Estadística (M) - Segundo cuatrimestre 2018

Clase práctica del 03/09 - Sesgo y Error Cuadrático Medio

- 1. Dada una m.a. X_1, \ldots, X_n de una distribución $\mathcal{U}(0, \theta)$, sea $\widehat{\theta}_n$ el EMV de θ y $\widetilde{\theta}_n$ el estimador de θ basado en el primer momento.
 - (a) Probar que $\widetilde{\theta}_n$ es insesgado y que $\widehat{\theta}_n$ es asintóticamente insesgado.
 - (b) Calcular el ECM de ambos estimadores. ¿Qué estimador preferiría desde este punto de vista?
 - (c) ¿Cómo modificaría el estimador $\widehat{\theta}_n$ para que quede insesgado? Llamemos $\widehat{\theta}_n^{\text{mod}}$ a este estimador modificado.
 - (d) (Para hacer con R y para entregar) Para cada uno de los valores n=6,10,20,40,80,200 hacer lo siguiente:
 - i. Generar una muestra de tamaño n de una distribución $\mathcal{U}(0,\theta_0)$ con $\theta_0=3$.
 - ii. Evaluar los tres estimadores $\widetilde{\theta}_n$, $\widehat{\theta}_n$ y $\widehat{\theta}_n^{\mathrm{mod}}$ en dicha muestra.
 - iii. Repetir los dos pasos anteriores m=1000 veces, obteniendo así, para cada uno de los tres estimadores, replicaciones $\theta_1^*, \ldots, \theta_m^*$.
 - iv. Para cada estimador, computar el ECM

$$\frac{1}{m}\sum_{i=1}^{m}(\theta_i^*-\theta_0)^2$$

Finalmente, graficar n vs. el ECM de cada uno de los tres métodos anteriores (todo en un mismo gráfico). Hacer boxplots con los valores estimados $\theta_1^*, \ldots, \theta_m^*$. ¿Tienen sentido estos resultados respecto de los resultados teóricos de los primeros items?

- 2. Supongamos que queremos estimar la proporción p de gente que probó drogas duras. Si nadie tuviese pudor de contestar, lo más lógico sería tomar una muestra aleatoria de n personas de la población, hacerles la pregunta y estimar p con la proporción muestral de personas que respondieron que sí. El problema es que probablemente muchas personas que sí probaron drogas duras van a mentir y contestar que no.
 - (a) Pensar alguna forma alternativa de hacer una encuesta de modo que se pueda estimar p de manera tal que los encuestados no sientan que se viola su privacidad. **IMPORTANTE:** tratar de no pasar a la siguiente página sin haber pensado este item un poco.

1



- (b) Consideremos el siguiente procedimiento para la encuesta:
 - Se le pide a la persona que tire un dado y no le muestre el resultado al encuestador.
 - Si en el dado sale 1 o 2, el encuestado debe contestar la verdad respecto de si probó o no drogas duras.
 - Si en el dado sale 3, 4, 5 o 6, el encuestado debe contestar al azar de la siguiente forma: tira una moneda equilibrada, si sale cara responde SÍ y si sale ceca responde NO.
 - El encuestador no ve los resultados del procedimiento del encuestado. Sólamente sabe si el encuestado responde SÍ o NO (ni siquiera sabe si fue necesaria la tirada de la moneda)

Probar que si los encuestados siguen este procedimiento honestamente, se puede estimar p de forma insesgada.

- (c) Hallar el ECM del estimador obtenido en el item anterior. y compararlo con el ECM del estimador en una encuesta clásica (suponiendo que nadie tiene pudor de contestar).
- (d) (Opcional para entregar) Realizar una simulación con condiciones que considere razonables para validar los resultados del item anterior.
- (e) (Opcional para entregar) Hacer un análisis similar si se le ocurrió algún estimador en el item (a).