

Act 4 Intervalos de confianza

Diego Rodríguez

2023-08-22

Problema del marcapasos

```
library(dplyr)

##
## Attaching package: 'dplyr'
## The following objects are masked from 'package:stats':
##
##   filter, lag
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union

library(ggplot2)

df <- read.csv("El marcapasos.csv")
summary(df)

## Periodo.entre.pulsos Intensidad.de.pulso Marcapasos
## Min. :0.690 Min. :0.0050 Length:102
## 1st Qu.:0.830 1st Qu.:0.1090 Class :character
## Median :0.935 Median :0.1910 Mode :character
## Mean :1.001 Mean :0.2015
## 3rd Qu.:1.100 3rd Qu.:0.2923
## Max. :2.800 Max. :0.5850

df <- df %>%
  mutate(Marcapasos = ifelse(Marcapasos == "Con MP", 1, 0))

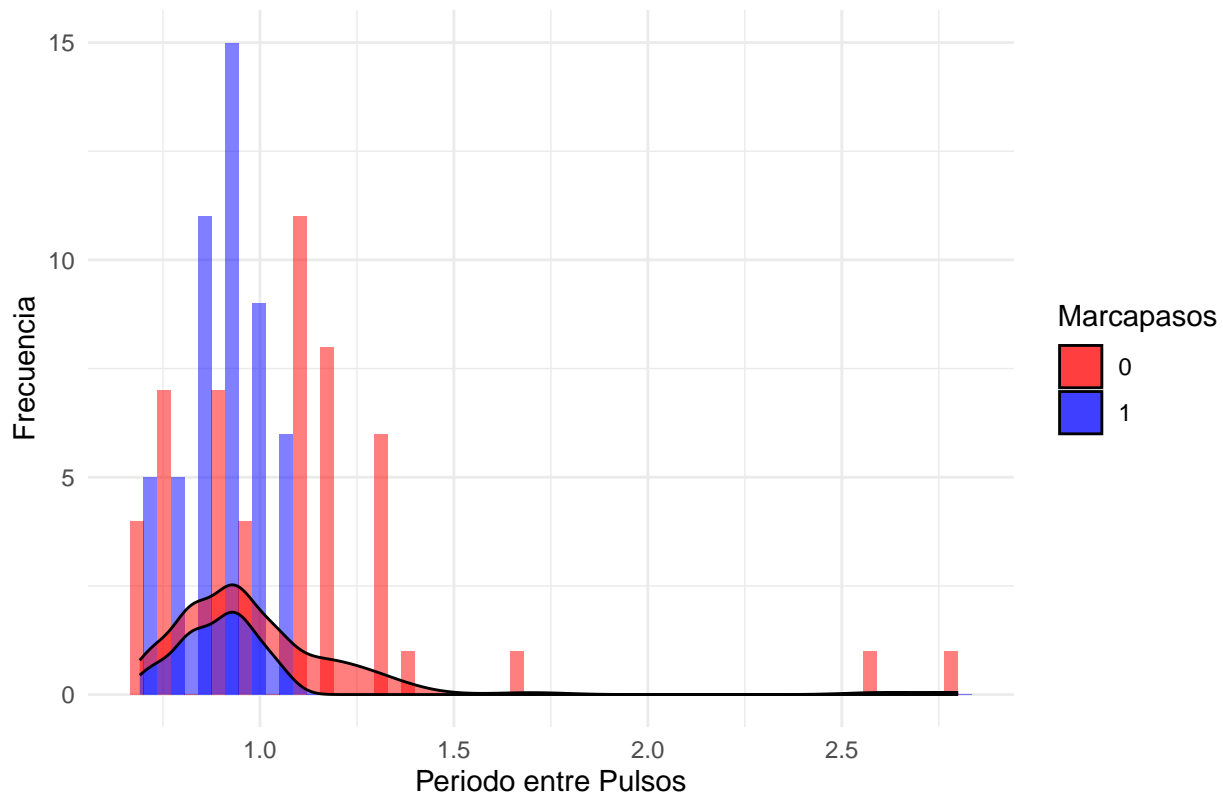
#con_marcapasos = subset(df, df$Marcapasos == "Con MP")

#sin_marcapasos = subset(df, df$Marcapasos == "Sin MP")

ggplot(df, aes(x = Periodo.entre.pulsos, fill = factor(Marcapasos))) +
  geom_histogram(binwidth = 0.07, position = "dodge", alpha = 0.5) +
  geom_density(aes(y = ..count.. * 0.01), position = "stack", alpha = 0.5, color = "black") +
  labs(title = "Histograma de Periodo entre Pulsos",
       x = "Periodo entre Pulsos",
       y = "Frecuencia",
       fill = "Marcapasos") +
  scale_fill_manual(values = c("0" = "red", "1" = "blue")) +
  theme_minimal()
```

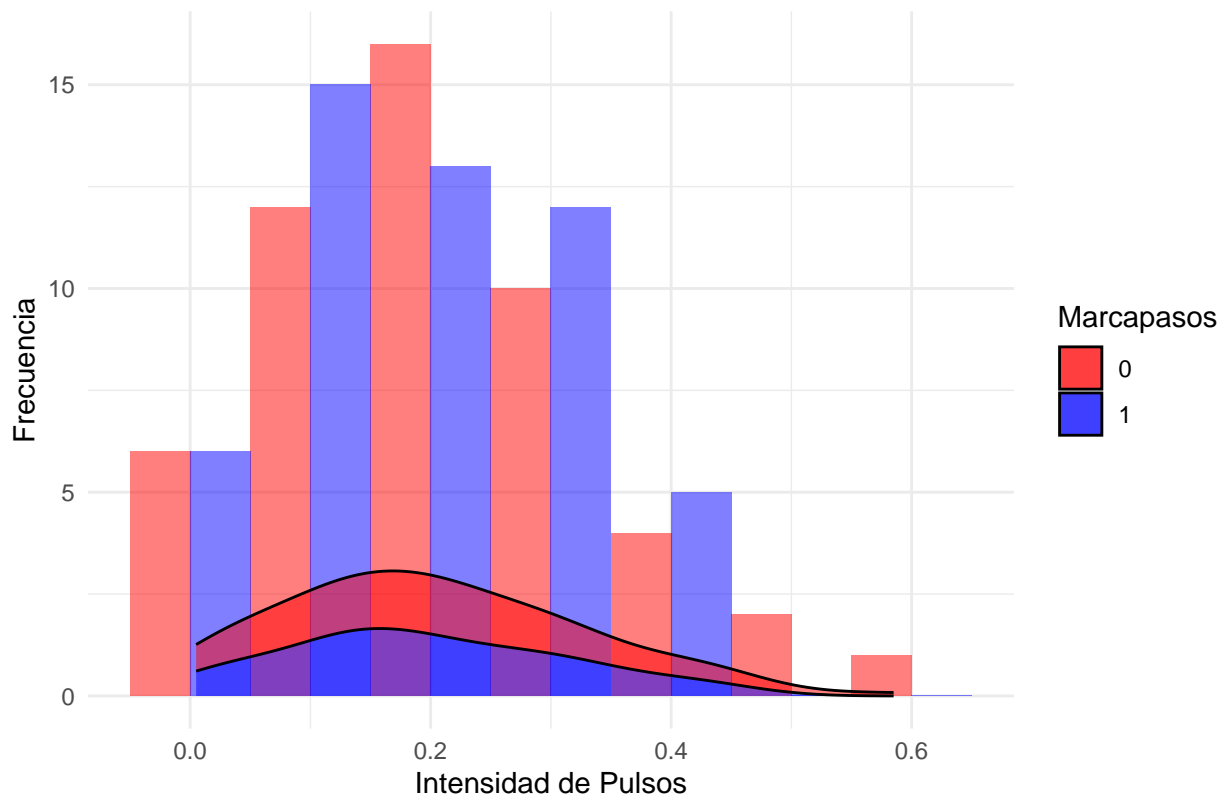
```
## Warning: The dot-dot notation (`..count..`) was deprecated in ggplot2 3.4.0.
## i Please use `after_stat(count)` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```

Histograma de Periodo entre Pulsos



```
ggplot(df, aes(x = Intensidad.de.pulso, fill = factor(Marcapasos))) +
  geom_histogram(binwidth = 0.1, position = "dodge", alpha = 0.5) +
  geom_density(aes(y = ..count.. * 0.01), position = "stack", alpha = 0.5, color = "black") +
  labs(title = "Histograma y Distribución de Intensidad de Pulsos",
       x = "Intensidad de Pulsos",
       y = "Frecuencia",
       fill = "Marcapasos") +
  scale_fill_manual(values = c("0" = "red", "1" = "blue")) +
  theme_minimal()
```

Histograma y Distribución de Intensidad de Pulsos



las graficas muestran que las personas sin marcapasos tienden a tener una intensidad de pulso promedio mayor al de las personas con marcapasos, ademas de un periodo también mayor.

Intervalos de confianza

```
# Calcula el intervalo de confianza para la media de la Intensidad de pulsos
intensidad_ci <- t.test(Intensidad.de.pulso ~ Marcapasos, data = df)
```

```
# Calcula el intervalo de confianza para la media del Periodo entre pulso
periodo_ci <- t.test(Periodo.entre.pulsos ~ Marcapasos, data = df)
```

```
ci_data <- data.frame(
  Variable = c("Intensidad de pulsos", "Periodo entre pulso"),
  Marcapasos = c("Con MP", "Sin MP"),
  Lower_CI = c(intensidad_ci$conf.int[1], periodo_ci$conf.int[1]),
  Upper_CI = c(intensidad_ci$conf.int[2], periodo_ci$conf.int[2])
)
```

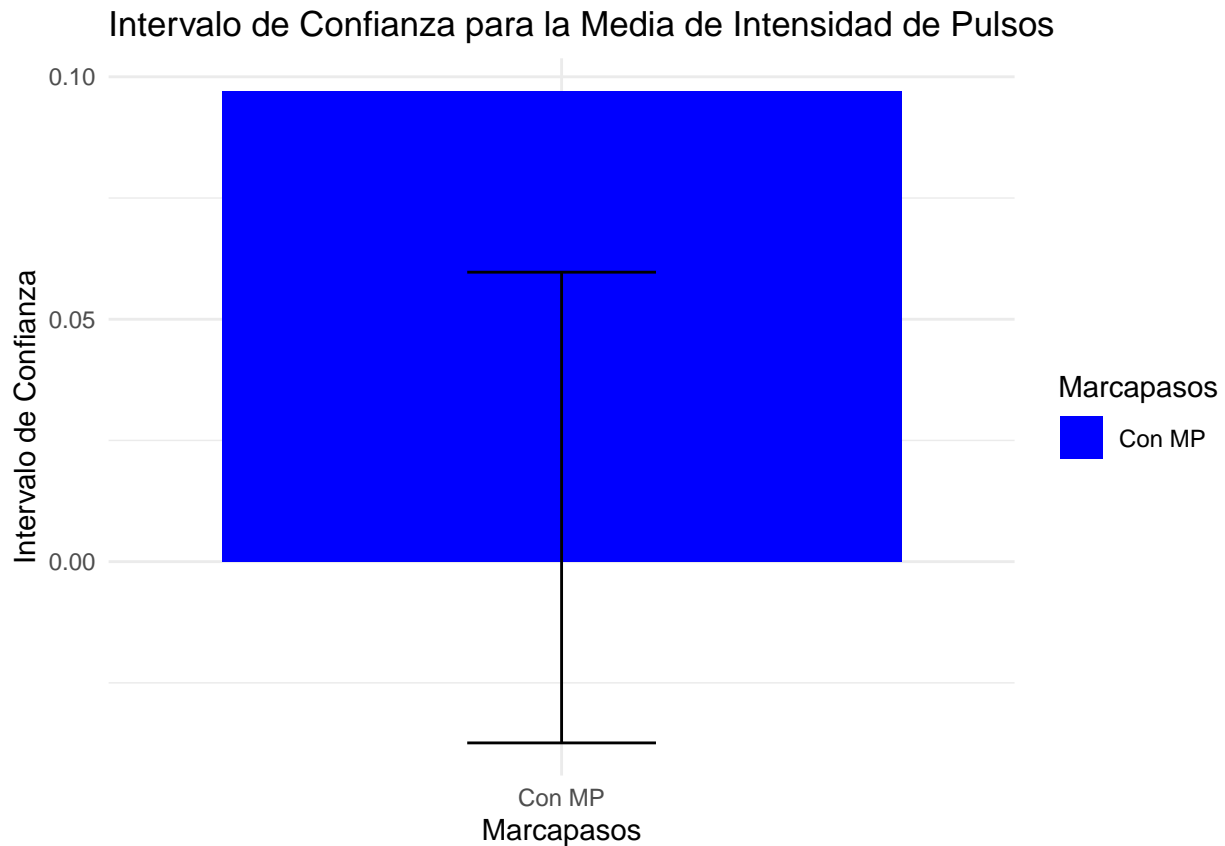
```
ci_data
```

```
##           Variable Marcapasos   Lower_CI   Upper_CI
## 1 Intensidad de pulsos      Con MP -0.03738952 0.05970324
## 2 Periodo entre pulso      Sin MP  0.10863004 0.33254643
```

```
intensidad_plot <- ggplot(ci_data[ci_data$Variable == "Intensidad de pulsos", ], aes(x = Marcapasos, y = 
  geom_bar(stat = "identity", position = "dodge") + 
  geom_errorbar(aes(ymin = Lower_CI, ymax = Upper_CI), position = position_dodge(width = 0.9), width = 0.5))
```

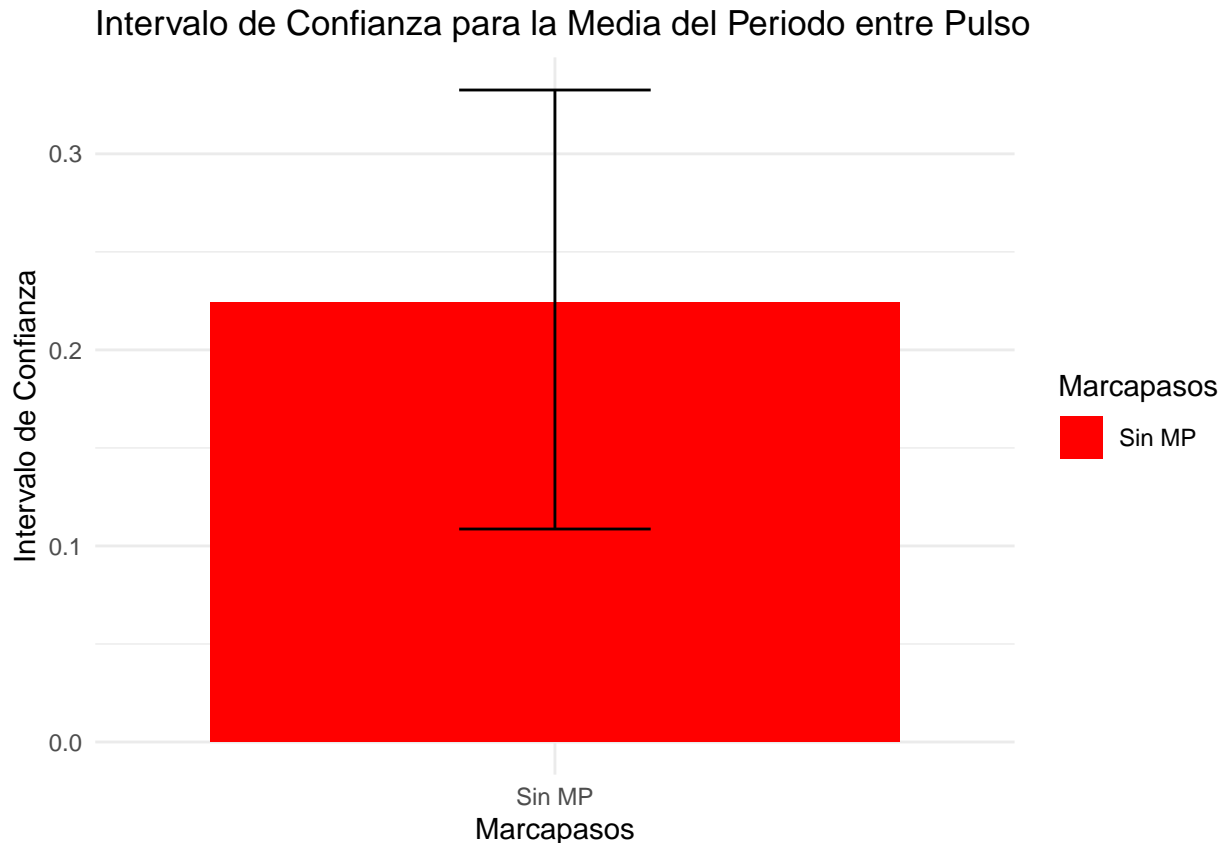
```
labs(title = "Intervalo de Confianza para la Media de Intensidad de Pulsos",
      y = "Intervalo de Confianza") +
scale_fill_manual(values = c("Con MP" = "blue", "Sin MP" = "red")) +
theme_minimal()
```

intensidad_plot



```
periodo_plot <- ggplot(ci_data[ci_data$Variable == "Periodo entre pulso", ], aes(x = Marcapasos, y = Upper_CI)) +
  geom_bar(stat = "identity", position = "dodge") +
  geom_errorbar(aes(ymin = Lower_CI, ymax = Upper_CI), position = position_dodge(width = 0.9), width = 0.5) +
  labs(title = "Intervalo de Confianza para la Media del Periodo entre Pulso",
        y = "Intervalo de Confianza") +
  scale_fill_manual(values = c("Con MP" = "blue", "Sin MP" = "red")) +
  theme_minimal()
```

periodo_plot



Intensidad de pulsos (Con MP):

El intervalo de confianza para la media de la intensidad de pulsos en el grupo “Con Marcapasos” va desde -0.03738952 hasta 0.05970324. Esto significa que podemos estar 95% seguros de que la verdadera media de la intensidad de pulsos en el grupo “Con Marcapasos” cae dentro de este rango. Dado que el intervalo incluye valores positivos y negativos cercanos a cero, no podemos concluir con certeza si hay una diferencia significativa en la intensidad de pulsos entre los pacientes “Con Marcapasos” y los que no lo tienen.

Periodo entre pulso (Sin MP):

El intervalo de confianza para la media del periodo entre pulsos en el grupo “Sin Marcapasos” va desde 0.10863004 hasta 0.33254643. Esto significa que podemos estar 95% seguros de que la verdadera media del periodo entre pulsos en el grupo “Sin Marcapasos” cae dentro de este rango. Dado que el intervalo está completamente por encima de cero y no incluye valores negativos, podemos concluir que hay una diferencia estadísticamente significativa en el periodo entre pulsos entre los pacientes “Sin Marcapasos” y los que sí lo tienen, siendo mayor en el grupo “Sin Marcapasos”.

##Problema Helio

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?

La fórmula general es:

$$n = \left(\frac{Z \cdot \sigma}{E} \right)^2$$

Donde: - n es el tamaño de la muestra que queremos determinar. - Z es el valor crítico de la distribución normal estándar para un nivel de confianza del 95%. Para un nivel de confianza del 95%, Z es aproximadamente igual a 1.96. - σ es la desviación estándar verdadera de la población, que en este caso es 0.75. - E es el margen de error que deseamos, que en este caso es 0.4.

```
sigma <- 0.75
E <- 0.4

conf_level <- 0.95

Z <- qnorm((1 + conf_level) / 2)

n <- ((Z * sigma) / E)^2
n <- ceiling(n)

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

```
## El tamaño de muestra necesario es: 14
```

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
E <- 0.2
conf_level <- 0.99
Z <- qnorm((1 + conf_level) / 2)
n <- ((Z * sigma) / E)^2
n <- ceiling(n)

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

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