Conceptos básicos de inferencia estadística | Semestre 2018-1

Tarea 9

Fecha de entrega: 24 de noviembre

1. Sea X una variable aleatoria continua con densidad

$$f(x\,|\,\theta) = \begin{cases} \mathbb{1}_{(0,1)}(x), & \text{si } \theta = 0. \\ \frac{1}{2\sqrt{x}}\mathbb{1}_{(0,1)}(x), & \text{si } \theta = 1. \end{cases}$$

Se tiene interés en contrastar la hipótesis

$$H_0 : \theta = 0 \qquad \text{vs.} \qquad H_1 : \theta = 1.$$

A partir de una sola observación X, se propone rechazar H_0 si $X < 0.25$.

   a) Calcular el tamaño de la prueba.

   b) Calcular la potencia de la prueba.

2. En el ejercicio anterior. Encontrar la región de rechazo de la prueba basada en el cociente de verosimilitudes con tamaño α, $\alpha \in (0,1)$.

   a) Calcular y simplificar $\frac{L(0\,|\,\mathbf{x})}{L(1\,|\,\mathbf{x})}$.

   b) Mostrar que

$$\frac{L(0\,|\,\mathbf{x})}{L(1\,|\,\mathbf{x})} < k \qquad \Leftrightarrow \qquad x < c,$$

   donde k y c son constantes a determinar según el tamaño de la prueba.

   c) ¿Cuál debe ser el valor de c para que el tamaño de la prueba sea α?
   Hint: Para que el tamaño de la prueba sea α se debe cumplir que $P(X < c\,|\,H_0) = \alpha$.

   d) ¿Cuál es la potencia de esta prueba?

3. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una población $Exp(\lambda)$, $\lambda > 0$, con la parametrización tal que $E(X_i) = \lambda$. Se tiene interés en contrastar las hipótesis

$$H_0 : \lambda = \lambda_0 \qquad \text{vs.} \qquad H_1 : \lambda = \lambda_1,$$

con $\lambda_0 < \lambda_1$. Encontrar la prueba basada en el cociente de verosimilitudes con tamaño α, $\alpha \in (0,1)$.

   a) Indicar cuál es la verosimilitud bajo H_0, $L(\lambda_0\,|\,\mathbf{x})$.

   b) Indicar cuál es la verosimilitud bajo H_1, $L(\lambda_1\,|\,\mathbf{x})$.

   c) Calcular y simplificar $\frac{L(\lambda_0\,|\,\mathbf{x})}{L(\lambda_1\,|\,\mathbf{x})}$.

   d) Mostrar que

$$\frac{L(\lambda_0\,|\,\mathbf{x})}{L(\lambda_1\,|\,\mathbf{x})} < k \qquad \Leftrightarrow \qquad \sum_{i=1}^{n} x_i > c.$$

   donde k y c son constantes a determinar según el tamaño de la prueba.

   e) Indicar cuál debe ser el valor de c para que el tamaño de la prueba sea α.
   Hint: ¿Cuál es la distribución de $\sum_{i=1}^{n} X_i$ bajo H_0?

   f) ¿Cuál es la potencia de esta prueba?

4. En el ejercicio anterior, suponer que $H_0 : \lambda = 3$ vs. $H_1 : \lambda = 5$ y se tiene un tamaño de muestra $n = 10$.

   a) Determinar la región de rechazo de la prueba basada en el cociente de verosimilitudes con tamaño $\alpha = 0.01$.

   b) Calcular la potencia de esta prueba.

   c) Determine si se rechaza H_0 si los datos observados son

   $$\{0.31, 5.75, 5.59, 1.07, 1.38, 0.80, 1.98, 0.02, 2.39, 0.12\}$$

   d) Calcular el *p-value* de esta prueba.

5. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una población $Poi(\lambda)$, $\lambda > 0$. Se tiene interés en contrastar las hipótesis

   $$H_0 : \lambda = \lambda_0 \qquad \text{vs.} \qquad H_1 : \lambda = \lambda_1,$$

   con $\lambda_0 < \lambda_1$. Encontrar la prueba basada en el cociente de verosimilitudes con tamaño α, $\alpha \in (0, 1)$.

   a) Indicar cuál es la verosimilitud bajo H_0, $L(\lambda_0 \,|\, \mathbf{x})$.

   b) Indicar cuál es la verosimilitud bajo H_1, $L(\lambda_1 \,|\, \mathbf{x})$.

   c) Calcular y simplificar $\frac{L(\lambda_0 \,|\, \mathbf{x})}{L(\lambda_1 \,|\, \mathbf{x})}$.

   d) Mostrar que

   $$\frac{L(\lambda_0 \,|\, \mathbf{x})}{L(\lambda_1 \,|\, \mathbf{x})} < k \qquad \Leftrightarrow \qquad \sum_{i=1}^{n} x_i > c.$$

   donde k y c son constantes a determinar según el tamaño de la prueba.

   e) Indicar cómo se debe calcular el valor de c para que el tamaño de la prueba sea α. ¿Es posible determinar c para cualquier valor de α?
   Hint: ¿Cuál es la distribución de $\sum_{i=1}^{n} X_i$ bajo H_0?

   f) ¿Cuál es la potencia de esta prueba?

6. En el ejercicio anterior, suponer que las hipótesis son $H_0 : \lambda = 1$ vs. $H_1 : \lambda = 4$ y se tiene un tamaño de muestra $n = 12$.

   a) Encontrar la región de rechazo de la prueba basada en el cociente de verosimilitudes para un tamaño $\alpha = 0.063$.

   b) Calcular la potencia de la prueba anterior.

   c) ¿Cuál es el tamaño de una prueba con región de rechazo $C = \left\{\mathbf{x} \in \mathbf{X}(\Omega) : \sum_{i=1}^{12} x_i > 15\right\}$?

   d) ¿Cuál es la potencia de la prueba anterior? 79

   e) Usando la prueba del inciso *a*), determine si se rechaza H_0 si los datos observados son

   $$\{4, 1, 0, 0, 1, 3, 1, 0, 1, 2, 0, 0\}$$

   f) Calcular el *p-value* de esta prueba.

7. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una población $Exp(\lambda)$, con $\lambda > 0$

   $$f(x \,|\, \lambda) = \lambda e^{-\lambda x} \mathbb{1}_{(0,\infty)}(x).$$

a) Defina la variable T como $T = \frac{1}{\lambda}\sum_{i=1}^{n} X_i$. Justificar que T es una cantidad pivotal para λ.

b) Suponer que se observó una muestra de tamaño 8 de la población anterior. Calcular un intervalo de confianza 90 % para λ.

$$\{0.69, 0.32, 1.01, 0.36, 1.14, 0.46, 0.35, 0.01\}$$

c) Utilizar el intervalo del inciso anterior para contrastar la hipótesis $H_0 : \lambda = 2$, con un tamaño de prueba $\alpha = 0.1$.

8. Sea $X_1, \ldots, X_{10}$ una muestra aleatoria de una distribución $B(\theta)$, $\theta \in (0, 1)$. Se tiene interés en contrastar las hipótesis.

$$H_0 : \theta = \frac{1}{2} \qquad \text{vs.} \qquad H_1 : \theta = \frac{1}{4}.$$

a) Encontrar la prueba más potente de tamaño $\alpha = 0.0547$.

b) Encontrar la potencia de la prueba anterior.

9. Sea X una observación de la densidad

$$f(x \,|\, \theta) = \theta x^{\theta-1} \mathbb{1}_{(0,1)}(x), \qquad \theta > 0.$$

Se desea contrastar las siguientes hipótesis

$$H_0 : \theta = 1 \qquad \text{vs.} \qquad H_1 : \theta = 2.$$

a) Encontrar la prueba más potente de tamaño $\alpha = 0.05$.

b) Encontrar la potencia de la prueba anterior.

10. Sea $X_1, \ldots, X_{10}$ una muestra aleatoria de una distribución $B(\theta)$, $\theta \in (0, 1)$. Se tiene interés en contrastar las hipótesis.

$$H_0 : \theta \leq \frac{1}{2} \qquad \text{vs.} \qquad H_1 : \theta > \frac{1}{2}.$$

Se usa una prueba con región de rechazo $C = \left\{ \mathbf{x} \in \mathbf{X}(\Omega) : \sum_{i=1}^{10} x_i > 6 \right\}$

a) ¿Cuál es el tamaño de la prueba?

b) Encontrar la función potencia de la prueba y graficarla.

11. Sea X una observación de la densidad

$$f(x \,|\, \theta) = \theta x^{\theta-1} \mathbb{1}_{(0,1)}(x), \qquad \theta > 0.$$

Se desea contrastar las siguientes hipótesis

$$H_0 : \theta \leq 1 \qquad \text{vs.} \qquad H_1 : \theta > 1.$$

Se usa una prueba con regla de decisión: rechazar H_0 si $X > 0.5$.

a) ¿Cuál es el tamaño de la prueba?

b) Encontrar la función potencia de la prueba y graficarla.

12. La SEDESOL está interesada en estudiar la distribución del ingreso en los hogares de cierta zona marginada al sur del Estado de México. Para tal efecto se seleccionó una muestra aleatoria de $n = 2500$ hogares y se registró su ingreso mensual (en miles de pesos). Se reportaron los siguientes resultados

$$\sum_{i=1}^{2500} X_i = 33,916.69 \qquad \text{y} \qquad \sum_{i=1}^{2500} X_i^2 = 490,676.1$$

Suponer que la distribución del ingreso se puede modelar como $N(\mu, \sigma^2)$. Responder lo siguiente.

a) Reportar estimaciones de la media y la varianza del ingreso.

b) Construir un intervalo de confianza 99 % para la media del ingreso. Interpretar los resultados en el contexto del problema.

c) Construir un intervalo de confianza 90 % para la varianza del ingreso.

d) Utilizar el intervalo de confianza del inciso *b)* para contrastar si la media del ingreso en los hogares marginados al sur del Estado de México es igual a la media nacional reportada por INEGI de \$13,850 MXN. ¿Cuál es la significancia de esta prueba?

13. Un investigador está interesado en comparar la efectividad de un nuevo fármaco con la de uno ya existente, en el tratamiento de cierta enfermedad. Para medir la efectividad se considera la probabilidad (θ) de que un paciente con la enfermedad se cure, la cual se asume constante para todos los pacientes. El investigador corrió un estudio con $n = 149$ pacientes enfermos, a los que se aplicó un tratamiento con el nuevo fármaco y se registró como resultado si los pacientes se curaron o no. Se asume que los resultados de los tratamientos son independientes.

a) Si la efectividad del fármaco anterior es $\theta = 0.75$, plantear el juego de hipótesis estadísticas para contrastar si el nuevo fármaco es **más efectivo** que el anterior.

b) Suponer que se utiliza como estadístico de prueba el número de pacientes curados T. Justificar que la distribución aproximada de T bajo H_0 es aproximadamente normal e indicar con qué parámetros.

c) Definir una regla de decisión que permita rechazar la hipótesis nula con una significancia aproximada $\alpha = 0.01$.

d) Si el estudio arrojó que 140 pacientes se curaron, ¿qué se puede concluir acerca de la efectividad del nuevo fármaco?

14. En un experimento para determinar la toxicidad de una sustancia se administra una dosis de ésta a cada uno de 300 conejos, y se registra el numero de muertos, que resulta ser de 192. Sea θ la probabilidad de que un conejo elegido al azar muera a causa de una dosis de la sustancia.

a) Calcular un intervalo de confianza 90 % para θ.

b) Utilizar el intervalo anterior para contrastar la hipótesis $H_0 : \theta \leq 0.5$.