

Tarea 5

Daniel Camarena

7/9/2017

Tarea 5

Se realiza un experimento en el que se seleccionan 7 ratones de manera aleatoria de un total de 16 ratones. A los siete seleccionados se les suministra un tratamiento mientras que los restantes formarán el grupo de control. El objetivo del tratamiento es prolongar la supervivencia de los ratones. La siguiente tabla muestra el tiempo de supervivencia en días después de suministrar el tratamiento.

Grupo	Datos	Tamaño de muestra
Tratamiento	94, 197, 16, 38, 99, 141, 23	7
Control	52, 104, 146, 10, 51, 30, 40, 27, 46	9

Tu objetivo es determinar si el tratamiento prolongó la supervivencia.

1. Usa las medias de las muestras para determinar si hay diferencias en los grupos, esto es calcula $\bar{x} - \bar{y}$

```
library(tidyverse)
x <- c(94,197,16,38,99,141,23)
y <- c(52,104,146,10,51,30,40,27,46)

bar_x<-mean(x)
bar_y<-mean(y)

result <- bar_x - bar_y

print(result)
```

```
## [1] 30.63492
```

2. Estima el error estándar de la diferencia usando bootstrap.

```
mediaBoot <- function(x, y){
  n <- length(x)
  n2<- length(y)
  muestra_boot_x <- sample(x, size = n, replace = TRUE)
  muestra_boot_y <- sample(y, size = n2, replace = TRUE)
  mean_bx <- mean(muestra_boot_x)
  mean_by <- mean(muestra_boot_y)
  result <- mean_bx - mean_by
  return(result)
}

thetas_boot <- rerun(1000, mediaBoot(x,y)) %>% flatten_dbl()
sd(thetas_boot)
```

```
## [1] 27.59508
```

3. Corroborar tus resultados del punto anterior usando que $\hat{se} = \sqrt{\hat{se}(x^2) + \hat{se}(y^2)}$

```
se <- function(x) sqrt(sum((x - mean(x)) ^ 2)) / length(x)
se_x<-se(x)*se(x)
se_y<-se(y)*se(y)
sqrt(se_x + se_y)
```

```
## [1] 26.90811
```

4. ¿Dirías que el tratamiento incrementó la supervivencia de los ratones? Si
5. Supongamos que deseamos comparar los grupos usando las medianas en lugar de las medias, estima la diferencia de las medias y usa bootsrtap para estimar el error estándar de la diferencia. ¿Cuál es tu conclusión de los resultados?

```
median_x<-median(x)
median_y<-median(y)
result <- median_x - median_y
print(result)
```

```
## [1] 48
```

```
medianBoot <- function(x, y){
  n <- length(x)
  n2<- length(y)
  muestra_boot_x <- sample(x, size = n, replace = TRUE)
  muestra_boot_y <- sample(y, size = n2, replace = TRUE)
  median_bx <- median(muestra_boot_x)
  median_by <- median(muestra_boot_y)
  result <- median_bx - median_by
  return(result)
}

thetas_boot <- rerun(1000, medianBoot(x,y)) %>% flatten_dbl()
sd(thetas_boot)
```

```
## [1] 39.13482
```

Podemos ver que el error estándar es más grande