

Trabajo Práctico I. Inferencia Estadística

Aclaraciones:

- El TP se puede hacer en grupos de **tres o cuatro personas**.
- La resolución debe resentarse en formato notebook y debe incluir razonamientos, gráficos y código utilizado. El código debe correr sin problemas
- La fecha límite de entrega es el **Viernes 03 de Junio**.

Sean X_1, \dots, X_n una muestra aleatoria con distribución $\mathcal{U}[0, b]$ con b un parámetro desconocido.

1. Para cada uno de los siguientes estimadores implementar una función que tome como argumento una muestra y devuelva el valor estimado de b

$$\hat{b}_{mom} = 2\bar{X}$$

$$\hat{b}_{mv} = \max\{X_1, \dots, X_n\}$$

$$\hat{b}_{med} = 2 \times \text{mediana}\{X_1, \dots, X_n\}.$$

2. Utilizando $b = 1$, sortear una muestra de tamaño $n = 15$. Calcular cada uno de los estimadores con la muestra obtenida y reportar el valor de cada estimador y el error de estimación.
3. Hacer una simulación para obtener el sesgo, varianza y error cuadrático medio (ECM) de cada uno de los estimadores. Para lograr esto:
 - a) Generar una muestra con $b = 1$, $n = 15$.
 - b) Para la muestra obtenida, calcular \hat{b}_{mv} , \hat{b}_{mom} , \hat{b}_{med} y almacenar los resultados.
 - c) Repetir $N_{rep} = 1000$ veces los pasos (a) y (b).
 - d) Obtener una aproximación del sesgo restando el valor verdadero de b a la media muestral de cada estimador.
 - e) Obtener la aproximación de la varianza a partir de la varianza muestral de cada estimador.
 - f) Obtener la aproximación del ECM a través de la fórmula que lo relaciona con el sesgo y la varianza.
4. Implementar las funciones *simulacion_mv(b, n)*, *simulacion_mom(b, n)* y *simulacion_med(b, n)* que devuelven una aproximación del sesgo y de la varianza de cada uno de los estimadores correspondientes al b y al n .
5. Comparar mediante gráficos, el sesgo, la varianza y el ECM de cada estimador con $n = 15$ y una grilla de valores con $0 < b < 2$. ¿Qué observa? ¿Qué estimador elige?
6. Realizar un gráfico de los ECM con $b = 1$ y $n = 15, 30, 60, 120, 240$. ¿Qué observa? ¿Qué estimador elige?

7. Calcular los estimadores en la siguiente muestra. ¿Observa algo extraño? ¿A qué cree que se debe?

0.917, 0.247, 0.384, 0.530, 0.798, 0.912, 0.096, 0.684,
0.394, 20.1, 0.769, 0.137, 0.352, 0.332, 0.670

8. Aproximar sesgo, varianza y error cuadrático medio para los estimadores bajo el siguiente escenario con datos contaminados:

Una muestra aleatoria uniforme con $b = 1$ y $n = 15$ que con probabilidad $p = 0.05$ un elemento de la muestra se multiplica por 100 (coma corrida dos lugares a la derecha).

- a) ¿Qué estimador prefiere en este escenario?
b) Explique en sus palabras lo que crea que esté sucediendo.

Aclaración: Para genera una muestra en estas condiciones basta generar una muestra como antes y luego sortear con probabilidad $p = 0.05$ multiplicar por 100 al primer elemento de la muestra.