

Hoja 7 de problemas y prácticas con R

Estadística Computacional I. Grado en Estadística

Marta Venegas Pardo

Contents

1	Ejercicio 1	1
2	Ejercicio 2	1
3	Ejercicio 3	1
4	Ejercicio 4	2
5	Ejercicio 5	2
6	Ejercicio 6	2
7	Ejercicio 7	2
8	Ejercicio 8	2

1 Ejercicio 1

1. Realizar un contraste bootstrap unilateral de hipótesis para comparar las desviaciones típicas a partir de las siguientes muestras: $x=c(137.9, 143, 143.2, 140, 140.2, 139.3, 141.4, 140.1, 142, 137.2, 139.5, 142.7, 141.3)$

$y=c(141.6, 138.9, 140, 141.9, 140.5, 138.6, 141.5, 141.5, 140.7, 141.5, 140.4, 142, 141)$

$$H_0 : \sigma_1 \leq \sigma_2$$

Escribir las instrucciones R sin y con la librería boot.

2 Ejercicio 2

2. Realizar mediante procedimientos bootstrap el contraste de hipótesis unilateral relativo a la posibilidad de que la primera componente principal del fichero iris de R explique más del 70% de la varianza total. Hacerlo directamente y con la ayuda de **boot**.

3 Ejercicio 3

3. Las siguientes instrucciones permiten definir dos variables en R que contienen las medidas de la corrosión (y) en 13 aleaciones de níquel-cobre, cada una de ellas con un contenido de hierro x. Es de interés el cambio en la corrosión cuando aumenta el nivel de hierro, comparado con la pérdida de corrosión cuando no hay hierro: β_1/β_0 en el modelo de regresión lineal.

Calcular intervalos de confianza bootstrap para dicho cociente.

`x<- c(0.01,0.48,0.71,0.95,1.19,0.01,0.48,1.44,0.71,1.96,0.01,1.44,1.96)`

`y<- c(127.6,124,110.8,103.9,101.5,130.01,122,92.3,113.1,83.7,128,91.4,86.2)`

Se sabe que un estimador de la varianza de β_1/β_0 mediante el método delta es:

$$\left(\frac{\hat{\beta}_1}{\hat{\beta}_0}\right)^2 \left(\frac{\hat{v}(\hat{\beta}_1)}{\hat{\beta}_1^2} + \frac{\hat{v}(\hat{\beta}_0)}{\hat{\beta}_0^2} - \frac{2cov(\hat{\beta}_0, \hat{\beta}_1)}{\hat{\beta}_0 \hat{\beta}_1} \right)$$

La función utilizada para calcular el estadístico de interés debe incluir también esta estimación de su varianza, para usarla en el método bootstrap-t.

4 Ejercicio 4

- El fichero “salmon.dat” contiene datos anuales sobre una población de salmones. Las variables que aparecen son: R = “recruits” número de salmones que entran, y S = “spawners” número de salmones que están poniendo huevos, que mueren en cuanto lo hacen. Para que la población se estabilice se requiere: R = S (en otro caso, o hay demasiados salmones para los mismos recursos, o bien no se repone la población). Se desea calcular un I.C. bootstrap para el punto donde R = S, trabajando con el modelo de Beverton- Holt:

$$R = \frac{1}{\beta_0 + \beta_1/s}$$

5 Ejercicio 5

- Diseñar un estudio empírico para analizar la efectividad de la transformación Z de Fisher en el cálculo de intervalos de confianza para el coeficiente de correlación lineal poblacional. Por ejemplo, generar 100 muestras de tamaño 10 de una Normal bivalente de media 0, desviaciones típicas 1 y correlación 0.6 y calcular los I.C. al 95% con los métodos Percentil, Normal y BCa. Comparar los cubrimientos y las longitudes medias utilizando procedimientos numéricos y gráficos. 1

6 Ejercicio 6

- Estimar el sesgo de la razón (cociente de las medias de las variables x y u) en el fichero “city” de R mediante el bootstrap balanceado (fichero disponible en la librería “boot”):
 - Escribiendo directamente las instrucciones.
 - Empleando la librería boot.

7 Ejercicio 7

- Estimar el error de clasificación para el modelo de análisis discriminante lineal sobre los datos “cats” de la librería MASS.

8 Ejercicio 8

- Estimar el error de predicción para los datos “Renta.txt”, siendo “rentsqm” (precio del alquiler por m2) la variable dependiente de un modelo de regresión lineal. Utilizar el criterio RECM.