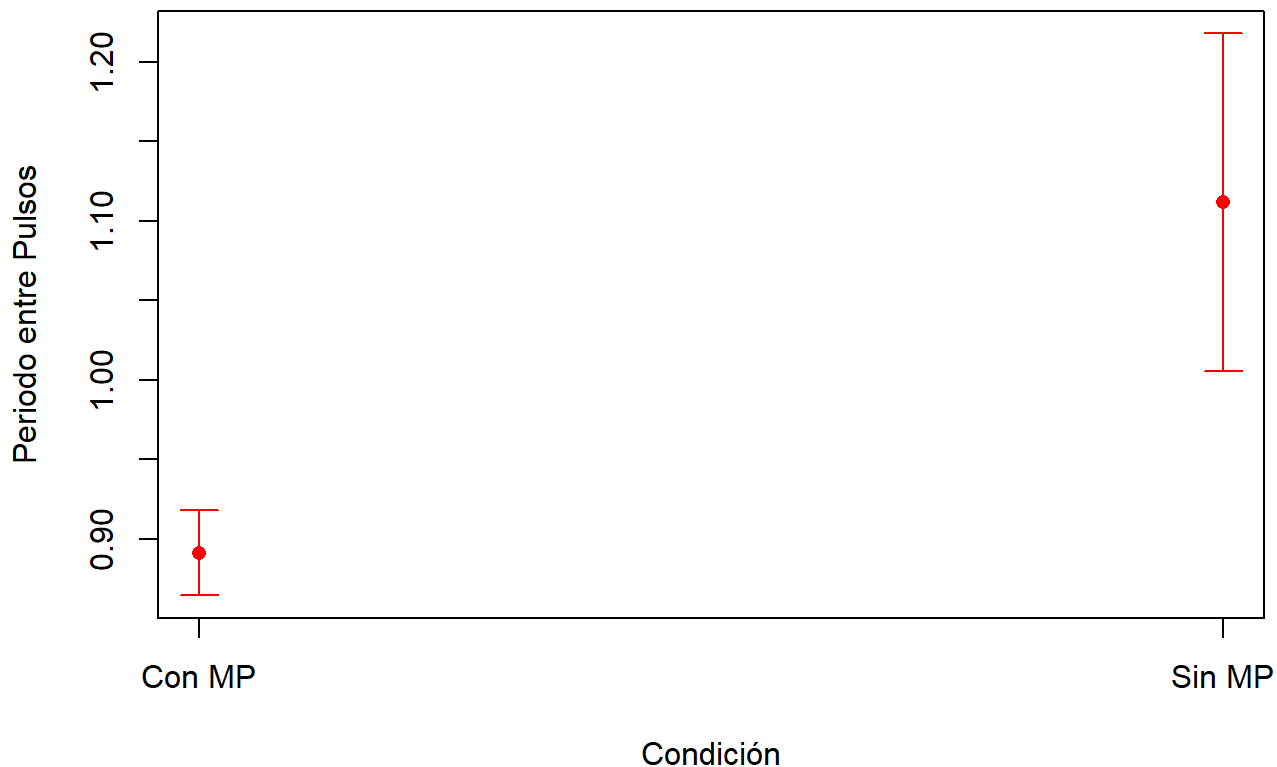
Intervalos de Confianza de la Intensidad de Pulso




```
# Periodo
```

```
plot(c(1, 2), c(medias_muest3, medias_muest4), ylim = c(min(A3, A4), max(B3, B4)),
     xaxt = "n", pch = 16, col = "red",
     xlab = "Condición", ylab = "Periodo entre Pulsos",
     main = "Intervalos de Confianza del Periodo entre Pulsos")
axis(1, at = c(1, 2), labels = c("Con MP", "Sin MP"))
arrows(1, A3, 1, B3, angle = 90, code = 3, length = 0.1, col = "red")
arrows(2, A4, 2, B4, angle = 90, code = 3, length = 0.1, col = "red")
```

Intervalos de Confianza del Periodo entre Pulsos



- Compara los intervalos obtenidos e interpreta los gráficos. Concluye sobre ambas variables en la presencia y ausencia de marcapasos

En el caso de intensidad podemos ver que en ambos intervalos de confianza se tienen resultados muy parecidos que hasta cierto punto se empalman. Esto nos dice que realmente no hay mucha diferencia entre la intensidad de pulso de una persona que usa o no marcapasos. Esto se puede observar mejor en la gráfica, ya que justamente podemos ver que hay muchos valores que comparten dentro del intervalo, lo que nos reafirma que no hay gran diferencia entre pacientes con o sin marcapasos.

Por otro lado, en el caso del pulso podemos ver que hay una diferencia más evidente en cuanto a los intervalos, que igual está reflejada en la gráfica, donde observamos que el de con MP es más angosto y pequeño que el de Sin MP, el cual es más amplio. Esto nos podría decir que el marcapasos puede ayudar a las personas a “regularizar” el periodo entre pulsos, a comparación de las personas que no lo usan.