

Tarea Examen - Probabilidades y Estadística - Otoño 2010

Profesor: Fernando Lema

Auxiliares: Abelino Jiménez - Juan Carlos Piña

Pregunta 1.

a) Sea $X \rightarrow e(\lambda)$, $\beta = \frac{1}{\lambda}$ y $X_1, X_2, \dots, X_n$ muestra aleatoria.

i.- ¿Cuál es el E.M.V. de β ? ¿Qué Distribución tiene para n grande?

ii.- Se desea estudiar

$$H_0 : \beta \leq \beta_0$$

$$H_1 : \beta > \beta_0$$

Usando i.- determine la condición para rechazar H_0 y evalúe si $n=100$; $\beta_0 = 5$.

$\bar{X} = 5.7$ y $\alpha = 0.05$ calcule p-value y la potencia si $\beta = 4.8$

b) Sea $X \rightarrow e(\lambda_X)$, $\beta_X = \frac{1}{\lambda_X}$ y $X_1, X_2, \dots, X_n$ muestra aleatoria, asimismo:

$Y \rightarrow e(\lambda_Y)$, $\beta_Y = \frac{1}{\lambda_Y}$ y $Y_1, Y_2, \dots, Y_m$ muestra aleatoria, siendo X e Y independientes

i.- Se desea estudiar el test

$$H_0 : \beta_X \geq \beta_Y$$

$$H_1 : \beta_X < \beta_Y$$

Indique el procedimiento para resolver el test

ii.- ¿Cómo cambia su respuesta si ahora

$$H_0 : \beta_X \geq k\beta_Y$$

$$H_1 : \beta_X < k\beta_Y$$

con k constante?.

Pregunta 2.

a) Un laboratorio requiere determinar la concentración de ADN de una bacteria en un extracto dado. Considerando que se cometen errores de medición, se replica 20 veces la medición. En la tabla adjunta se presenta los datos obtenidos de manera aleatoria e independiente. Se sabe que las mediciones se ajustan a una distribución $N(\mu, \sigma^2)$, donde μ y σ son desconocidos y μ es el valor real de la concentración de ADN.

Réplica	Concentración	Réplica	Concentración	Réplica	Concentración	Réplica	Concentración
1	0.503	2	0.249	3	1.047	4	0.670
5	0.882	6	0.858	7	0.805	8	0.549
9	0.563	10	0.706	11	0.503	12	0.536
13	1.168	14	0.813	15	0.496	16	0.790
17	0.518	18	0.463	19	0.735	20	1.038

i.- Encuentre el intervalo de confianza al 95% para μ a partir de los datos de la tabla.

ii.- Para el laboratorio es más arriesgado sobreestimar el valor real de μ de la concentración. Encuentre un intervalo $[a,b]$ de modo que la concentración de ADN pertenezca a dicho intervalo con un nivel de confianza de 95% y tal que $P(\mu < a) = \frac{1}{4}P(\mu > b)$.

iii.- Para demostrar al laboratorio la baja calidad de las mediciones, se solicita a usted que encuentre un intervalo al cual pertenezca σ^2 , el valor real de la varianza de la medición, con un nivel de confianza de 90%

b) En el año 2000, en las salas cunas de la Región Metropolitana se obtuvieron datos sobre 622 niñas y 694 niños recién nacidos. Supongamos que X e Y son las variables aleatorias que representan los perímetros del cráneo de las niñas y de los niños que se distribuyen según las distribuciones normales $N(\mu_1, \sigma_1^2)$ y $N(\mu_2, \sigma_2^2)$ respectivamente. Se observó en las muestras que las medias de cada grupo son $\bar{x} = 34,643 \text{ cm}$, $\bar{y} = 35,062 \text{ cm}$ y las varianzas de $S_1^2 = 2,18 \text{ cm}^2$, $S_2^2 = 2,16 \text{ cm}^2$.

i.- Se sabe que en la totalidad del país, los datos indican que el perímetro del cráneo de los recién nacidos de sexo masculino es en promedio 35cm con una desviación estándar de 1,35cm. ¿Los recién nacidos de sexo masculino de la Región Metropolitana poseen un perímetro de cráneo significativamente diferentes a la totalidad del país?

ii.- Queremos ahora comparar los perímetros del cráneo de los recién nacidos niños y niñas de la Región Metropolitana. ¿Podemos admitir que $\sigma_1^2 = \sigma_2^2$? Compare los perímetros del cráneo de los recién nacidos niños y niñas de la Región Metropolitana.

Pregunta 3.

a) Un supermercado quiere ver si la espera de sus clientes se distribuye como una Uniforme(0,10), para así saber cómo plantear una nueva forma de publicidad. Para realizar el estudio, se consideró una muestra de clientes durante 3 horas, llegándose a la siguiente tabla:

5.4	5.1	1.5	3.6	4.2	1.4	1	7.8	4.9	3.4
7.4	5.7	2	8	5.9	9.4	7.4	2.9	0.8	4.8
4	2.4	6.5	8.4	3.4	2.6	4.3	7	2.4	1.4
4.6	4.8	6.5	3.2	4.6	3.7	4.8	8.6	3.2	1

Tabla de tiempo de espera para 40 clientes

i.- Realice un test de ajuste considerando un nivel de significación de $\alpha = 0,05$.

ii.- ¿Se puede decir que menos del 40% de los clientes espera más de 5 minutos? (Realice un Test de Proporciones).

b) A una empresa de estudios de mercado se le encargó el estudio de la elección del nombre de una nueva marca de cigarrillos. Con este fin se realiza una encuesta cuyo resultado se presenta en la tabla de contingencia.

	Nombre de la Marca			
Percepción	Alazán	Corsario	Zodiaco	TOTAL
Viril	24	46	10	80
Ridículo	6	30	36	72
Sobrio	46	2	4	52
Vulgar	8	64	14	86
TOTAL	84	142	64	290

i.- Analice si existe una relación entre la percepción del consumidor y el nombre de la marca. Use su resultado para elegir el nombre de la marca.

ii.- Un estudiante dice que la proporción de personas de la población que encuentra vulgar los nombres propuestos se encuentra entre 0.28 y 0.35, ¿Con que confianza lo dice? (Note que el intervalo no es simétrico).

Pregunta 4.

Se quiere estudiar la relación entre la Viscosidad del aceite (X), y el Volumen de desgaste del acero dulce (Y), para un determinado tipo de motor.

Por medio de una muestra, se obtuvieron los siguientes datos:

Y	240	181	193	155	172	110	113	75	94
X	1,6	9,4	15,5	20	22	35,5	43	40,5	33

i.- Haga un gráfico a partir de los datos anteriores

ii.- Realice la estimación de mínimos cuadrados a partir de la tabla.

iii.- Estudie las Hipótesis

$$\begin{array}{ll} H_0 : \alpha = 0 & H_0 : \beta = 0 \\ H_1 : \alpha \neq 0 & H_1 : \beta \neq 0 \end{array}$$

y concluya.

iv.-Construya un intervalo de confianza al 95% para α y β .

v.- Construya un intervalo de confianza al 90% para el volumen de desgaste promedio si se tiene una viscosidad de 20.

vi.- ¿Qué viscosidad tendría que existir para que el volumen de desgaste sea cero? ¿Parece razonable?

Pregunta 5.

a) Considere el modelo de regresión $E(Y | X = x) = \beta x$, donde β es desconocido. Suponga que obtenemos los valores independientes $(x_1, y_1), \dots, (x_n, y_n)$.

i.- Determine el estimador de mínimos cuadrado de β .

ii.- Demuestre que b , el estimador por mínimos cuadrado de β , es insesgado.

iii.- Suponiendo que $\text{var}(Y | X = x) = \sigma^2$, demuestre que

$$\text{Var}(b | X_1 = x_1, \dots, X_n = x_n) = \frac{\sigma^2}{\sum_{i=1}^n x_i^2}$$

iv.- Con las condiciones anteriores, demuestre que

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - bx_i)^2$$

es un estimador insesgado de σ^2 .

b) Sea $y = \alpha + \beta x + u$ un modelo lineal y $\hat{\alpha}, \hat{\beta}$ los estimadores de mínimos cuadrados basados en una muestra $\{(x_i, y_i)\}_{i=1}^n$. Se consideran las transformaciones

$$y'_i = \frac{y_i - \bar{y}}{S_y} \quad x'_i = \frac{x_i - \bar{x}}{S_x}$$

variables “estandarizadas” y el modelo $y' = a + bx' + u'$

Dé una expresión para los estimadores $\hat{a}, \hat{b}$ como función de $\{(x_i, y_i)\}_{i=1}^n$ e indique su relación con $\hat{\alpha}, \hat{\beta}$.

NOTA: Esta tarea es de carácter obligatorio para los alumnos eximidos.

Fecha de Entrega: día del Examen.