

Generación de intervalos de confianza al 95%

Karla E. Juárez Contreras

29/10/2023

Generación de intervalos de confianza al 95%

Primero se genera una población de 10000 intervalos de confianza para muestras $n=100$, $\mu = np = 100 \times 0.35 = 35$, $\sigma = npq = 100 \times 0.35 \times 0.65 = 22.75$

```
# Se establece una semilla porque se va a usar una función aleatoria  
#de esta manera, se asegura que el resultado sea reproducible  
set.seed(1)  
  
# Se generan 2 vectores numéricos en donde se almacenarán 10000 números  
#en este momento, solo estarán llenos de 0  
lower <- numeric(10000)  
upper <- numeric(10000)  
  
# Se generan 10000 conjuntos de datos, y cada uno de ellos consiste en n=100 números aleatorios  
#de distribución normal con media =5 y #desviación estándar =5. Para cada uno de esos  
#10000 conjuntos de datos, se calculan los límites superiores e inferiores del intervalo al 95%  
#considerando  $\mu \pm 1.96 * \text{desviación estándar} / \text{raíz cuadrada}(n)$   
#desviación estándar= 5 y raíz cuadrada(100)=10  
  
for(i in 1:10000) {  
  
  Y <- rnorm(100, mean = 35, sd = 22.75)  
  lower[i] <- mean(Y) - 1.96 * 22.75 / 10  
  upper[i] <- mean(Y) + 1.96 * 22.75 / 10  
  
}
```

Con la siguiente instrucción, se guardan los límites superiores e inferiores de cada intervalo

```
CIs <- cbind(lower, upper)
```

A continuación se pregunta en qué intervalos, de los primeros 100 intervalos (1:100) está contenida $\mu=35$. Para ello se pide que el límite inferior ($CIs[,1]$) sea mayor o igual a 35 y que el límite superior ($CIs[,2]$) sea menor o igual a 35. Los intervalos en donde se indica TRUE son los que si contienen a μ y aquellos en los que se indica FALSE son los que no la contienen

```
which((CIs[1:100, 1] <= 35 & 35 <= CIs[1:100, 2]))
```

```
## [1] 1 2 3 4 5 6 8 9 10 11 12 14 15 16 17 18 19 20 21  
## [20] 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40  
## [39] 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59  
## [58] 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78
```

```
## [77] 79 80 81 82 83 84 86 87 88 89 91 92 93 94 95 96 97 98 99
## [96] 100
```

Se pide a continuación que identifique en qué intervalos no está μ . Para ello se hace la pregunta anterior pero un! antes de la definición de los límites. Es decir, se indican qué intervalos de esos primeros 100 no cumplen con qué su límite inferior sea menor o igual a 35, y su límite superior sea mayor o igual a 35. Esos casos se guardan con el nombre ID

```
which(!(CIs[1:100, 1] <= 35 & 35 <= CIs[1:100, 2]))
```

```
## [1] 7 13 85 90
```

```
ID <- which(!(CIs[1:100, 1] <= 35 & 35 <= CIs[1:100, 2]))
```

Se genera la gráfica, especificando que los intervalos identificados con ID serán coloreados con rojo

```
plot(0,
      xlim = c(20, 50),
      ylim = c(1, 100),
      ylab = "Muestra",
      xlab = "Porcentaje de votos",
      main = "Intervalos de confianza")

# establecer colores de la gráfica
colors <- rep(gray(0.6), 100)
colors[ID] <- "red"

#Se agrega una línea punteada a la altura de mu
abline(v = 35, lty = 2, col="black", lwd=2)

# Se agregan líneas horizontales que representan los intervalos de confianza
for(j in 1:100) {
  lines(c(CIs[j, 1], CIs[j, 2]),
        c(j, j),
        col = colors[j],
        lwd = 2)
}
```

Intervalos de confianza

