

Actividad 7 Intervalos de confianza

Adrian Pineda Sanchez

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.

A. Para ello haz el gráfico de 150 intervalos de confianza obtenidos de la misma población. Guiáte de los siguientes pasos:

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

```
library(plotrix)
```

```
# Parámetros
```

```
m <- 150 # Número de muestras
```

```
n <- 150 # Tamaño de cada muestra
```

```
mu <- 70 # Media de la población
```

```
sigma <- 9 # Desviación estándar de la población
```

```
alpha <- 0.03 # Nivel de significancia (para un intervalo de confianza del 97%)
```

```
# Simulación de las medias muestrales
```

```
xb <- rnorm(m, mean=mu, sd=sigma/sqrt(n))
```

```
xb
```

```
## [1] 70.76682 70.00297 69.52106 70.79580 69.90365 71.06032 69.68582  
69.07225
```

```
## [9] 69.80064 70.25841 70.79013 69.73527 69.39051 70.23141 69.80106  
69.30000
```

```
## [17] 69.43578 69.38162 71.30968 70.94397 70.44855 70.73640 69.29612  
71.36276
```

```
## [25] 70.03545 70.44334 70.13418 70.36910 70.29759 69.59783 69.96080  
69.19886
```

```
## [33] 69.17841 69.33435 68.57479 70.15821 68.45096 68.97747 71.18869  
69.19460
```

```
## [41] 69.25540 70.48927 70.32639 70.72487 70.44571 69.28919 70.31163  
69.39032
```

```
## [49] 70.10164 70.23138 70.68628 68.65796 70.34538 69.28863 70.25941
```

```

70.89597
## [57] 70.88621 70.11696 68.80317 69.84112 69.54614 69.10011 69.56775
71.09610
## [65] 69.33821 70.67741 68.44909 69.35820 70.58785 69.65599 70.42051
69.68026
## [73] 69.30333 70.87482 70.23408 69.16401 70.46068 69.78766 70.95140
70.00157
## [81] 71.50795 68.25418 70.22867 70.74378 69.91674 69.10784 69.89763
69.76524
## [89] 71.49714 69.82359 71.09272 69.68851 70.74335 70.46514 70.72955
70.34582
## [97] 69.78114 69.63915 69.75323 70.61328 69.68781 70.63925 69.57806
70.36304
## [105] 69.79652 70.11128 68.75542 70.14275 70.63194 70.85951 69.69825
68.92525
## [113] 69.99025 70.26057 70.40842 70.64544 69.85198 69.64399 71.55874
71.39641
## [121] 70.85852 70.29126 70.35684 69.29003 71.27741 70.12294 70.00159
69.97804
## [129] 69.22970 69.50532 70.51672 70.59162 69.87167 69.60496 70.09384
69.69854
## [137] 71.22443 70.26061 70.06565 70.34387 68.38502 70.36603 70.26048
69.60746
## [145] 70.64625 70.47244 70.28509 69.72358 68.61463 71.14793

```

Calcula el margen de error con un nivel de confianza del 97% para esas medias.

```
# Cálculo del margen de error
```

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

```
cat("el margen de error con un nivel de confianza del 97% para esas medias es de:",E)
```

```
## el margen de error con un nivel de confianza del 97% para esas medias es de: 1.594684
```

Grafica las medias con el margen de error calculado en el inciso B. El comando plotCI calcula automáticamente los intervalos de confianza si le proporcionas las medias y el margen de error. Grafica la media poblacional ($\mu = 70$) como una línea horizontal

```
# Graficar los intervalos de confianza
```

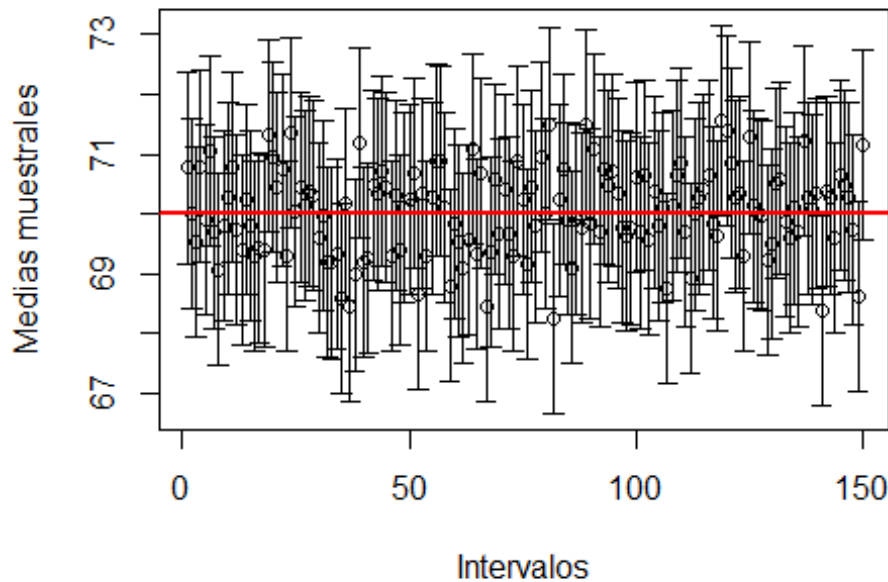
```
plotCI(1:m, xb, E, main="Gráfico de 150 Intervalos de Confianza (97%)",
```

```
      xlab="Intervalos", ylab="Medias muestrales",
```

```
      ylim=c(min(xb-E), max(xb+E)))
```

```
abline(h=mu, col="red", lwd=2)
```

Gráfico de 150 Intervalos de Confianza (97%)



B. Cuenta cuántos intervalos de confianza contienen a la verdadera media, ¿qué porcentaje representan? (Si quieres tener los 150 intervalos de confianza, calcula $\bar{x} - E$ para las cotas inferiores y $\bar{x} + E$ para las cotas superiores de los intervalos. Esto puede ayudarte a contar de manera automática cuántos intervalos contienen a la verdadera media)

```
# Cálculo de Los límites inferiores y superiores de Los intervalos
limites_inferiores <- xb - E
limites_superiores <- xb + E

# Contar cuántos intervalos contienen La media verdadera
contiene_mu <- (limites_inferiores <= mu) & (limites_superiores >= mu)
num_intervals_containing_mu <- sum(contiene_mu)

# Calcular el porcentaje de intervalos que contienen La media verdadera
porcentaje_contiene_mu <- num_intervals_containing_mu / m * 100

# Imprimir el resultado
cat("Número de intervalos que contienen la verdadera media:",
    num_intervals_containing_mu, "\n")

## Número de intervalos que contienen la verdadera media: 148

cat("Porcentaje de intervalos que contienen la verdadera media:",
    porcentaje_contiene_mu, "%\n")
```

```
## Porcentaje de intervalos que contienen la verdadera media: 98.66667 %
```

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.

X: Porosidad helio

$$X \sim N(\mu = ?, \sigma = 0.75)$$

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

```
n = 20
sigma = 0.75
desv_est = sigma / sqrt(n)
alfa = 0.03
xb = 4.85

E = abs(qnorm(0.03/2))*desv_est
A = xb - E
B = xb + E

cat("La verdadera media esta entre:", A, "y", B)

## La verdadera media esta entre: 4.486065 y 5.213935
```

B. 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

```
n2 = 16
sigma = 0.75
desv_est = sigma / sqrt(n2)
alfa = 0.03
xb2 = 4.56

E2 = abs(qnorm(0.03/2))*desv_est
A2 = xb2 - E2
```

```
B2 = xb2 + E2
```

```
cat("La verdadera media esta entre:", A2, "y", B2)
```

```
## La verdadera media esta entre: 4.153108 y 4.966892
```

C. ¿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  
helio")
```

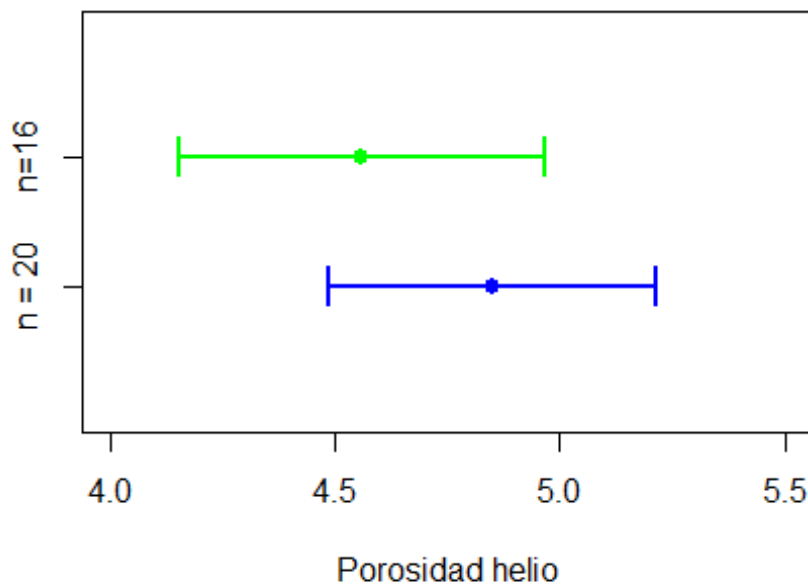
```
axis(2, at=c(1,2), labels=c("n = 20", "n=16"))
```

```
arrows(A, 1, B, 1, angle=90, code=3, length = 0.1, lwd = 2, col= "blue")
```

```
arrows(A2, 2, B2, 2, angle=90, code=3, length = 0.1, lwd = 2, col = "green")
```

```
points(xb, 1, pch=19, cex=1.1, col = "blue")
```

```
points(xb2, 2, pch=19, cex=1.1, col = "green")
```



Con los resultados actuales obtenidos para los intervalos de confianza al 97% para las muestras de tamaño 20 y 16, podemos realizar un análisis para determinar si la porosidad al helio ha disminuido en comparación con la media histórica de 5.3.

Para la muestra de tamaño 20 (n = 20): Media de la muestra: 4.85 Intervalo de confianza al 97%: [4.486065, 5.213935]

Para la muestra de tamaño 16 (n = 16): Media de la muestra: 4.56 Intervalo de confianza al 97%: [4.153108, 4.966892]

Esto indica que con un nivel de confianza del 97%, se puede afirmar que la porosidad al helio ha disminuido. Ambos intervalos de confianza sugieren que la verdadera media poblacional es menor que 5.3, lo que respalda la hipótesis de una disminución en la porosidad.

Por lo tanto, sí podemos afirmar que la porosidad del helio ha disminuido basándonos en estos resultados.

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.

A. ¿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
cant_desv_est = sigma / sqrt(n)
alfa = 0.05
xb = 4.85

# Observar que si el error es .4 entonces se reparte en 2, y de la ecuacion
original de E, obtenemos n

E = 0.2

n = (abs(qnorm(alfa/2))*sigma/(.2))^2

n

## [1] 54.02051

cat("El tamaño de la muestra debe ser minimo:",55,"si se desea que el ancho
del intervalo con un 95% de confianza no sobrepase de 0.4")

## El tamaño de la muestra debe ser minimo: 55 si se desea que el ancho del
intervalo con un 95% de confianza no sobrepase de 0.4
```

B. ¿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
cant_desv_est = sigma / sqrt(n)
alfa = 0.01
xb = 4.85

# Observar que si el error es .4 entonces se reparte en 2, y de la ecuacion
```

original de E, obtenemos n

```
E = 0.2
```

```
n = (abs(qnorm(alfa/2))*sigma/(.2))^2
```

```
n
```

```
## [1] 93.30323
```

```
cat("El tamaño de la muestra debe ser minimo:",94,"para estimar la porosidad  
promedio verdadera dentro de 0.2 unidades alrededor de la media muestral con  
una confianza de 99%")
```

```
## El tamaño de la muestra debe ser minimo: 94 para estimar la porosidad  
promedio verdadera dentro de 0.2 unidades alrededor de la media muestral con  
una confianza de 99%
```

Problema 3.

Con el archivo de datos de El Marcapasos Download El Marcapasos haz los intervalos de confianza para la media de las siguientes variables:

```
M = read.csv("El marcapasos.csv")
```

Intensidad de pulsos con y sin Marcapasos (2 intervalos de confianza)

Separar en dos DataFrames según la columna "Marcapasos"

```
i_sin_mp <- subset(M, Marcapasos == "Sin MP")[, "Intensidad.de.pulso",  
drop=FALSE]
```

```
i_con_mp <- subset(M, Marcapasos == "Con MP")[, "Intensidad.de.pulso",  
drop=FALSE]
```

Mostrar los primeros datos de cada DataFrame para verificar

```
head(i_sin_mp)
```

```
## Intensidad.de.pulso
```

```
## 1 0.131
```

```
## 2 0.303
```

```
## 3 0.297
```

```
## 4 0.416
```

```
## 5 0.585
```

```
## 6 0.126
```

```
head(i_con_mp)
```

```
## Intensidad.de.pulso
```

```
## 52 0.140
```

```
## 53 0.296
```

```
## 54 0.281
```

```
## 55 0.355
```

```
## 56          0.441
## 57          0.135

# Función para calcular el intervalo de confianza para la media usando qnorm
calcular_IC_media_qnorm <- function(variable, nivel_confianza = 0.95) {
  n <- length(variable)
  media <- mean(variable, na.rm = TRUE)
  cant_desv_estandar <- sd(variable, na.rm = TRUE) / sqrt(n)

  z_score <- qnorm(1 - (1 - nivel_confianza) / 2)
  error_margin <- z_score * cant_desv_estandar
  lower_bound <- media - error_margin
  upper_bound <- media + error_margin

  return(c(lower_bound, upper_bound))
}

# Calcular el intervalo de confianza para 'Intensidad.de.pulso' en el grupo
"Sin MP" usando qnorm
ic_sin_mp_qnorm <- calcular_IC_media_qnorm(i_sin_mp$Intensidad.de.pulso)

# Calcular el intervalo de confianza para 'Intensidad.de.pulso' en el grupo
"Con MP" usando qnorm
ic_con_mp_qnorm <- calcular_IC_media_qnorm(i_con_mp$Intensidad.de.pulso)

# Mostrar los resultados

cat("Intervalo de confianza de Intensidad para 'Con MP' usando qnorm:",
ic_con_mp_qnorm[1], "a", ic_con_mp_qnorm[2], "\n", "\n")

## Intervalo de confianza de Intensidad para 'Con MP' usando qnorm: 0.1645811
a 0.2273013
##

cat("Intervalo de confianza de Intensidad para 'Sin MP' usando qnorm:",
ic_sin_mp_qnorm[1], "a", ic_sin_mp_qnorm[2])

## Intervalo de confianza de Intensidad para 'Sin MP' usando qnorm: 0.1708292
a 0.2433669
```

Periodo entre pulso con y sin Marcapasos (2 intervalos de confianza)

```
# Separar en dos DataFrames según la columna "Marcapasos"
p_sin_mp <- subset(M, Marcapasos == "Sin MP")[, "Periodo.entre.pulsos",
drop=FALSE]
p_con_mp <- subset(M, Marcapasos == "Con MP")[, "Periodo.entre.pulsos",
drop=FALSE]

# Mostrar los primeros datos de cada DataFrame para verificar
head(p_sin_mp)
```



```
## Periodo.entre.pulsos
## 1 1.2
## 2 0.9
## 3 0.9
## 4 0.8
## 5 0.7
## 6 1.2

head(p_con_mp)

## Periodo.entre.pulsos
## 52 0.94
## 53 0.81
## 54 0.82
## 55 0.73
## 56 0.69
## 57 0.94

ic_p_sin_mp <- calcular_IC_media_qnorm(p_sin_mp$Periodo.entre.pulsos)

ic_p_con_mp <- calcular_IC_media_qnorm(p_con_mp $Periodo.entre.pulsos)

# Mostrar Los resultados

cat("Intervalo de confianza de periodo para 'Con MP' usando qnorm:",
ic_p_con_mp[1], "a", ic_p_con_mp[2], "\n", "\n")

## Intervalo de confianza de periodo para 'Con MP' usando qnorm: 0.8644566 a
0.9178964
##

cat("Intervalo de confianza de periodo para 'Sin MP' usando qnorm:",
ic_p_sin_mp [1], "a", ic_p_sin_mp [2])

## Intervalo de confianza de periodo para 'Sin MP' usando qnorm: 1.005521 a
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

```
# Configurar Los datos de Los intervalos de confianza
intensidad_con <- c(0.16458, 0.2273)
intensidad_sin <- c(0.17083, 0.24337)

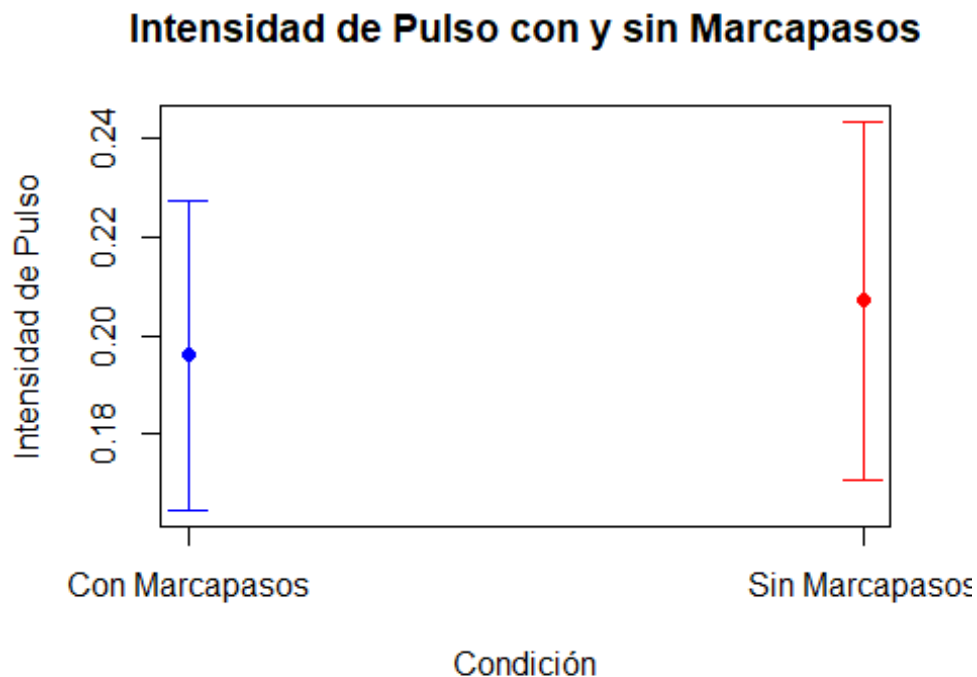
# Crear La gráfica
plot(1, type="n", xlim=c(1, 2), ylim=c(min(c(intensidad_con,
intensidad_sin)), max(c(intensidad_con, intensidad_sin))),
      xlab="Condición", ylab="Intensidad de Pulso", xaxt="n", main="Intensidad
```

```

de Pulso con y sin Marcapasos")
axis(1, at=1:2, labels=c("Con Marcapasos", "Sin Marcapasos"))

# Agregar Los intervalos de confianza
arrows(1, intensidad_con[1], 1, intensidad_con[2], angle=90, code=3,
length=0.1, col="blue")
arrows(2, intensidad_sin[1], 2, intensidad_sin[2], angle=90, code=3,
length=0.1, col="red")
points(1, mean(intensidad_con), pch=19, col="blue")
points(2, mean(intensidad_sin), pch=19, col="red")

```



```

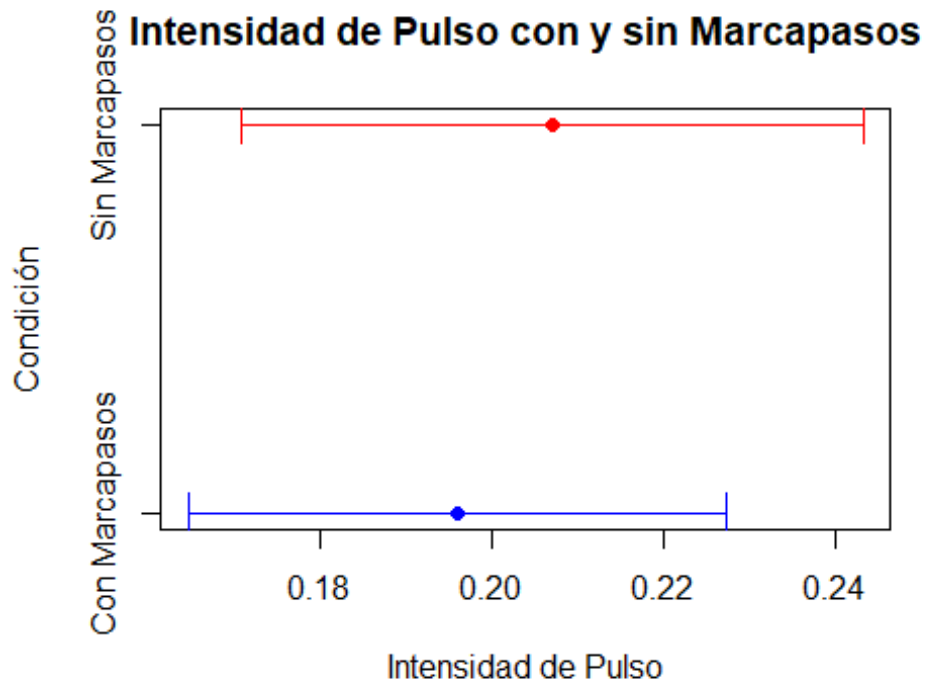
# Configurar Los datos de Los intervalos de confianza
intensidad_con <- c(0.16458, 0.2273)
intensidad_sin <- c(0.17083, 0.24337)

# Crear La gráfica horizontal
plot(1, type="n", ylim=c(1, 2), xlim=c(min(c(intensidad_con,
intensidad_sin)), max(c(intensidad_con, intensidad_sin))),
      ylab="Condición", xlab="Intensidad de Pulso", yaxt="n", main="Intensidad
de Pulso con y sin Marcapasos")
axis(2, at=1:2, labels=c("Con Marcapasos", "Sin Marcapasos"))

# Agregar Los intervalos de confianza
arrows(intensidad_con[1], 1, intensidad_con[2], 1, angle=90, code=3,
length=0.1, col="blue")
arrows(intensidad_sin[1], 2, intensidad_sin[2], 2, angle=90, code=3,

```

```
length=0.1, col="red")
points(mean(intensidad_con), 1, pch=19, col="blue")
points(mean(intensidad_sin), 2, pch=19, col="red")
```



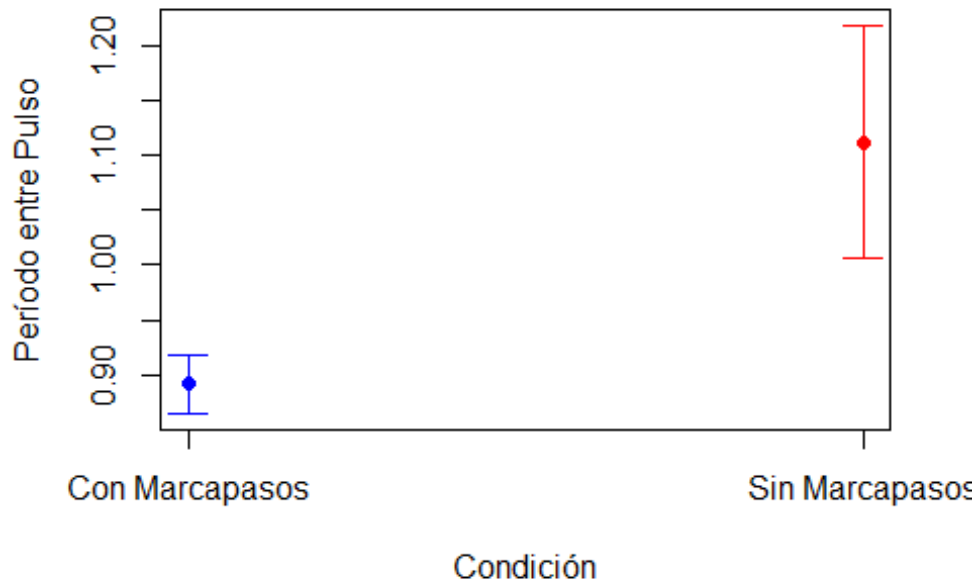
Grafica en un mismo eje coordenado el periodo entre pulso con y sin marcapasos

```
# Configurar Los datos de Los intervalos de confianza
periodo_con <- c(0.86446, 0.9179)
periodo_sin <- c(1.00552, 1.21881)

# Crear La gráfica
plot(1, type="n", xlim=c(1, 2), ylim=c(min(c(periodo_con, periodo_sin)),
max(c(periodo_con, periodo_sin))),
     xlab="Condición", ylab="Período entre Pulso", xaxt="n", main="Período
entre Pulso con y sin Marcapasos")
axis(1, at=1:2, labels=c("Con Marcapasos", "Sin Marcapasos"))

# Agregar Los intervalos de confianza
arrows(1, periodo_con[1], 1, periodo_con[2], angle=90, code=3, length=0.1,
col="blue")
arrows(2, periodo_sin[1], 2, periodo_sin[2], angle=90, code=3, length=0.1,
col="red")
points(1, mean(periodo_con), pch=19, col="blue")
points(2, mean(periodo_sin), pch=19, col="red")
```

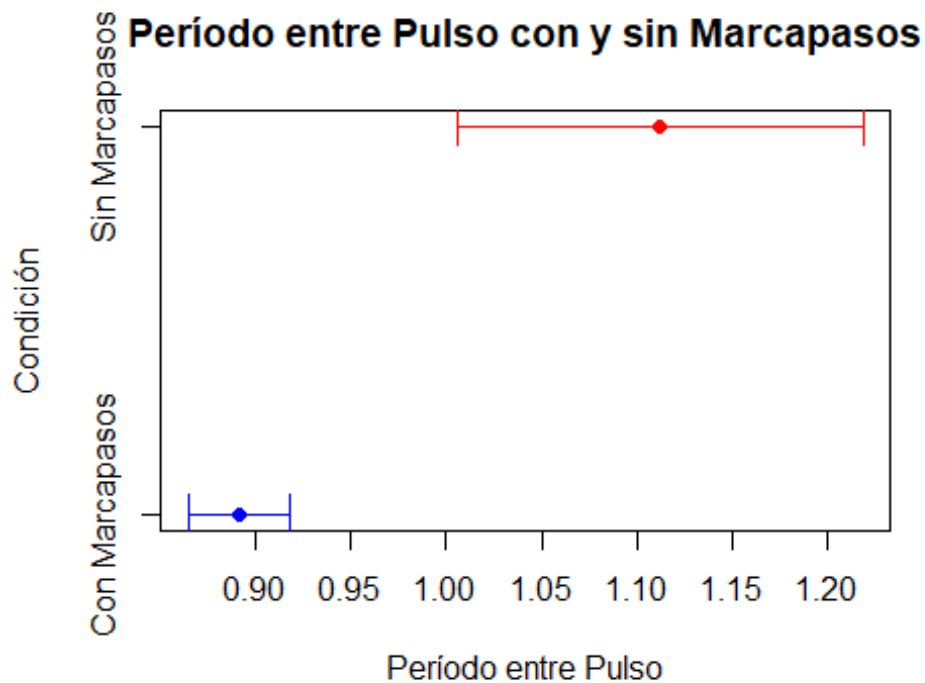
Período entre Pulso con y sin Marcapasos



```
# Configurar Los datos de Los intervalos de confianza
periodo_con <- c(0.86446, 0.9179)
periodo_sin <- c(1.00552, 1.21881)

# Crear La gráfica horizontal
plot(1, type="n", ylim=c(1, 2), xlim=c(min(c(periodo_con, periodo_sin)),
    max(c(periodo_con, periodo_sin))),
    ylab="Condición", xlab="Período entre Pulso", yaxt="n", main="Período
entre Pulso con y sin Marcapasos")
axis(2, at=1:2, labels=c("Con Marcapasos", "Sin Marcapasos"))

# Agregar Los intervalos de confianza
arrows(periodo_con[1], 1, periodo_con[2], 1, angle=90, code=3, length=0.1,
col="blue")
arrows(periodo_sin[1], 2, periodo_sin[2], 2, angle=90, code=3, length=0.1,
col="red")
points(mean(periodo_con), 1, pch=19, col="blue")
points(mean(periodo_sin), 2, pch=19, col="red")
```



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

Intensidad de Pulso con y sin Marcapasos:

-Con Marcapasos: Intervalo de confianza [0.1645, 0.2273] -Sin Marcapasos: Intervalo de confianza [0.1708, 0.2433]

Existe una superposición entre los intervalos de confianza para la intensidad del pulso con y sin marcapasos. Esto sugiere que la diferencia en la intensidad de pulso entre ambas condiciones puede no ser estadísticamente significativa. Aunque la media parece ser un poco más alta sin marcapasos, la superposición indica que esta diferencia podría no ser concluyente.

Aunque los intervalos de confianza muestran que la intensidad del pulso sin marcapasos es ligeramente mayor, la superposición de los intervalos indica que la diferencia no es estadísticamente significativa. Esto sugiere que el marcapasos no tiene un efecto considerable en la intensidad del pulso.

Período entre Pulso con y sin Marcapasos:

-Con Marcapasos: Intervalo de confianza [0.8645, 0.9178] -Sin Marcapasos: Intervalo de confianza [1.0055, 1.2180]

Los intervalos de confianza no se superponen. Esto sugiere una diferencia significativa en el período entre pulso cuando se utiliza un marcapasos en comparación con cuando no se

utiliza. Específicamente, el período entre pulsos es menor cuando se usa un marcapasos, lo que es coherente con la función de un marcapasos para regular y mantener un ritmo cardíaco más constante y rápido.

El uso del marcapasos reduce significativamente el período entre pulsos, lo que confirma su eficacia en regular el ritmo cardíaco. La diferencia observada entre los intervalos de confianza es lo suficientemente grande como para concluir que el marcapasos tiene un impacto positivo en la regulación del tiempo entre latidos.