

Instrucciones: Esta guía es un bonus track del TP1 a fin de afianzar conceptos relativos a la distribución de un estimador.

Ejercicio Error Cuadrático Medio

Sea X_1, \dots, X_n una muestra aleatoria (m.a.), es decir una colección de n variables independientes e idénticamente distribuidas, tales que $X_i \sim F_\theta$. Supongamos que θ es un parámetro desconocido y llamemos $\hat{\theta}_n = \hat{\theta}_n(X_1, \dots, X_n)$ a un estimador de θ basado en la muestra.

B1. Probar que $\text{MSE}(\hat{\theta}_n) = \mathbb{B}^2(\hat{\theta}_n) + \mathbb{V}(\hat{\theta}_n)$.

B2. Asumamos que nos interesa estimar $q(\theta)$, donde q es una función conocida. Por ejemplo, si $q(t) = t$ estaríamos diciendo que nos interesaría estimar el parámetro desconocido en sí mismo, mientras que si $q(t) = t^2$ estaríamos interesados en su cuadrado. Supongamos que estimamos a $q(\theta)$ por $\widehat{q(\theta)} = q(\hat{\theta}_n)$, en ese caso ¿cómo quedaría la fórmula del ítem anterior para el Error Cuadrático Medio de $\widehat{q(\theta)}$?

Ejercicio Simulación (Para hacer con R)

Dada una m.a. X_1, \dots, X_n de una distribución $\mathcal{U}(0, \theta)$, consideremos para θ los estimadores $\hat{\theta}_n = \max(X_1, \dots, X_n)$ y $\tilde{\theta}_n = 2\bar{X}_n$. De aquí en más evitaremos el uso del subíndice n para indicar el tamaño de la muestra.

B3. Para cada uno de los valores $n = 6, 10, 20, 40, 80, 200$ hacer lo siguiente:

- Generar n datos correspondientes a una muestra X_1, \dots, X_n i.i.d. de variables aleatorias distribuidas como X con una distribución $\mathcal{U}(0, \theta)$ con $\theta = 3$.
- Para cada conjunto de datos evaluar $\tilde{\theta}$ y $\hat{\theta}$.
- Repetir los dos pasos anteriores $m = 1000$ veces, obteniendo así, para cada uno de los dos estimadores, replicaciones $\theta_1^*, \dots, \theta_m^*$.
- Para cada estimador, computar el ECM mediante la expresión

$$\frac{1}{m} \sum_{i=1}^m (\theta_i^* - \theta)^2$$

B4. Graficar n vs. el ECM de cada uno de los dos métodos anteriores (todo en un mismo gráfico).

B5. Para cada valor de n hacer histogramas y boxplots con los valores estimados $\theta_1^*, \dots, \theta_m^*$.

B6. ¿Qué sugieren estos gráficos?