

Intervalos de confianza

Danilo Ibáñez Rojas

26 de junio de 2018

En esta primera función se calcula una población al azar donde tendrá n alumnos, con una media y desviación estandar fijadas por el usuario. Se incluye la semilla 123 para que todos los alumnos que realizamos este taller podamos obtener la misma muestra aunque esta sea aleatoria.

```
poblacion <- function(n, media, desviacion){  
  set.seed(123)  
  ppa <- rnorm(n,media,desviacion)  
  return(ppa)  
}
```

En las siguientes líneas se llama a la función anterior con una población de 5000 alumnos, con una media de 4 y desviación estandar de 0.5. Luego se guardan los datos en una matriz y se obtiene una muestra de 300 para esta población, a la cual se le calcula su media, su desviación estandar y sus valores límites calculados de manera teórica los intervalos de confianza.

```
ppa <- poblacion(5000,4,0.5)  
matrixPpa <- matrix(ppa, ncol = 1)  
  
muestraOriginal <- sample(ppa, 300, replace = FALSE)  
mediaOriginal <- mean(muestraOriginal)  
desviacionEstandarOriginal <- sd(ppa)  
valorSuperior = mean(muestraOriginal) + qnorm(1-(0.05/2))*(desviacionEstandarOriginal/sqrt(length(muestraOriginal)))  
valorInferior = mean(muestraOriginal) - qnorm(1-(0.05/2))*(desviacionEstandarOriginal/sqrt(length(muestraOriginal)))
```

Las siguientes funciones sirven para calcular 500 muestras de largo 300 de la muestra obtenida anteriormente, además de calcular la media de cada muestra y agregarla a un vector llamado `medias`.

```
matrixMuestra <- matrix(c(1:500*300),nrow = 300, ncol = 500)  
for(i in 1:500){  
  matrixMuestra[,i] = apply(matrix(muestraOriginal,nrow = 300), 2, sample, replace = TRUE)  
}  
  
medias <- c()  
for(i in 1:500){  
  medias[i] = mean(matrixMuestra[,i])  
}
```

Las próximas líneas ordenan el vector calculado anteriormente y se calculan su percentiles para 0.975 y 0.025.

```
medias = sort(medias)  
valorSuperiorMuestra = quantile(medias,0.975)  
valorInferiorMuestra = quantile(medias,0.025)
```

Por último se imprime en pantalla los valores obtenidos en la primera instancia en forma teórica y los obtenidos en las líneas anteriores mediante el proceso de calcular varias muestras y obtener sus medias y percentiles correspondientes.

```
print('Los intervalos obtenidos con la muestra original fueron:')  
print(paste('Valor inferior:',valorInferior,'-', 'Valor superior:',valorSuperior))
```

```
## [1] "Los intervalos obtenidos con la muestra original fueron:"  
## [2] "Valor inferior: 3.97443335577508 - Valor superior: 4.08697758587017"
```

```
print('Los intervalos obtenidos con las 500 muestras fueron:')  
print(paste('Valor inferior:',valorInferiorMuestra,'-', 'Valor superior:',valorSuperiorMuestra))
```

```
## [1] "Los intervalos obtenidos con las 500 muestras fueron:"  
## [2] "Valor inferior: 3.97783622067325 - Valor superior: 4.0810285230078"
```

Se puede apreciar que los valores son practicamente similares, por tanto, es una buena aproximación al método formal.