

BOOTSTRAP

Ejemplo:

Considere una muestra aleatoria de tamaño $n=10$ con las siguientes observaciones: $x_1 = -2.41$, $x_2 = 4.86$, $x_3 = 6.06$, $x_4 = 9.11$, $x_5 = 10.2$, $x_6 = 12.81$, $x_7 = 13.17$, $x_8 = 14.1$, $x_9 = 15.77$, $x_{10} = 15.79$, cuya media y error estándar son $\bar{x} = 9.946$ y $s = 5.46$ respectivamente. La siguiente tabla ilustra el procedimiento Bootstrap para $B = 10$.

Muestras	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x^*	$x^* - \bar{x}^*$
1	1	0	1	0	3	0	1	1	2	1	10.885	0.318
2	0	1	0	0	2	0	1	2	1	3	12.977	2.41
3	2	3	2	1	0	0	1	0	1	0	5.993	-4.574
4	1	0	1	0	3	1	1	1	1	1	10.589	0.022
5	1	1	0	2	1	1	1	1	2	0	10.249	-0.318
6	0	2	0	2	0	2	2	0	2	0	11.144	0.577
7	0	0	1	2	1	1	2	0	3	0	12.094	1.527
8	0	0	0	2	1	2	2	2	1	0	12.435	1.868
9	0	1	2	0	0	1	0	2	2	2	12.111	1.544
10	2	2	2	0	1	0	1	0	1	1	7.195	-3.372

Donde:

Media muestral $\bar{x} = 9.946$

Media muestral Bootstrap $\Rightarrow \bar{x}^* = \frac{1}{B} \sum_{i=1}^B x_i^* \Rightarrow \bar{x}^* = 10.567$

Error estándar de $\bar{x} = 5.4619$

Error estándar de las $\bar{x}^* \Rightarrow s = \sqrt{\frac{1}{B-1} \sum_{i=1}^B (x_i^* - \bar{x}^*)^2} \Rightarrow 2.2813$

Intervalo de Confianza Bootstrap Básico

Para su construcción se obtienen B muestras bootstrap y se calculan los correspondientes cuantiles de $\hat{\theta}^*$. Así se tiene:

$$\left[2\bar{x} - \hat{\theta}_B^{*(\alpha)}, 2\bar{x} - \hat{\theta}_B^{*(1-\alpha)} \right]$$

donde $\hat{\theta}_B^{*(\alpha)}$ es el α -ésimo percentil de la réplicas B bootstrap y $\hat{\theta}_B^{*(1-\alpha)}$ es el $(1-\alpha)$ -ésimo percentil de las B réplicas bootstrap.

Ejemplo: Para los datos anteriores, se calculan los límites de confianza bootstrap básico. Para esto se usó $B = 1000$, $n = 10$ muestras aleatorias generadas de la FDE con media $\bar{x}=9.9460$. Los valores ordenados de x^* son el $1000(0.025)$ -ésimo y el $1000(0.975)$ -ésimo, es decir corresponden a la posición 25-ésima y 975 -ésima, los cuales tienen un valor de 6.360 y 13.234 respectivamente. Por lo tanto, los límites de confianza del 95% obtenidos son:

$$2 \times 9.9460 - 13.2340 = 6.658, \quad 2 \times 9.9460 - 6.3600 = 13.532$$

Así que el intervalo de confianza aproximado será $[6.658, 13.532]$

Intervalo de Confianza Percentil

Se generan B muestras bootstrap y a partir de estas se obtiene una estimación de la función de distribución acumulada \hat{F} de $\hat{\theta}^*$, el intervalo $1-2\alpha$ percentil se define por los percentiles α y $1-\alpha$ de \hat{F} . Primero se procede a generar B conjuntos de muestras bootstrap independientes $x_1^*, x_2^*, \dots, x_B^*$ y luego se calculan las replicas bootstrap.

Sea $\hat{\theta}^{*(\alpha)}$ el α -ésimo cuantil empírico de los valores de $\hat{\theta}^*(b)$, esto es el $B\alpha$ -ésimo valor en la lista ordenada de las B repeticiones de $\hat{\theta}^*$. De la misma forma $\hat{\theta}_B^{*(1-\alpha)}$ es el $(1-\alpha)$ -ésimo percentil empírico.

El intervalo $(1-2\alpha)$ percentil aproximado es:

$$\left[\hat{\theta}_{\text{inf}}, \hat{\theta}_{\text{sup}} \right] \approx \left[\hat{\theta}_B^{*(\alpha)}, \hat{\theta}_B^{*(1-\alpha)} \right]$$

Ejemplo: Continuando con los datos del ejemplo anterior, se tiene que para $B=1000$ y $\alpha=0.05$ el $\theta_{1000}^{*(0.05)} = 6.360$, el 50-ésimo valor ordenado de las 1000 réplicas bootstrap y $\theta_{1000}^{*(0.95)}=13.234$, el 950-ésimo valor ordenado de las 1000 réplicas bootstrap.

Asimismo, el intervalo de confianza percentil será $[6.360,13.234]$, bastante cerca del intervalo bootstrap básico.