

Ejercicio Concurso estadística

Avila Aylén

2022-11-29

Consigna: Construir un intervalo de confianza al 95% para el contenido de Cinc en muestras fisiológicas de pacientes con cierta enfermedad ($\sigma = 5$ ppm).

Los datos y el script están disponibles en mi GitHub (<https://github.com/lmoPupato/Estadistica-Ejemplo-1>)

- Primero seteamos el directorio de trabajo, cargamos los datos y los exploramos:

```
setwd("G:/Mi unidad/Git/Estadistica-Concurso/Estadistica-Ejemplo-1")
data <- read.csv("G:/Mi unidad/Git/Estadistica-Concurso/Estadistica-Ejemplo-1/data.txt", sep=
"")
summary(data)
```

```
##      Muestra      Conc.
## Min.   : 1.00   Min.   :138.1
## 1st Qu.: 25.75   1st Qu.:145.7
## Median : 50.50   Median :149.2
## Mean   : 50.50   Mean   :149.6
## 3rd Qu.: 75.25   3rd Qu.:153.8
## Max.   :100.00   Max.   :160.5
```

- Luego, cargamos la librería para evaluar la normalidad de nuestros datos:
Si bien ya sabemos que por tener un $n > 30$, la variable tendrá una distribución normal, en caso de que no lo sepamos o dudemos siempre debemos verificarlo.

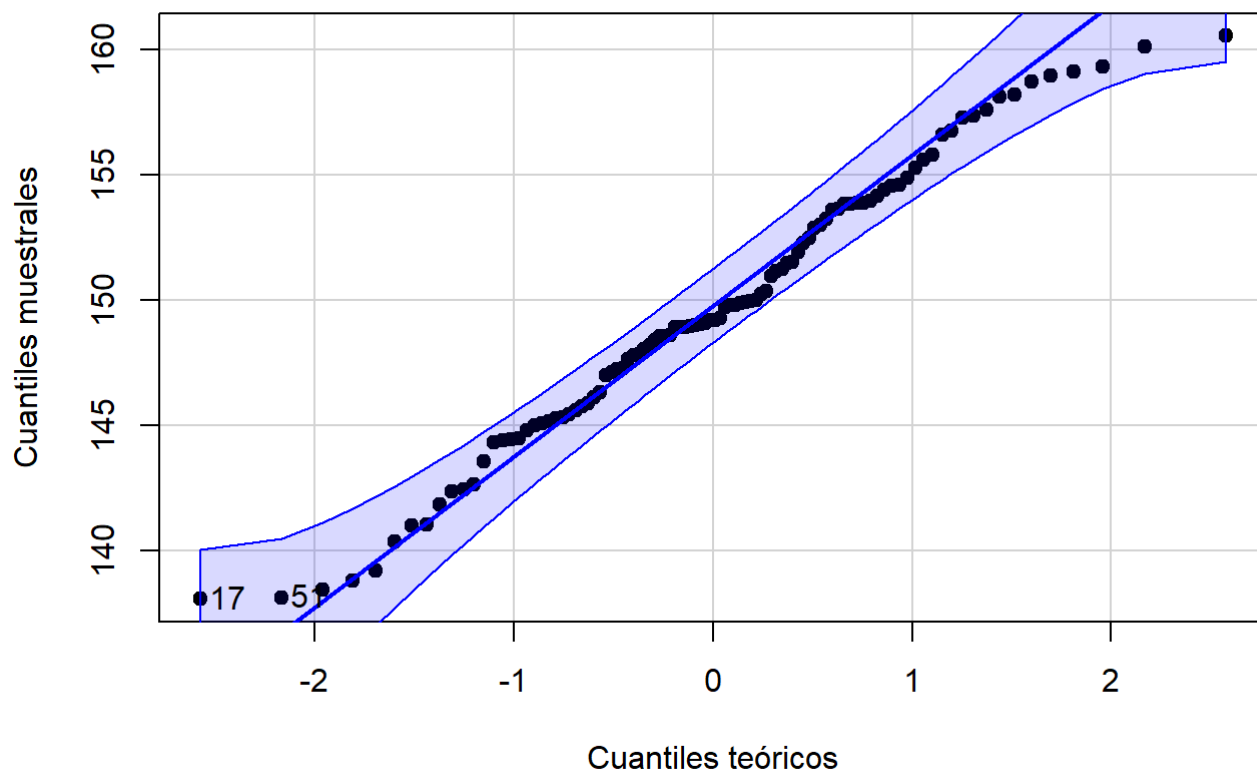
```
library(car)
```

```
## Loading required package: carData
```

- Gráficos de normalidad y distribución

```
qqPlot(data$Conc., pch=19,
        main='QQplot para la Concentración de Cinc en muestras fisiológicas',
        xlab='Cuantiles teóricos',
        ylab='Cuantiles muestrales')
```

QQplot para la Concentración de Cinc en muestras fisiológicas

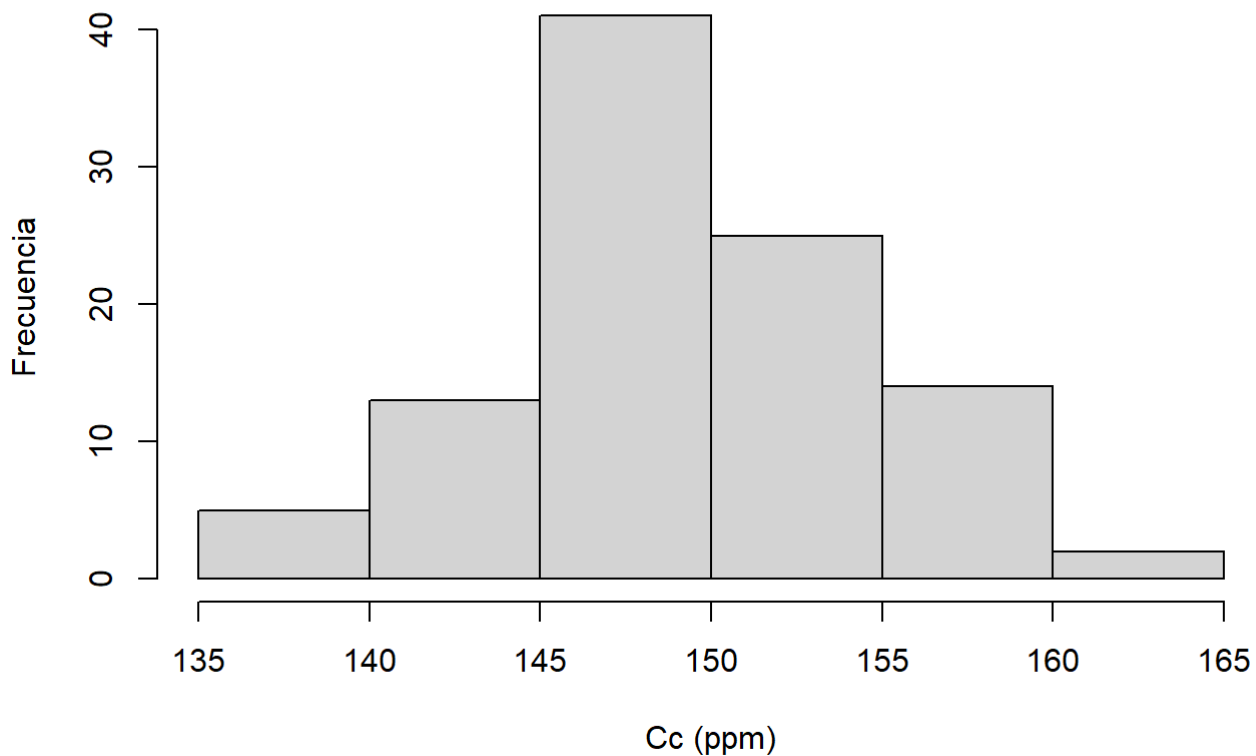


```
## [1] 17 51
```

Podemos observar que los puntos se alinean sobre la recta y no hay ninguno por fuera de las líneas de control. Podemos decir, de la lectura del gráfico de probabilidad normal, que la concentración de cinc en muestras fisiológicas sigue una distribución normal.

```
hist(data$Conc., freq=TRUE,  
      main='Histograma de la Concentración de Cinc en muestras fisiológicas',  
      xlab='Cc (ppm)',  
      ylab='Frecuencia')
```

Histograma de la Concentración de Cinc en muestras fisiológicas



De la inspección del histograma podemos ver que la distribución es simétrica.

- Test de normalidad:

Lo apropiado es realizar el test para despejar cualquier duda.

```
shapiro.test(data$Conc.)
```

```
##  
##  Shapiro-Wilk normality test  
##  
## data:  data$Conc.  
## W = 0.98443, p-value = 0.2887
```

Como $p < 0.05$ no rechazamos H_0 y por lo tanto aceptamos que la variable se distribuye de manera normal.

- Creación del IC: con $\text{desvío} = 5\text{ppm}$ y $P(EI) = \alpha = 0.05$

```
media<-mean(data$Conc.)  
n<-length(data$Conc.)  
desvio<-5  
errorst<-desvio/sqrt(n)  
Z_0.975<-1.96 # Recordar que  $\alpha=0.05$ ,  $1-\alpha/2 = 0.975$ 
```

```
lim_inf<-media-(Z_0.975*errorst)
lim_sup<-media+(Z_0.975*errorst)

interval_m<-data.frame(n, media, desvio, lim_inf, lim_sup)

interval_m
```

```
##      n      media desvio  lim_inf  lim_sup
## 1 100 149.6156      5 148.6356 150.5956
```

Se concluye que, con un nivel de confianza del 95% el intervalo (148.64, 151.60) contiene al verdadero valor de la concentración de cinc en la población de pacientes con cierta enfermedad.