

ESTADÍSTICA TEÓRICA II. FEBRERO 2010 2ª SEMANA

PRIMERA PARTE: CUESTIONES

1.- Distribución de la diferencia de medias muestrales cuando no se conoce la varianza poblacional.

Respuesta.-

Sean X e Y dos variables independientes, normales $N(\mu_x, \sigma_x)$ y $N(\mu_y, \sigma_y)$, pertenecientes a dos poblaciones distintas cuyas varianzas respectivas σ_x y σ_y sean desconocidas. Tomemos sendas muestras de tamaños n_x y n_y y medias y varianzas muestrales \bar{X} , \bar{Y} , S_x^2 , S_y^2 respectivamente.

Si $\sigma_x = \sigma_y = \sigma$, entonces
$$\frac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{\sqrt{\frac{1}{n_x} + \frac{1}{n_y}} \sqrt{(n_x - 1)S_x^2 + (n_y - 1)S_y^2}} \sqrt{n_x + n_y - 2}$$
 se distribuye

como una t-Student con $n_x + n_y - 2$ grados de libertad.

Si $\sigma_x \neq \sigma_y$ y los tamaños muestrales son grandes ($n_x \geq 30$, $n_y \geq 30$), entonces
$$\frac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{\sqrt{\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y}}}$$
 es aproximadamente normal $N(0, 1)$.

Si $\sigma_x \neq \sigma_y$ y los tamaños muestrales son pequeños, entonces
$$\frac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{\sqrt{\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y}}}$$
 es

aproximadamente t-Student con v grados de libertad, donde v es el entero más próximo a

$$\frac{\left(\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y}\right)^2}{\frac{\left(\frac{S_x^2}{n_x}\right)^2}{n_x - 1} + \frac{\left(\frac{S_y^2}{n_y}\right)^2}{n_y - 1}}$$

2.- Estimador eficiente. Expresión analítica.

Respuesta.-

Un estimador $\hat{\theta}$ del parámetro θ es eficiente si es insesgado y su varianza es la mínima, es decir, coincide con la cota de Frechet-Cramer-Rao:

$$\text{Var}(\hat{\theta}) = \frac{1}{nE\left[\left(\frac{\partial \ln f(x; \theta)}{\partial \theta}\right)^2\right]}$$

3.- Tamaño de la muestra para estimar la proporción de una población.

Respuesta.-

Se trata de hallar el tamaño muestral n para construir un intervalo de confianza de longitud L , al nivel de confianza $100(1-\alpha)\%$. Puesto que tal intervalo es:

$$\left[\hat{p} - z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}}, \hat{p} + z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}} \right]$$

cuya longitud es $L = 2 z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}}$, podemos despejar n, obteniéndose: $n = \frac{4z_{\frac{\alpha}{2}}^2 \hat{p}\hat{q}}{L^2}$.

En particular, como el error $e = |\hat{p} - p| = \frac{L}{2}$, se tiene que $n = \frac{4z_{\frac{\alpha}{2}}^2 \hat{p}\hat{q}}{4e^2} = \frac{z_{\frac{\alpha}{2}}^2 \hat{p}\hat{q}}{e^2}$

4.- Lema de Neyman-Pearson.

Respuesta.-

Proporciona un criterio para hallar la región crítica de tamaño α en un contraste de hipótesis, que haga mínimo el error de tipo II, $\beta = P[\text{Aceptar } H_0 / H_0 \text{ es falsa}]$. Diremos que es la mejor región crítica. Establece que, obtenida una muestra aleatoria simple (X_1, X_2, \dots, X_n) de una población con función de densidad $f(x; \theta)$, siendo $L(x_1, \dots, x_n; \theta) = \prod_{i=1}^n f(x_i; \theta)$ la función de verosimilitud de una muestra, si queremos efectuar el contraste:

$$H_0: \theta = \theta_0$$

$$H_1: \theta = \theta_1$$

siendo $k > 0$ y C un subconjunto del espacio muestral \mathbb{R}^n tal que:

$$\frac{L(x_1, \dots, x_n; \theta_0)}{L(x_1, \dots, x_n; \theta_1)} \leq k \text{ si } (x_1, \dots, x_n) \in C$$

$$\frac{L(x_1, \dots, x_n; \theta_0)}{L(x_1, \dots, x_n; \theta_1)} > k \text{ si } (x_1, \dots, x_n) \notin C$$

$$P((X_1, \dots, X_n) \in C / \theta = \theta_0) = \alpha$$

Entonces C es la mejor región crítica de tamaño α para efectuar el contraste.

PROBLEMAS

1.- El Director Comercial de una determinada empresa X asegura haber realizado un magnífico trabajo en el lanzamiento del nuevo producto “Neo”, competidor del “Clásico” de otra compañía Y. En una reunión del Comité de Dirección insiste hasta el extremo de renunciar a su “bonus” anual si no obtiene un 70% del mercado de la empresa Y. Se realiza un muestreo aleatorio simple con 800 compradores y resulta que 496 se manifiestan a favor de “Neo”. Con un nivel de significación del 10%, contraste la hipótesis de que el citado Director deba renunciar al “bonus”.

Solución.-

Sea p la proporción (poblacional) de mercado de la empresa Y obtenido por la empresa X después del lanzamiento del nuevo producto “Neo”. Plantearemos el contraste:

$$H_0: p \geq 0,7$$

$$H_1: p < 0,7$$

Bajo la hipótesis nula, la proporción muestral \hat{p} se distribuirá aproximadamente normal

$N\left(0,7; \sqrt{\frac{0,7 \cdot 0,3}{800}}\right) \cong N(0,7; 0,0162)$. Rechazaremos H_0 si la proporción muestral hallada

$\frac{496}{800} = 0,62 < x_{0,1}$, siendo $x_{0,1}$ tal que $0,1 = P[\hat{p} < x_{0,1}] = P\left[Z < \frac{x_{0,1} - 0,7}{0,0162}\right]$. De las tablas de la

normal $N(0, 1)$ obtenemos que $\frac{x_{0,1} - 0,7}{0,0162} = -1,28$, de donde $x_{0,1} \cong 0,679$. Como $0,62 < 0,679$, se debe rechazar la hipótesis nula, es decir, el director deberá renunciar al “bonus”.

2.- Dada una población definida por la función de cuantía:

$$P(\xi = -1) = \frac{1-\theta}{2}$$

$$P(\xi = 0) = \frac{\lambda+\theta}{2}$$

$$P(\xi = 1) = \frac{1-\lambda}{2}$$

Siendo $0 < \theta < 1$ y $0 < \lambda < 1$

Estime ambos parámetros por el método de los momentos, estudiando si son insesgados los estimadores obtenidos.

Solución.-

$$E(\xi) = \frac{\theta-1}{2} + \frac{1-\lambda}{2} = \frac{\theta-\lambda}{2}$$

$$E(\xi^2) = \frac{1-\theta}{2} + \frac{1-\lambda}{2} = \frac{2-\theta-\lambda}{2}$$

Igualando respectivamente a los momentos muestrales:

$$\left. \begin{aligned} \frac{\theta-\lambda}{2} &= \bar{X} \\ \frac{2-\theta-\lambda}{2} &= \overline{X^2} \end{aligned} \right\} \text{de donde se obtienen los estimadores: } \left. \begin{aligned} \hat{\theta} &= \bar{X} - \overline{X^2} + 1 \\ \hat{\lambda} &= -\bar{X} - \overline{X^2} + 1 \end{aligned} \right\}$$

Comprobemos si son insesgados:

$$\left. \begin{aligned} E(\hat{\theta}) &= E(\bar{X}) - E(\overline{X^2}) + 1 = \frac{\theta-\lambda}{2} - \frac{2-\theta-\lambda}{2} + 1 = \theta \\ E(\hat{\lambda}) &= -E(\bar{X}) - E(\overline{X^2}) + 1 = -\frac{\theta-\lambda}{2} - \frac{2-\theta-\lambda}{2} + 1 = \lambda \end{aligned} \right\}$$

Luego ambos son insesgados.