

PROBABILIDAD Y ESTADÍSTICA B - CURSO 27

ENTREGA 2 - INTERVALOS DE CONFIANZA CON BOOTSTRAP

En este ejercicio pondremos a prueba el método *bootstrap*, que consiste en realizar remuestreos de una muestra aleatoria para estudiar estimadores cuya distribución es difícil de obtener o no se conoce. El trabajo práctico se desarrollará en R. El grupo deberá entregar un archivo .R con el código y una carilla en .pdf resumiendo brevemente los resultados.

Punto de partida: Tenemos una muestra aleatoria (X_1, X_2, \dots, X_n) de variables *iid* de una distribución que se supone desconocida. Nos interesa estimar la mediana m_e de dicha distribución. Un posible estimador es:

$$\hat{m}_e = \min\{X_i : \hat{F}_n(X_i) \geq 0.5\},$$

donde \hat{F}_n es la función de distribución empírica asociada a la muestra. Queremos usar este estimador para obtener intervalos de confianza de nivel aproximado $1 - \alpha$ para la verdadera mediana, con métodos basados en *bootstrap*. Usando las ayudas vistas en clase:

- Armar la función `intervalo_boot` que, tomando como argumentos un vector numérico `x` y una variable `nivel`, realice $B = 1000$ estimaciones de la mediana con *bootstrap* y que devuelva los límites de confianza con el nivel deseado. Cada grupo debe elegir qué metodología usar para calcular los límites y describirla en el informe.
- Armar la función `cubrimiento_empirico` que, teniendo como único argumento el tamaño de muestra `n`, simule 1000 muestras aleatorias de tamaño `n` con distribución provenientes de una distribución F_Θ asignada al grupo con los correspondientes parámetros. Calcule para cada una los límites del intervalo de confianza para la mediana y devuelva una lista de dos valores: La proporción de los 1000 intervalos que contienen a la verdadera mediana, y la longitud promedio de los intervalos.
- Para tres tamaños de muestra $n = 10$, $n = 100$ y $n = 500$ usar las funciones programadas anteriormente para analizar el cubrimiento empírico y la longitud promedio de los intervalos. Organizar los resultados en una tabla y realizar conclusiones.