

PARCIAL 2 - ESTADÍSTICA II

ITAM, Primavera 2020

13/04/2020

Instrucciones

El examen es para resolver en casa. Se debe contestar individualmente y entregarse a más tardar a las 23:59 del lunes 13 de abril. La entrega será via email a la dirección: salvador.garcia.gonzalez@itam.mx (la misma que aparece en comunidad). En cuanto se envíe el examen, favor de escribir un mensaje para que les confirme de recibido. El examen cuenta con 9 preguntas a desarrollar y 1 opcional. Se debe cuidar la formalidad al escribir los resultados, ya que es parte de la calificación del problema. En caso de no tener el desarrollo de la pregunta, o bien se llegó a la respuesta sin una justificación se podrá anular la respuesta. Cualquier práctica fraudulenta será sancionada de acuerdo al reglamento del departamento.

Seccion A: Propiedades de estimadores (30 pts)

1. (10 pts) Suponga que el número de pacientes de una enfermedad en los países del G20 de la sigue la siguiente distribución. Donde x es el número de enfermos.

$$f(x; \theta) = \frac{3\theta^3}{x^4} \quad \text{para } x > \theta \quad (1)$$

Se está evaluando el ingreso de otros 5 países, pero para su ingreso se debe determinar el número máximo de enfermos por dicha enfermedad. Como estimador se propone a $\hat{\theta} = b\bar{X}$. Determine el valor de b , tal que el estimador $\hat{\theta}$ sea insesgado (Hint: No olvidar que, si $E(x)$ no es proporcionada, se puede calcular con $\int_a^b x f(x; \theta) dx$)

2. (20 pts) Sea X_1, X_2, \dots, X_n una muestra aleatoria de una población con media μ y varianza σ^2 . Considere los tres estimadores siguientes para μ :

$$\hat{\mu}_1 = \frac{1}{2}(X_1 + X_2) \quad \hat{\mu}_2 = \frac{1}{4}X_1 + \frac{X_2 + \dots + X_{n-1}}{2(n-2)} + \frac{1}{4}X_n \quad \hat{\mu}_3 = \bar{X} \quad (2)$$

1. Determine si son insesgados
2. Encuentre la varianza de cada estimador e identifique el más eficiente
3. Determine las eficiencias relativa: $\frac{ECM(\hat{\mu}_2)}{ECM(\hat{\mu}_3)}$ y $\frac{ECM(\hat{\mu}_1)}{ECM(\hat{\mu}_3)}$

Seccion B: Intervalos de confianza y tamaño de muestra (70 pts)

3. (10 pts) En una muestra aleatoria de 100 empleados de una empresa 50 están a favor del home office. Obtenga un intervalo de confianza (IC) del 95 % de probabilidad para la proporción poblacional.
4. (10 pts) El jefe del departamento de estadística va a estimar la calificación promedio de los estudiantes de estadística 2. Para ello, desea usar una muestra lo suficientemente grande para que la probabilidad que la media muestral no difiera de la media poblacional en más del 25 % de la desviación estándar poblacional. Tome $(1 - \alpha) = 95 \%$. ¿De qué tamaño debe elegir la muestra?
5. (10 pts) Se desea conocer la proporción de la población que está en contra de un cambio en política pública. ¿Cuántas personas se debe incluir en la muestra si se desea con una probabilidad de 96 % que la proporción muestral no diste de la verdadera proporción en más de 0.15. Se sabe de un estudio previo que σ^2 es de 1 (la varianza poblacional).
6. (10 pts) Se seleccionó una muestra aleatoria de 21 repartidores de pizza y se les preguntó el número de horas que trabajan a la semana. La desviación estándar muestral fue de 7 horas. Determine un IC del 90 % para la varianza de las horas de trabajo de todos los repartidores de pizza. (Asuma normalidad)
7. (10 pts) Se desea comparar la cantidad de venados con la cantidad de lobos en un parque estatal. Se supone que la cantidad de venados N_V depende de la cantidad de lobos N_L , por lo que se quiere estimar la diferencia en el número de lobos vs el número de venados $N_V - N_L$. Se tienen los siguientes datos obtenidos en distintos sectores del parque:

Sector	1	2	3	4	5	6	7	8	9	10	11	12
N_V	110	66	44	49	49	49	52	61	64	65	64	56
N_L	81	102	37	44	44	43	44	48	67	69	74	60

Encuentre un IC al 90 % para la diferencia de lobos y venados en el parque. Determine si podemos afirmar que hay una diferencia.

8. (10 pts) Los pangolines son unos mamíferos que pasan durmiendo la mayor parte del día. Pero se cree que los pangolines machos duermen más, pero con una varianza mayor que las pangolines hembras. Se tienen los siguientes datos para una muestra de 40 pangolines macho y 40 pangolines hembra:

Hembras	Machos
$\bar{x} = 5.35$	$\bar{y} = 4.31$
$s_x^2 = 2.31$	$s_y^2 = 1.21$

Construya un intervalo de confianza del 99 % para el cociente de varianzas $\frac{\sigma_x^2}{\sigma_y^2}$. Con base en este intervalo concluya si el supuesto de que las varianzas poblacionales son distintas.

9. (10 pts) Los empleados de una app de reparto de comida quieren conocer el tiempo promedio de entrega. Eligen aleatoriamente 16 pedidos y toman el tiempo promedio. Los resultados son: $\sum X_i = 336$ y $\sum X_i^2 = 7116$. Determine el intervalo de confianza del 99 % del tiempo promedio de entrega.

Pregunta de rescate - estimación puntual (5pts)

10. Sean X y Y dos poblaciones. Se toman 6 muestras aleatorias de los pesos de cada una, de manera que la muestra aleatoria para X es 5, 7, 9, 11, 13, 15 y la muestra aleatoria para Y es 10, 14, 18, 22, 26, 30:

1. Estime el promedio promedio de la población X (con $\bar{X} = \frac{1}{n} \sum_{i=1}^n$)
2. Estime la varianza de la población Y (con $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$)

3. Estime el coeficiente de correlación de la población X y Y

(con $r = \frac{S_{XY}}{S_X S_Y}$, usando $S_{XY} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}$, $S_X = \sqrt{S_X^2}$, $S_Y = \sqrt{S_Y^2}$)