Point estimation refers to providing a single "best guess" of some quantity of interest.
All of statistics. Wasserman

1 Estimación bajo modelo uniforme $\mathcal{U}[0, \theta]$:

Sean $(X_i)_{i\geq 1}$ variables aleatorias independientes identicamente distribuídas, con distribución uniforme en el intervalo $[0, \theta]$: $X_i \sim \mathcal{U}[0, \theta]$. Consideremos los siguientes estimadores de θ basados en una muestra $X_1, \dots, X_n$:

$$\widehat{\theta}_n = \widehat{\theta}_n(X_1, \dots, X_n) = 2\bar{X}_n \; , \quad \widetilde{\theta}_n = \widetilde{\theta}_n(X_1, \dots, X_n) = \max\{X_1, \dots, X_n\}. \tag{1}$$

1. Implemente las funciones `est1` y `est2` que tengan por argumento un conjunto de `datos` $(x_1, \dots, x_n)$ y devuelva el valor de la estimación $\widehat{\theta}_n(x_1, \dots, x_n)$ y $\widetilde{\theta}_n(x_1, \dots, x_n)$, para los estimadores definidos en (1), respectivamente.

2. Calcule el valor de los estimadores **est1** y **est2** en los datos

   `1.17 1.75 0.28 2.56 2.36 0.36 1.82 0.24 1.17 1.86`

3. Calcule el valor de los estimadores **est1** y **est2** en los datos

   `0.66 0.07 0.62 0.65 1.33 0.40 1.17 1.11 2.01 2.98`

   Simulación 1. A lo largo de esta simulación generaremos variables con distribución uniforme en el intervalo $[0, 3]$. Es decir, trabajaremos con v.a. $(X_i)_{i\geq 1}$ i.i.d., $X_i \sim \mathcal{U}[0, \theta]$ con $\theta = 3$.

4. Realice histogramas para emular la distribuciòn de cada uno de los estimadores con $n = 5$, $n = 30$, $n = 50$, haciendo $Nrep = 1000$ replicaciones. Comente las principales características que observa en los gráficos. Diría usted que la distribución de $\widehat{\theta}_n$ (est1) es aproximadamente normal? Diría usted que la distribución de $\widetilde{\theta}_n$ (est2) es aproximadamente normal?

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

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

Para obtener el ECM 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

$$\text{ECME} = \frac{1}{Nrep} \sum_{k=1}^{Nrep} (\widehat{\theta}_{n,k} - \theta)^2, \tag{3}$$

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

5. Presente en una tabla el error cuadrático medio empírico de los estimadores $\widehat{\theta}_n$ y $\widetilde{\theta}_n$ para muestras de tamaño $n = 5$, $n = 30$, $n = 50$, $n = 100$ y 500, utilizando $Nrep = 1000$ replicaciones en cada caso. Qué estimador elegiría?

2 ¿A medida?

Sea $(X_i)_{i\geq 1}$ una muestra aleatoria con distribucin F. Denotemos con X a un elemento con misma distribucin 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 funcin **est1** que tenga por argumento un conjunto de datos $(x_1, \ldots, x_n)$ muestra 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 distribucin 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 $\widehat{\theta}_n$ 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 funcin **est2** que tenga por argumento un conjunto de datos $(x_1, \ldots, x_n)$ muestra 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 6.37 6.10 0.70 3.48 2.82 9.55 2.21 0.72 9.09.

Simulacin 1:. A lo largo de esta simulacin generaremos variables con distribucin exponencial de paramtro $\lambda = 0.2$.

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 una muestras de tamao n=50 y calcule cada uno de los estimadores.

1.11 Genere $Nrep$= 1000 muestras de tamao 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.

1.13 Represente en una tabla el error cuadrtico medio (estimado) de los estimadores $\widehat{\theta}_n$ y $\widetilde{\theta}_n$ para muestras de tamao 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 distribucin 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 lmites:

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

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 simulacin generaremos variables con distribucin normal de media $\mu = 1/0.2$ y $\sigma^2 = 1/0.2^2$. Represente en una tabla el error cuadrtico medio (estimado) de los estimadores $\widehat{\theta}_n$, $\widetilde{\theta}_n$ y θ_n^* para muestras de tamao n=150, n=200, n=500 y n=1000, utilizando Nrep=1000 replicaciones en cada caso. Analice los resultados obtenidos y expliqye que estimador elegira bajo este escenario.