MÉTODOS ESTADÍSTICOS

Nombre(s): Equipo 4:
- Diana Zepeda Martínez
- José Juan García Romero

Nº 20	Descripción: La base de datos UsingR::brightness cuenta con el registro de la brillantez de 966 estrellas. Aplicando el concepto BootStrap, inferir la media de la población realizando:

a) 3 muestras simples de tamaño 50, 300, 600

Actividad 20.R

Source on Save

```r
# Diana Zepeda Martinez
# José Juan García Romero

library(boot)
library(dplyr)

# INCISO A

brillo = UsingR::brightness
mediaP = mean(brillo)

muestra1 = sample(brillo,replace = T,size = 50)
media1 = mean(muestra1)

muestra2 = sample(brillo,replace = T,size = 300)
media2 = mean(muestra2)

muestra3 = sample(brillo,replace = T,size = 600)
media3 = mean(muestra3)

medias = rbind(mediaP,media1,media2,media3)
medias

```

Console | Terminal | Jobs

R 4.1.3 · ~/Universidad/Métodos Estadisticos/R/

```
> # Diana Zepeda Martinez
> # José Juan García Romero
> # INCISO A
>
> library(boot)
> library(dplyr)
> brillo = UsingR::brightness
> mediaP = mean(brillo)
> muestra1 = sample(brillo,replace = T,size = 50)
> media1 = mean(muestra1)
> muestra2 = sample(brillo,replace = T,size = 300)
> media2 = mean(muestra2)
> muestra3 = sample(brillo,replace = T,size = 600)
> media3 = mean(muestra3)
> medias = rbind(mediaP,media1,media2,media3)
> medias
           [,1]
mediaP 8.417743
media1 8.508200
media2 8.378333
media3 8.373267
>
```

Data	
medias	num [1:4, 1] 8.42 8.37 8.4 8.39
values	
brillo	num [1:966] 9.1 9.27 6.61 8.06 8.55 ...
media1	8.3702
media2	8.40246666666667
media3	8.3886
mediaP	8.41774327122153
muestra1	num [1:50] 5.78 7.09 8.01 9.99 9.06 8.24 8.27 8.31 8.11 7.46 ...
muestra2	num [1:300] 9.77 6.18 9.05 8.56 8.75 ...
muestra3	num [1:600] 8.33 9.37 8.77 10.82 7.91 ...

b) Implementando una función obtener 200 muestras

```r
# Diana Zepeda Martinez
# José Juan García Romero
# INCISO B

medias = NULL
for(i in 1:200){
  muestra = sample(brillo, replace = T)
  medias[i] = mean(muestra)
}

mediaP
muestra200 = mean(medias)
muestra200

```

Console | Terminal | Jobs

R 4.1.3 · ~/Universidad/Métodos Estadisticos/R/

```
> # Diana Zepeda Martinez
> # José Juan García Romero
> # INCISO B
>
> medias = NULL
> for(i in 1:200){
+   muestra = sample(brillo, replace = T)
+   medias[i] = mean(muestra)
+ }
> mediaP
[1] 8.417743
> muestra200 = mean(medias)
> muestra200
[1] 8.413804
>
```

values	
brillo	num [1:966] 9.1 9.27 6.61 8.06 8.55 ...
i	200L
media1	8.3702
media2	8.40246666666667
media3	8.3886
mediaP	8.41774327122153
medias	num [1:200] 8.43 8.5 8.47 8.38 8.4 ...
muestra	num [1:966] 9.16 7.96 8.3 8.76 8.55 8.8 7.78 8.34 8.27 8.95 ...
muestra1	num [1:50] 5.78 7.09 8.01 9.99 9.06 8.24 8.27 8.31 8.11 7.46 ...
muestra2	num [1:300] 9.77 6.18 9.05 8.56 8.75 ...
muestra200	8.41309922360248
muestra3	num [1:600] 8.33 9.37 8.77 10.82 7.91 ...

c) Con la función boot:

c.1) obtener 100,000 réplicas

c.2) obtener sus intervalos de confianza al 99%

c.3) graficar las medias de las réplicas obtenidas en el c.1 y los intervalos del c.3

```r
# Diana Zepeda Martinez
# José Juan García Romero
# INCISO C
library(boot)

func_media <- function(m,d){ # v:datos d:indice
  (mean(m[d]))
}

boot_media <- boot(brillo, func_media, R = 100000)
media_boot <- mean(boot_media$t)
media_pob <- boot_media$t0
tabla = rbind(media_pob,media_boot)
tabla

# INCISO C2

IC <- boot.ci(boot_media,conf = 0.99,type = "norm")
IC

# INCISO C3

hist(boot_media$t,freq = F,ylim = c(0,12))
abline(v = boot_media$t0,col = "purple",lwd = 3)
abline(v = IC$normal[2],col = "green",lwd = 2,lty = 5)
abline(v = IC$normal[3],col = "blue",lwd = 2,lty = 5)
lines(density(boot_media$t),col = "red",lwd = 2)
```

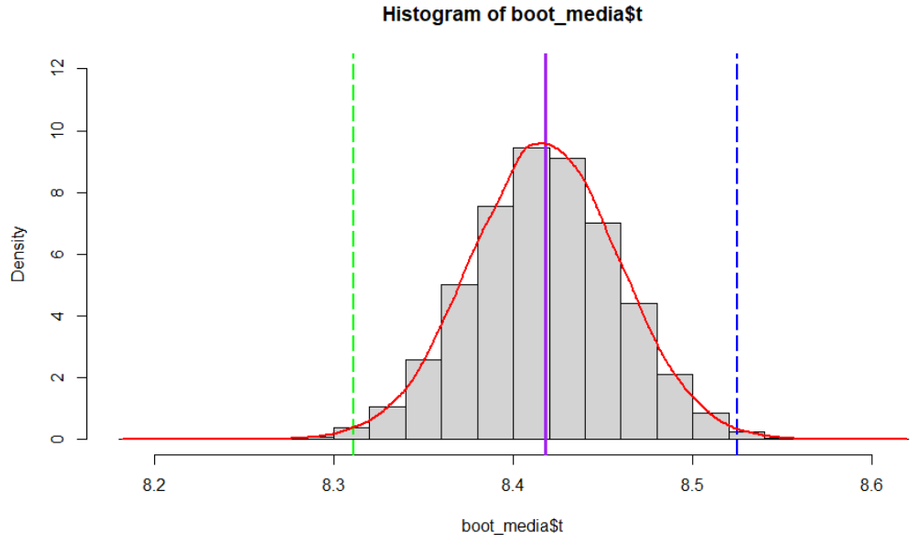
Console | Terminal | Jobs

R 4.1.3 · ~/Universidad/Métodos Estadisticos/R/

```
> # Diana Zepeda Martinez
> # José Juan García Romero
> # INCISO C
>
> # INCISO C1
>
> # Diana Zepeda Martinez
> # José Juan García Romero
> # INCISO C
> library(boot)
> func_media <- function(m,d){ # v:datos d:indice
+   (mean(m[d]))
+ }
> boot_media <- boot(brillo, func_media, R = 100000)
> media_boot <- mean(boot_media$t)
> media_pob <- boot_media$t0
> tabla = rbind(media_pob,media_boot)
> tabla
                [,1]
media_pob  8.417743
media_boot 8.417947
>
> # INCISO C2
>
> IC <- boot.ci(boot_media,conf = 0.99,type = "norm")
> IC
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 100000 bootstrap replicates

CALL :
boot.ci(boot.out = boot_media, conf = 0.99, type = "norm")

Intervals :
Level      Normal
99%   ( 8.310,  8.525 )
Calculations and Intervals on Original Scale
>
> # INCISO C3
>
> hist(boot_media$t,freq = F,ylim = c(0,12))
> abline(v = boot_media$t0,col = "purple",lwd = 3)
> abline(v = IC$normal[2],col = "green",lwd = 2,lty = 5)
> abline(v = IC$normal[3],col = "blue",lwd = 2,lty = 5)
> lines(density(boot_media$t),col = "red",lwd = 2)
>
```



Data	
boot_media	Large boot (11 elements, 839.1 kB)
IC	List of 4
tabla	num [1:2, 1] 8.42 8.42
Values	
brillo	num [1:966] 9.1 9.27 6.61 8.06 8.55 ...
i	200L
media_boot	8.41777374171843
media_pob	8.41774327122153
media1	8.3702
media2	8.40246666666667
media3	8.3886
mediaP	8.41774327122153
medias	num [1:200] 8.43 8.5 8.47 8.38 8.4 ...
muestra	num [1:966] 9.16 7.96 8.3 8.76 8.55 8.8 7.78 8.34 8.27 8.95 ...
muestra1	num [1:50] 5.78 7.09 8.01 9.99 9.06 8.24 8.27 8.31 8.11 7.46 ...
muestra2	num [1:300] 9.77 6.18 9.05 8.56 8.75 ...
muestra200	8.41309922360248
muestra3	num [1:600] 8.33 9.37 8.77 10.82 7.91 ...
Functions	
func_media	function (m, d)

d) Con la función bootstrap obtener 100,000 réplicas

```
66  # Diana Zepeda Martinez
67  # José Juan García Romero
68  # INCISO D
69  library(bootstrap)
70
71  funcion <- function(x){mean(x)}
72
73  replicas_bootstrap <- bootstrap(brillo,100000,funcion)
74  mediaP
75  media_bootstap <- mean(replicas_bootstrap$thetastar)
76  media_bootstap
77
```

Console Terminal Jobs

R 4.1.3 · ~/Universidad/Métodos Estadisticos/R/

```
> # Diana Zepeda Martinez
> # José Juan García Romero
> # INCISO D
>
> # Diana Zepeda Martinez
> # José Juan García Romero
> # INCISO D
> library(bootstrap)
> funcion <- function(x){mean(x)}
> replicas_bootstrap <- bootstrap(brillo,100000,funcion)
> mediaP
[1] 8.417743
> media_bootstap <- mean(replicas_bootstrap$thetastar)
> media_bootstap
[1] 8.417823
>
```

Data	
boot_media	Large boot (11 elements, 839.1 kB)
IC	List of 4
replicas_bootstrap	Large list (5 elements, 801.3 kB)
tabla	num [1:2, 1] 8.42 8.42
Values	
brillo	num [1:966] 9.1 9.27 6.61 8.06 8.55 ...
i	200L
media_boot	8.41777374171843
media_bootstap	8.41782250879917
media_pob	8.41774327122153
media1	8.3702
media2	8.40246666666667
media3	8.3886
mediaP	8.41774327122153
medias	num [1:200] 8.43 8.5 8.47 8.38 8.4 ...
muestra	num [1:966] 9.16 7.96 8.3 8.76 8.55 8.8 7.78 8.34 8.27 8.95 ...
muestra1	num [1:50] 5.78 7.09 8.01 9.99 9.06 8.24 8.27 8.31 8.11 7.46 ...
muestra2	num [1:300] 9.77 6.18 9.05 8.56 8.75 ...
muestra200	8.41309922360248
muestra3	num [1:600] 8.33 9.37 8.77 10.82 7.91 ...
Functions	
func_media	function (m, d)
funcion	function (x)

e) Realizar una tabla comparativa de medias y una conclusión sobre esta.

```
78  # Diana Zepeda Martínez
79  # José Juan García Romero
80  # INCISO E
81
82  tabla_general = rbind(mediaP,media1,media2,media3,muestra200,media_boot,media_bootstap)
83  rownames(tabla_general) = c("poblacion","muestra1","muestra2","muestra3","muestra200",
84                              "media_boot","media_bootstap")
85  colnames(tabla_general) = c("medias")
86  tabla_general
87
```

Console | Terminal | Jobs

R 4.1.3 · ~/Universidad/Métodos Estadísticos/R/

```
> # Diana Zepeda Martínez
> # José Juan García Romero
> # INCISO E
>
> tabla_general = rbind(mediaP,media1,media2,media3,muestra200,media_boot,media_bootstap)
> rownames(tabla_general) = c("poblacion","muestra1","muestra2","muestra3","muestra200","media_boot","media_bootstap")
> colnames(tabla_general) = c("medias")
> tabla_general
                 medias
poblacion      8.417743
muestra1       8.370200
muestra2       8.402467
muestra3       8.388600
muestra200     8.413099
media_boot     8.417774
media_bootstap 8.417823
>
```

Data	
boot_media	Large boot (11 elements, 839.1 kB)
IC	List of 4
replicas_bootstrap	Large list (5 elements, 801.3 kB)
tabla	num [1:2, 1] 8.42 8.42
tabla_general	num [1:7, 1] 8.42 8.37 8.4 8.39 8.41 ...
Values	
brillo	num [1:966] 9.1 9.27 6.61 8.06 8.55 ...
i	200L
media_boot	8.41777374171843
media_bootstap	8.41782250879917
media_pob	8.41774327122153
media1	8.3702
media2	8.40246666666667
media3	8.3886
mediaP	8.41774327122153
medias	num [1:200] 8.43 8.5 8.47 8.38 8.4 ...
muestra	num [1:966] 9.16 7.96 8.3 8.76 8.55 8.8 7.78 8.34 8.27 8.95 ...
muestra1	num [1:50] 5.78 7.09 8.01 9.99 9.06 8.24 8.27 8.31 8.11 7.46 ...
muestra2	num [1:300] 9.77 6.18 9.05 8.56 8.75 ...
muestra200	8.41309922360248
muestra3	num [1:600] 8.33 9.37 8.77 10.82 7.91 ...
Functions	
func_media	function (m, d)
funcion	function (x)

f) Describir las nuevas funciones
boot: Es la biblioteca de arranque y con esta se generan réplicas. **bootstrap:** Se obtiene la distribución del estadístico que se va a usar. **hist:** Crea histogramas, gráficas, etc. **abline:** Pinta líneas dentro de un histograma. **lines:** pinta una línea de tendencia en el histograma. **rownames:** Sirve para definir nombres de las filas de una tabla. **colnames:** Sirve para definir los nombres de las columnas de una tabla. **UsingR:** Es para usar una colección de datos, en este caso brightness. **rbind:** Sirve para unir varios arreglos de datos.