PROBABILIDADES Y ESTADÍSTICA (C)

SIMULACIÓN PRÁCTICA 7 - PARTE 1

Sea $(X_i)_{i\geq 1}$ una muestra aleatoria con distribución F. Denotemos con X a un elemento con misma distribución que X_i. Asuma que estamos interesados en estimar el la probabilidad de que X sea mayor a uno: $\theta(F) := \mathbb{P}_F(X > 1)$.

Estimador 1:

1.1 Proponga un estimador $\widehat{\theta}_n$ consistente para $\theta(F) = \mathbb{P}_F(X > 1)$.

1.2 Implemente una función **est1** que tenga por argumento un conjunto de datos $(x_1, \ldots, x_n)$ y devuelva el valor de la estimación obtenida utilizando $\widehat{\theta}_n$.

1.3 Calcule el valor de $\widehat{\theta}_n$ en el siguiente conjunto de datos:

$$12.23 \quad 6.37 \quad 6.10 \quad 0.70 \quad 3.48 \quad 2.82 \quad 9.55 \quad 2.21 \quad 0.72 \quad 9.09.$$

Mundo Exponencial: Calentando motores

1.4 Sea X una variabe aleatoria con distribución F, exponencial de parámetro $\lambda = 0.2$: $X \sim \mathcal{E}(0.2)$. Indique el valor de

$$\mathbb{E}(X) = \ldots\ , \quad \mathbb{V}(X) = \ldots\ , \quad \mathbb{P}(X > 1) = \ldots\ , \quad \text{cuando } X \sim \mathcal{E}(0.2).$$

1.5 Sea ahora X una variabe aleatoria con distribución F pertenecinete a la familia exponencial: es decir, $X \sim \mathcal{E}(\lambda)$ con λ DESCONOCIDO. Exprese cada uno de los siguientes objetos en función de λ:

$$\mathbb{E}(X) = \ldots\ , \quad \mathbb{V}(X) = \ldots\ , \quad \mathbb{P}(X > 1) = \ldots\ , \quad \text{cuando } X \sim \mathcal{E}(\lambda).$$

Mundo Exponencial: Haciendo Estadística

Sean $(X_i)_{i\geq 1}$ i.i.d., con misma distribución que X. Asuma ahora que F pertenece a la familia exponencial; es decir, $X \sim \mathcal{E}(\lambda)$, con λ DESCONOCIDO.

1.6 Proponga un nuevo estimador $\widetilde{\theta}_n$ consistente para $\theta(F) = \mathbb{P}_F(X > 1)$ bajo este nuevo escenario. Es decir, defina $\widetilde{\theta}_n = f_n(X_1, \ldots, X_n)$ de forma tal que

$$\widetilde{\theta}_n = f_n(X_1, \ldots, X_n) \xrightarrow{p} e^{-\lambda} \text{ cuando } X_i \sim \mathcal{E}(\lambda)\ , \forall\lambda > 0.$$

1.7 Implemente una función **est2** que tenga por argumento un conjunto de datos $(x_1, \ldots, x_n)$ y devuelva el valor de la estimación obtenida utilizando $\widetilde{\theta}_n$.

1.8 Calcule el valor de $\widetilde{\theta}_n$ en el siguiente conjunto de datos:

$$12.23 \quad 6.37 \quad 6.10 \quad 0.70 \quad 3.48 \quad 2.82 \quad 9.55 \quad 2.21 \quad 0.72 \quad 9.09.$$

Simulación 1:. A lo largo de esta simulación generaremos variables con distribución exponencial de paramétro $\lambda = 0.2$.

1.9 Indique cual es el veradero valor que estamos queriendo estimar: $\theta_0 = \mathbb{P}(X > 1)$, siendo $X \sim \mathcal{E}(0.2)$.

1.10 Genere un conjunto de n=50 datos y calcule cada uno de los estimadores.

1.11 Genere $Nrep$= 1000 conjunto de datos de tamaño n=50 y guarde los valores de cada uno de los dos estimadores calculados en cada uno de los $Nrep = 1000$ conjuntos de datos.

1.12 Realize un histograma de cada uno de los estimadores propuestos con los valores obtenidos en el item anterior. Comente los gráficos realizados. Indique que etimador prefiere en este escenario y explique a que atribuye sus bondades.

Recuerde que el error cuadrático medio de un estimador $\widehat{\theta}_n = \widehat{\theta}_n(X_1, \ldots, X_n)$ está dado por

$$\text{ECM} = \mathbb{E}\{(\widehat{\theta}_n - \theta)^2\}. \tag{0.1}$$

Para obtener el ECM en el punto θ necesitamos conocer la distribución de $\widehat{\theta}_n$. Sin embargo, cuando simulamos y generamos datos, podemos estimar el ECM con su versión empírica,(ECME) haciendo

$$\text{ECME} = \frac{1}{Nrep} \sum_{i=1}^{Nrep} (\widehat{\theta}_{n,i} - \theta)^2, \tag{0.2}$$

siendo que $\widehat{\theta}_{n,i}$ la estimación obtenida en la i-ésima replicación.

1.13 Represente en una tabla el error cuadrático medio (estimado) de los estimadores $\widehat{\theta}_n$ y $\widetilde{\theta}_n$ para muestras de tamaño n=150, n=200, n=500 y n=1000, utilizando Nrep=1000 replicaciones en cada caso. ¿Qué estimador prefiere bajo este escenario?

Mundo Normal: Ojo al Piojo! Considere ahora variables aleatorias X_i i.i.d. con distribución normal de media $\mu = 1/0.2$ y $\sigma^2 = 1/0.2^2$.

1.14 Calcule la probabilidad de que X_i supere el valor 1: $\mathbb{P}(X_i > 1)$

1.15 Calcule el valor de cada uno de los siguientes límites:

$$\lim_{n\to\infty} \widehat{\theta}_n(X_1, \ldots, X_n) = \ldots\ldots , \quad \lim_{n\to\infty} \widetilde{\theta}_n(X_1, \ldots, X_n) = \ldots\ldots$$

1.16 Propongo un nuevo estimador $\theta_n^* = \theta_n^*(X_1, \cdots, X_n)$ para $\theta(F) = \mathbb{P}_F(X_i > 1)$, asumiendo asumiendo ahora que F pertenece a la normal: $X_i \sim \mathcal{N}(\mu, \sigma^2)$.

Simulacion 2: A lo largo de esta simulación generaremos variables con distribución normal de media $\mu = 1/0.2$ y $\sigma^2 = 1/0.2^2$. Represente en una tabla el error cuadrático medio (estimado) de los estimadores $\widehat{\theta}_n$, $\widetilde{\theta}_n$ y θ_n^* para muestras de tamaño n=150, n=200, n=500 y n=1000, utilizando Nrep=1000 replicaciones en cada caso. Analice los resultados obtenidos y expliqye que estimador elegiría bajo este escenario.