

ESTADÍSTICA TEÓRICA II. ENERO 2009 1ª SEMANA

PRIMERA PARTE: CUESTIONES

1.- Teorema de factorización de Fisher-Neyman.

Respuesta.-

Si (X_1, X_2, \dots, X_n) es una muestra aleatoria simple, siendo $P(x_1, x_2, \dots, x_n, \theta)$ ó $f(x_1, x_2, \dots, x_n, \theta)$ las funciones de probabilidad o densidad, respectivamente, entonces el estadístico $T = T(X_1, X_2, \dots, X_n)$ es suficiente para el parámetro θ si y sólo si existen dos funciones $g(T(X_1, X_2, \dots, X_n), \theta)$ y $h(x_1, x_2, \dots, x_n)$, tales que:

$$P(x_1, x_2, \dots, x_n, \theta) = g(T(X_1, X_2, \dots, X_n), \theta) \cdot h(x_1, x_2, \dots, x_n)$$

ó

$$f(x_1, x_2, \dots, x_n, \theta) = g(T(X_1, X_2, \dots, X_n), \theta) \cdot h(x_1, x_2, \dots, x_n)$$

2.- Intervalo de confianza de una proporción para muestras pequeñas.

Respuesta.-

Consideramos el estimador de máxima verosimilitud que, para una muestra de tamaño n de una población binomial $B(1, p)$ es la proporción muestral $\hat{p} = \frac{\text{nº de éxitos en } n \text{ pruebas}}{\text{nº de pruebas}} = \frac{X}{n}$ cuya función de probabilidad

$$P\left(\frac{X}{n} = x\right) = P(X = nx) = \binom{n}{nx} p^{nx} (1-p)^{n-nx}, \quad x = 0, \frac{1}{n}, \frac{2}{n}, \dots, 1.$$

Los extremos p_i y p_s de un intervalo de nivel de confianza $100(1-\alpha)\%$ cumplirían las ecuaciones:

$$\left. \begin{aligned} P\left(\frac{X}{n} \leq p_i\right) &= \sum_{x=0}^{p_i} \binom{n}{nx} p^{nx} (1-p)^{n-nx} \leq \frac{\alpha}{2} \\ P\left(\frac{X}{n} \geq p_s\right) &= \sum_{x=p_s}^1 \binom{n}{nx} p^{nx} (1-p)^{n-nx} \leq \frac{\alpha}{2} \end{aligned} \right\}$$

dada la dificultad de resolución de estas ecuaciones, Clopper y Pearson elaboraron unos ábacos que proporcionan las soluciones para distintos valores de x , con n y α fijos.

3.- Contrastes uniformemente más potentes

Respuesta.-

Supongamos un contraste de una hipótesis simple H_0 frente a una hipótesis compuesta H_1 . Se dice que C es la *región crítica uniformemente más potente de tamaño α* si es la mejor región crítica de ese tamaño para contrastar H_0 , para cada una de las hipótesis simples de las que consta H_1 .

Si la región crítica de un contraste cumple esta propiedad, diremos que el contraste es el uniformemente más potente de tamaño α .

4.- Regla de decisión para el contraste de Pearson de bondad de ajuste cuando los parámetros poblacionales son conocidos.

Respuesta.-

Consideremos una partición en k clases del campo de variación de la variable aleatoria X y una muestra aleatoria simple de tamaño n . Sea n_i , $i = 1, \dots, k$ el número de observaciones muestrales contenidas en la clase i -ésima.

Supongamos que la hipótesis nula establece las probabilidades p_1, p_2, \dots, p_k de que una observación caiga en cada una de las clases. Entonces, rechazaremos H_0 , al nivel de significación α , si

$$\chi^2_{\text{exp}} = \sum_{i=1}^k \frac{(n_i - np_i)^2}{np_i} > \chi^2_{1-\alpha}$$

donde $\chi^2_{1-\alpha}$ es tal que $P[\chi^2_{k-1} > \chi^2_{1-\alpha} / H_0] = \alpha$ en donde χ^2_{k-1} es la variable χ^2 con $k-1$ grados de libertad.

PROBLEMA 1

1.- En la distribución de Poisson de parámetro λ y para muestras aleatorias simples de tamaño n , compruebe si el estadístico $T = \sum_{i=1}^n x_i$ es suficiente.

Solución.-

Un estimador es suficiente si la probabilidad de obtener una determinada muestra, condicionada por un valor del estimador, no depende del parámetro a estimar. En el caso del problema se trataría de hallar la probabilidad de una muestra (X_1, X_2, \dots, X_n) condicionada por que el estadístico $T = \sum_{i=1}^n x_i$ tome un valor t .

$$P(X_1=x_1, X_2=x_2, \dots, X_n=x_n / T=t)$$

El parámetro a estimar es λ y se cumple que $P(X=x) = \frac{\lambda^x}{x!} e^{-\lambda}$. Además, la suma de n variables aleatorias independientes de Poisson, de parámetro λ , es también de Poisson, de parámetro $n\lambda$. Así pues, tendremos:

$$\begin{aligned} & P(X_1=x_1, X_2=x_2, \dots, X_n=x_n / T=t) = \\ &= \frac{P(X_1=x_1, \dots, X_n=x_n) P(T=t / X_1=x_1, \dots, X_n=x_n)}{P(T=t)} = \\ &= \begin{cases} \frac{\lambda^{\sum_{i=1}^n x_i} e^{-n\lambda}}{x_1! \cdots x_n!}, & \text{si } t = \sum_{i=1}^n x_i \\ \frac{(n\lambda)^{\sum_{i=1}^n x_i} e^{-n\lambda}}{\sum_{i=1}^n x_i!}, & \\ 0, & \text{si } t \neq \sum_{i=1}^n x_i \end{cases} = \begin{cases} \frac{\sum_{i=1}^n x_i!}{x_1! \cdots x_n! n^{\sum_{i=1}^n x_i}}, & \text{si } t = \sum_{i=1}^n x_i \\ 0, & \text{si } t \neq \sum_{i=1}^n x_i \end{cases} \end{aligned}$$

Observamos que el resultado, en cualquier caso, no depende del parámetro a estimar λ . Por ello concluimos que T es suficiente.

PROBLEMA 2

2.- Una determinada empresa agrícola, con criterios muy rígidos en cuanto a la productividad y eficacia de sus empleados, contrata un nuevo operario cuya función es la comprobación del peso de las mercancías embaladas. La empresa prepara cajas de 150 grs y considera que el operario ha superado el período de prueba si el peso medio de un lote de 15 cajas, elegido aleatoriamente, es igual o inferior a los 150 grs. Los datos del

lote nos proporcionan la siguiente información: $\sum_{i=1}^n x_i = 2.268,51$; $\sum_{i=1}^n x_i^2 = 343.307,75$.

Si el peso de las mercancías embaladas se distribuye normalmente, con un nivel de significación del 10% , realice un contraste para comprobar si el nuevo empleado superará el período de prueba.

Solución.-

El contraste es:

$$H_0: \mu \leq 150$$

$$H_1: \mu > 150$$

Consideramos la variable $\frac{\bar{X} - \mu}{S} \sqrt{n}$ que se distribuye t-Student con n-1 grados de libertad. En el problema n = 14 y en las tablas de la t-Student con 14 grados de libertad encontramos que $0,1 = P[t > 1,345]$. Así pues, aceptaremos H_0 si $\frac{\bar{X} - \mu}{S} \sqrt{n} < 1,345$.

Tenemos:

$$\bar{X} = \frac{2268,51}{15} = 151,234$$

$$S = \sqrt{\frac{15}{14} \left(\frac{343307,75}{15} - 151,234^2 \right)} \cong 4,07$$

$$\text{Luego } \frac{\bar{X} - \mu}{S} \sqrt{n} = \frac{151,234 - 150}{4,07} \sqrt{15} \cong 1,174$$

Por tanto, aceptamos H_0 , es decir, el nuevo empleado superará el periodo de prueba.