

Tarea 7 - Intervalos de Confianza

Héctor Hibrán Tapia Fernández - A01661114

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

```
muestras = 150
tamano = 150
media = 70
desviacion = 9

set.seed(25022001)
simulacion <- replicate(muestras, rnorm(tamano, mean = media, sd = desviacion))
```

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

```
nivel_confianza = 0.97
alfa = 1 - nivel_confianza

error_estandar <- desviacion / sqrt(tamano)
z = qnorm(1 - alfa/2)

medias = colMeans(simulacion)
margen_error = z * error_estandar
```

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

```
library(gplots)
```

```
## Warning: package 'gplots' was built under R version 4.2.3
```

```
##  
## Attaching package: 'gplots'
```

```
## The following object is masked from 'package:stats':  
##  
##      lowess
```

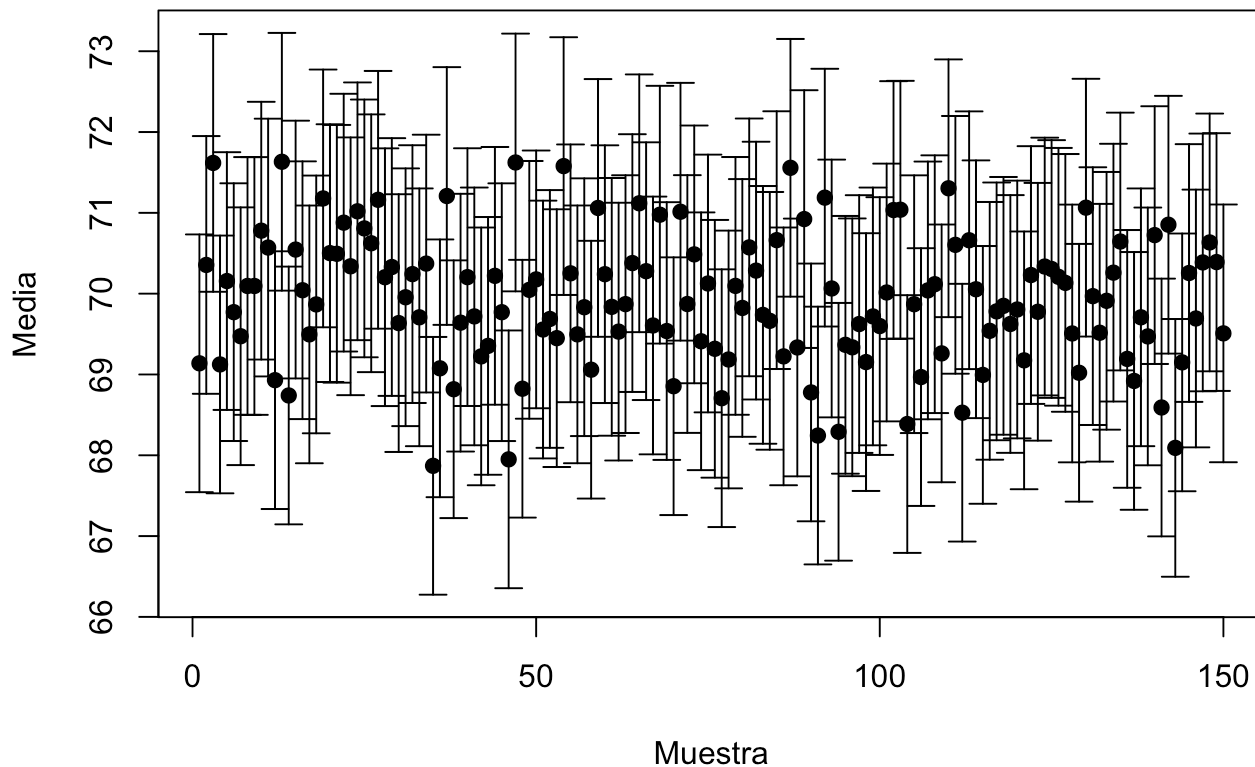
```
library(plotrix)
```

```
##  
## Attaching package: 'plotrix'
```

```
## The following object is masked from 'package:gplots':  
##  
##      plotCI
```

```
plotCI(1:muestras, medias, uiw = margen_error, liw = margen_error, xlab = "Muestra", ylab =  
      "Media", main = "Medias con Intervalos de Confianza del 97%", pch = 19)
```

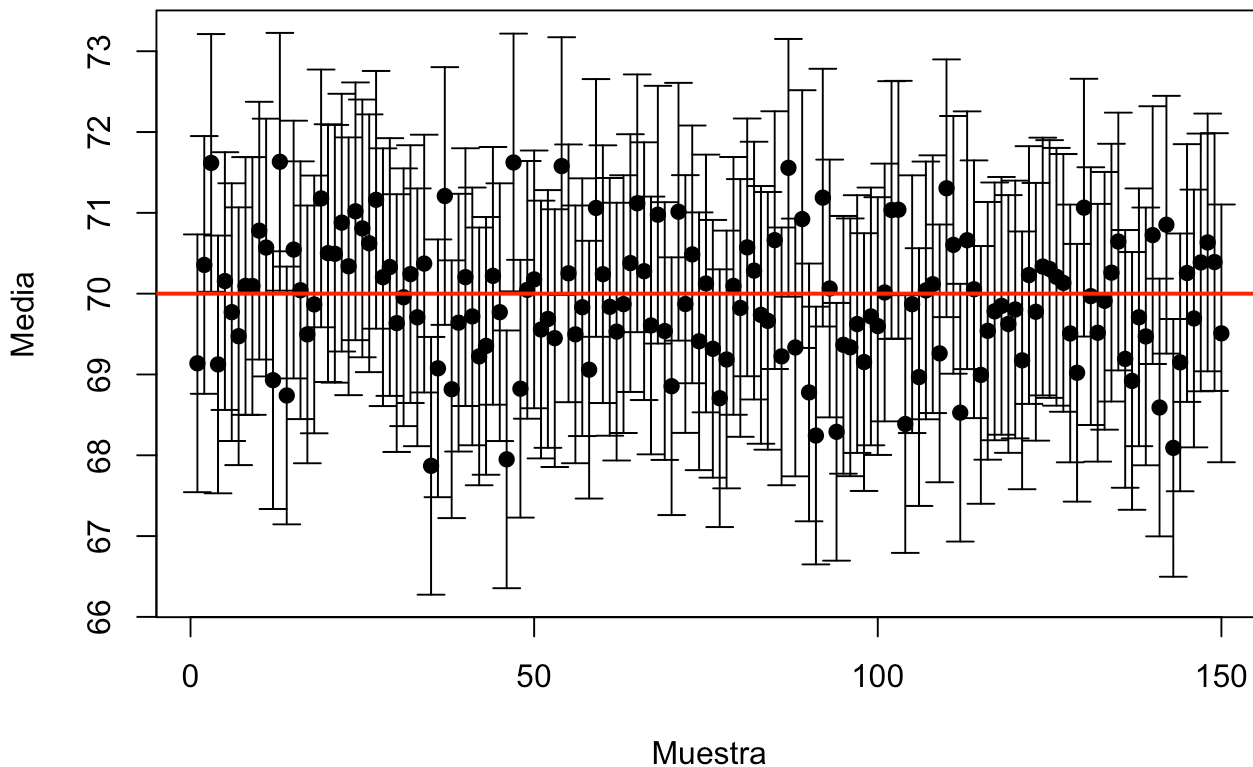
Medias con Intervalos de Confianza del 97%



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

```
plotCI(1:muestras, medias, uiw = margen_error, liw = margen_error, xlab = "Muestra", ylab = "Media", main = "Medias con Intervalos de Confianza del 97%", pch = 19)
abline(h = media, col = "red", lwd = 2)
```

Medias con Intervalos de Confianza del 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)

```
cotas_inferiores = medias - margen_error
cotas_superiores = medias + margen_error

intervalos_contienen_media = sum(cotas_inferiores <= media & cotas_superiores >= media)

porcentaje_contienen_media = (intervalos_contienen_media / 150) * 100

cat("El número de intervalos que contienen la verdadera media:", intervalos_contienen_media, "\n")
```

```
## El número de intervalos que contienen la verdadera media: 141
```

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

```
## Porcentaje de intervalos que contienen la verdadera media: 94 %
```

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  
alfa = 0.03  
xb1 = 4.85  
n1 = 20  
  
E1 = abs(qnorm(0.03/2))*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
alfa = 0.03
xb2 = 4.56
n2 = 16

E2 = abs(qnorm(0.03/2))*sigma/sqrt(n2)
A2 = xb2 - E2
B2 = xb2 + E2

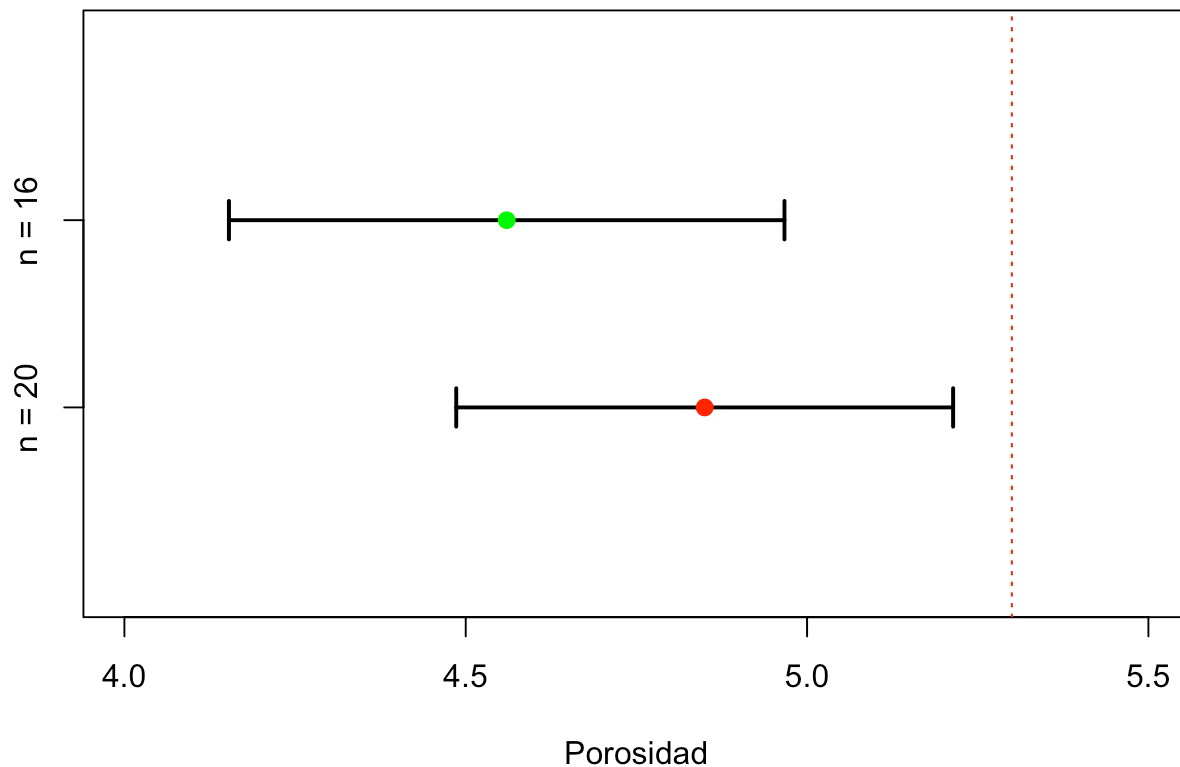
cat("La verdadera media se encuentra entre: ", A2, "y", B2)
```

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



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?

```
desviacion = 0.75
confianza1 = 0.95
alpha = 1 - confianza1
W1 = 0.4

Z1 = abs(qnorm(alpha / 2))
n1 = (Z1 * desviacion / (W1 / 2)) ^ 2
n1 = ceiling(n1)

cat("El tamaño de muestra debe ser:", n1)
```

```
## El tamaño de muestra debe ser: 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%?

```
desviacion = 0.75
W2 = 0.2
confianza2 = 0.99
alpha = 1 - confianza2

Z2 = qnorm(alpha / 2)
n2 = (Z2 * desviacion / W2) ^ 2
n2 = ceiling(n2)

cat("El tamaño de muestra debe ser:", n2)
```

```
## El tamaño de muestra debe ser: 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)

```
df = read.csv("./El marcapasos.csv", header = TRUE, sep = ",")
```



```

con_marcapasos <- df[df$Marpapasos == "Con MP", "Intensidad.de.pulso"]
sin_marcapasos <- df[df$Marpapasos == "Sin MP", "Intensidad.de.pulso"]

n_con_mp <- length(con_marcapasos)
n_sin_mp <- length(sin_marcapasos)

media_con_mp <- mean(con_marcapasos)
media_sin_mp <- mean(sin_marcapasos)

error_estandar_con_mp <- sd(con_marcapasos) / sqrt(n_con_mp)
error_estandar_sin_mp <- sd(sin_marcapasos) / sqrt(n_sin_mp)

margen_error_con_mp <- qt(0.975, df = n_con_mp-1) * error_estandar_con_mp
margen_error_sin_mp <- qt(0.975, df = n_sin_mp-1) * error_estandar_sin_mp

inferior_con_mp <- media_con_mp - margen_error_con_mp
superior_con_mp <- media_con_mp + margen_error_con_mp

inferior_sin_mp <- media_sin_mp - margen_error_sin_mp
superior_sin_mp <- media_sin_mp + margen_error_sin_mp

cat("La media de la intensidad de pulso con marcapasos se encuentra entre:", inferior_co
n_mp, "y", superior_con_mp, "\n")

```

```

## La media de la intensidad de pulso con marcapasos se encuentra entre: 0.1638035 y 0.2
280788

```

```

cat("La media de la intensidad de pulso sin marcapasos se encuentra entre:", inferior_si
n_mp, "y", superior_sin_mp, "\n")

```

```

## La media de la intensidad de pulso sin marcapasos se encuentra entre: 0.16993 y 0.244
2661

```

Periodo entre pulso con y sin Marcapasos (2 intervalos de

confianza)

```

periodo_con_mp <- df[df$Marcapasos == "Con MP", "Periodo.entre.pulsos"]
periodo_sin_mp <- df[df$Marcapasos == "Sin MP", "Periodo.entre.pulsos"]

n_periodo_con_mp <- length(periodo_con_mp)
n_periodo_sin_mp <- length(periodo_sin_mp)

media_periodo_con_mp <- mean(periodo_con_mp)
media_periodo_sin_mp <- mean(periodo_sin_mp)

error_estandar_periodo_con_mp <- sd(periodo_con_mp) / sqrt(n_periodo_con_mp)
error_estandar_periodo_sin_mp <- sd(periodo_sin_mp) / sqrt(n_periodo_sin_mp)

margen_error_periodo_con_mp <- qt(0.975, df = n_periodo_con_mp-1) * error_estandar_perio
do_con_mp
margen_error_periodo_sin_mp <- qt(0.975, df = n_periodo_sin_mp-1) * error_estandar_perio
do_sin_mp

inferior_periodo_con_mp <- media_periodo_con_mp - margen_error_periodo_con_mp
superior_periodo_con_mp <- media_periodo_con_mp + margen_error_periodo_con_mp

inferior_periodo_sin_mp <- media_periodo_sin_mp - margen_error_periodo_sin_mp
superior_periodo_sin_mp <- media_periodo_sin_mp + margen_error_periodo_sin_mp

cat("La media del período entre pulsos con marcapasos se encuentra entre:", inferior_per
iodo_con_mp, "y", superior_periodo_con_mp, "\n")

```

```

## La media del período entre pulsos con marcapasos se encuentra entre: 0.8637941 y 0.91
85589

```

```

cat("La media del período entre pulsos sin marcapasos se encuentra entre:", inferior_per
iodo_sin_mp, "y", superior_periodo_sin_mp, "\n")

```

```

## La media del período entre pulsos sin marcapasos se encuentra entre: 1.002887 y 1.220
643

```

Grafica los intervalos de confianza obtenidos en “El marcapasos”:

Grafica en un mismo eje coordinado la intensidad de pulso con

y sin marcapasos

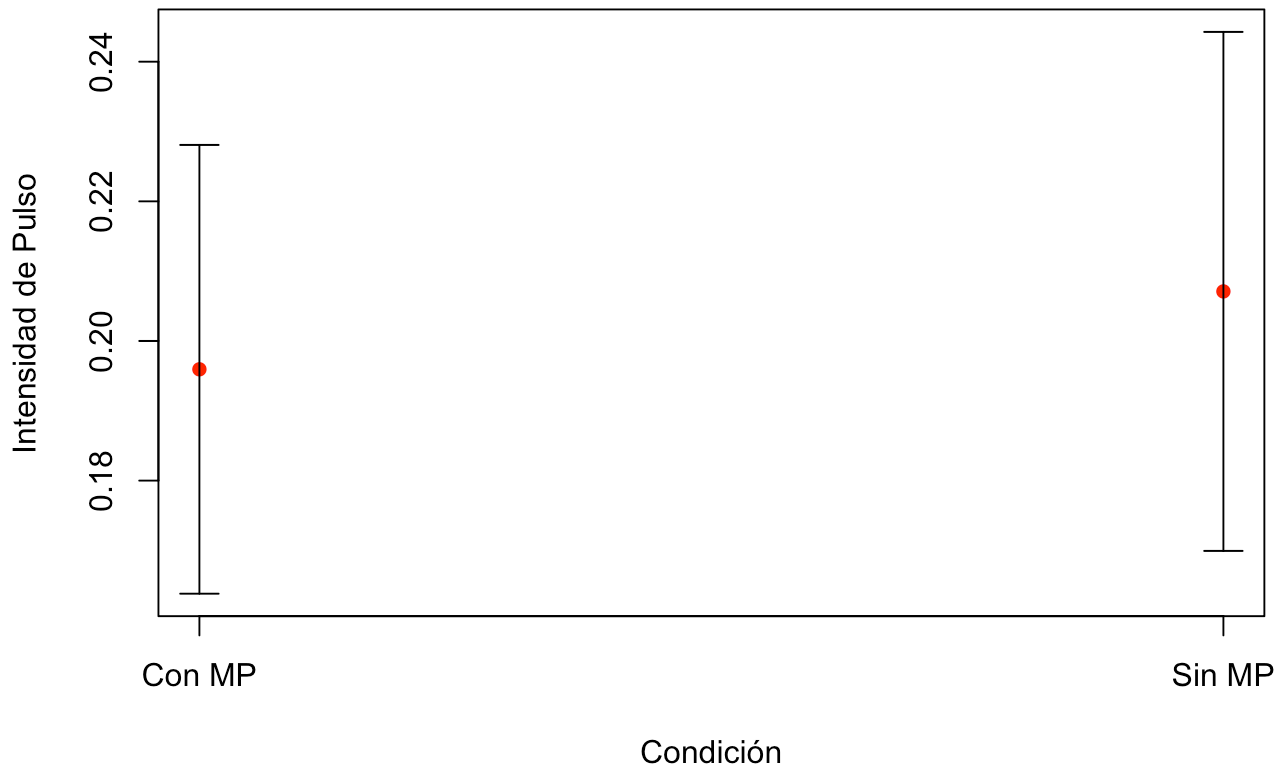
```
medias_muest1 <- media_con_mp
medias_muest2 <- media_sin_mp

A1 <- inferior_con_mp
B1 <- superior_con_mp

A2 <- inferior_sin_mp
B2 <- superior_sin_mp

plot(c(1, 2), c(medias_muest1, medias_muest2), ylim = c(min(A1, A2), max(B1, B2)), xaxt =
"n", pch = 16, col = "red", xlab = "Condición", ylab = "Intensidad de Pulso", main = "Int
ervalos 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)
arrows(2, A2, 2, B2, angle = 90, code = 3, length = 0.1)
```

Intervalos de Confianza de la Intensidad de Pulso



Grafica en un mismo eje coordinado el periodo entre pulso con

y sin marcapasos

```
medias_muest1 <- media_periодо_con_mp
medias_muest2 <- media_periодо_sin_mp

A1 <- inferior_periодо_con_mp
B1 <- superior_periодо_con_mp

A2 <- inferior_periодо_sin_mp
B2 <- superior_periодо_sin_mp

plot(c(1, 2), c(medias_muest1, medias_muest2), ylim = c(min(A1, A2), max(B1, B2)), xaxt =
"n", pch = 16, col = "red", xlab = "Condición", ylab = "Período entre pulsos", main = "I
ntervalos de Confianza del Período entre Pulsos")
axis(1, at = c(1, 2), labels = c("Con MP", "Sin MP"))
arrows(1, A1, 1, B1, angle = 90, code = 3, length = 0.1)
arrows(2, A2, 2, B2, angle = 90, code = 3, length = 0.1)
```

Intervalos de Confianza del Período entre Pulsos



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

Vayamos por gráfico:

- Gráfico de Intensidad de Pulso: En este gráfico se observa una pequeña diferencia en las medias de la intensidad de pulso con y sin marcapasos, lo que nos dice que esta diferencia podría no ser estadísticamente significativa, la intensidad del pulso es similar en ambas condiciones.
- Gráfico de Período entre Pulsos: En este gráfico se observa que existe una diferencia enorme y estadísticamente significativa en el período entre pulsos con y sin marcapasos. Con marcapasos se reduce notablemente el tiempo entre pulsos, lo que nos dice que el dispositivo cumple con su función de regular el ritmo cardíaco.