Cálculo de probabilidades y estadística **Examen Segundo Parcial**

EJERCICIO 2 (SOLUCIONES) **20 de mayo de 2024**

1.– Una empresa de fabricación de cables decide realizar un control de calidad de las bobinas realizando $n = 4$ ensayos destructivos. De esta manera se obtienen 4 medidas experimentales del límite elástico del cable. Se supone que las medidas son independientes y responden a una distribución $N(m, \sigma^2)$, donde σ^2 es desconocida. El control se efectúa mediante un contraste de hipótesis con significación $\alpha = 0.1$ sobre

$$H_o: \; m \geq \; 400, \quad H_1: < \; 400$$

En una prueba se obtuvo una media muestral de 375 y una desviación típica $S^* = 25$. Se pide:

a) Determinar el estadístico adecuado para el contraste y la correspondiente región de aceptación.

b) Determinar la probabilidad del error Tipo II si el verdadero límite elástico medio fuera de 380.

c) Determinar la probabilidad de rechazar la hipótesis primaria si el verdadero límite elástico medio fuera 405.

d) Supóngase que con la misma significación y el mismo número de datos, pero con otros valores muestrales, se ha obtenido que el error Tipo II para $m = 380$ es 0.08484 y la probabilidad de rechazar la hipótesis primaria para $m = 405$ es 0.29530. Determínese cual fue en este caso la región de aceptación y cuál es el valor de la desviación típica muestral obtenida

Nota: todos los valores de los límites elásticos se indican en MPa

——————SOLUCIÓN——————

a) El contraste de hipótesis que debemos realizar es

$H_o: m \geq 400$

$H_1: m < 400$

que como sabemos es equivalente, desde un punto de vista operativo a

$H_o: m = 400$

$H_1: m < 400$

Como la varianza es desconocida utilizaremos el estadístico $t = \frac{\bar{x}-m}{S*/\sqrt{n}} \equiv t_{n-1}$. En nuestro caso $n = 4$. Dado el tipo de contraste, la región crítica estará en la cola inferior de la distribución del estadístico t. Si suponemos que la desviación típica muestral que hemos obtenido es la insesgada, es decir $S^* = 25$, en el supuesto de la hipótesis primaria $t = \frac{\bar{x}-400}{25/\sqrt{4}} \equiv t_3$. Por tanto nuestra regla de decisión, para $\alpha = 0.10$, será rechazar la hipótesis primaria si $t < c$, donde $F_t(c) = 0.10$. De las tablas $c = -1.6378$. De nuestra muestra el estadístico t vale

$t = \frac{375-400}{25/\sqrt{4}} = -2$ y por lo tanto debemos rechazar la hipótesis primaria a este nivel de significación.

b) La probabilidad de error Tipo II es

$$\beta = P[t \geq c|H_1] = P\Big[\frac{\bar{x}-400}{25/\sqrt{4}} \geq c|H_1\Big] = P\Big[\bar{x} \geq 400 + c\frac{25}{\sqrt{4}}|H_1\Big] =$$

$$= P\Big[\bar{x} - m_1 \geq 400 - m_1 + c\frac{25}{\sqrt{4}}|H_1\Big] = P\Big[\frac{\bar{x}-m_1}{25/\sqrt{4}} \geq c + \frac{400-m_1}{25/\sqrt{4}}|H_1\Big] = 1 - F_t\Big(c + \frac{400-m_1}{25/\sqrt{4}}\Big)$$

Para $m_1 = 380$,

$$\beta = 1 - F_t\Big(c + \frac{400-380}{25/\sqrt{4}}\Big) = 1 - F_t\Big(-1.6378 + \frac{400-380}{25/\sqrt{4}}\Big) = 1 - F_t(-0.0378) = F_t(0.0378) = 0.51388$$

c) La probabilidad q de rechazar la hipótesis primaria si el verdadero límite elástico medio fuera 405 es

$$q = P[t < c|m_1 = 405] = P\Big[\frac{\bar{x}-400}{25/\sqrt{4}} < c|m_1 = 405\Big] = P\Big[\frac{\bar{x}-m_1}{25/\sqrt{4}} < c + \frac{400-m_1}{25/\sqrt{4}}|m_1 = 405\Big] =$$

$$= F_t\Big(-1.6378 + \frac{400-405}{25/\sqrt{4}}\Big) = F_t(-2.0378) = 0.06716$$

d) Nos dicen por una parte que

$$\beta = 0.08484 = P[t \geq c|H_1] = 1 - F_t\Big(c + \frac{400-380}{S^*/\sqrt{4}}\Big)$$

y por otra

$$q = 0.29530 = F_t\Big(c + \frac{400-405}{S^*/\sqrt{4}}\Big)$$

Consecuentemente tenemos estas dos ecuaciones con dos incógnitas (c y S^*):

$$0.91516 = F_t\Big(c + \frac{400-380}{S^*/\sqrt{4}}\Big) = F_t(c + \frac{40}{S^*})$$

$$0.29530 = F_t\Big(c + \frac{400-405}{S^*/\sqrt{4}}\Big) = F_t(c - \frac{10}{S^*})$$

Y de las tablas obtenemos

$$c + \frac{40}{S^*} = 1.8$$

$$c - \frac{10}{S^*} = -0.6$$

de donde

$$S^* = 20.833, \qquad c = -0.12$$

por lo que nuestra región de aceptación es, usando el estadístico $t = \frac{\bar{x}-400}{18.048/\sqrt{4}} = \frac{\bar{x}-400}{9.024} \equiv t_3$,

$$(-0.12, \infty)$$

2.– Determinar los estimadores de los parámetros de una distribución logarítmico normal por el método de máxima verosimilitud.

————————SOLUCIÓN————————

La función de densidad de una distribución lognormal puede escribirse como

$$f_X(x) = \frac{1}{x\sigma\sqrt{2\pi}} \quad e^{-\frac{1}{2\sigma^2}\left(\ln(\frac{x}{m})\right)^2}$$

donde por σ representamos la desviación logarítmica y por m la mediana de la distribución. Con esta notación, la función de verosimilitud es

$$L = \prod_{i=1}^{n} \frac{1}{x_i\sigma\sqrt{2\pi}} \quad e^{-\frac{1}{2\sigma^2}\left(\ln(\frac{x_i}{m})\right)^2} = \frac{1}{\sigma^n(\sqrt{2\pi})^n}\prod_{i=1}^{n} \frac{1}{x_i} \quad e^{-\frac{1}{2\sigma^2}\left(\ln(\frac{x_i}{m})\right)^2}.$$

Tomando logaritmos

$$\ln L = -n\ln\sigma - n\ln\sqrt{2\pi} + \sum_{i=1}^{n} \ln\left(\frac{1}{x_i}\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(\ln(\frac{x_i}{m})\right)^2 =$$

$$= -n\ln\sigma - n\ln\sqrt{2\pi} - \sum_{i=1}^{n} \ln x_i - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(\ln(\frac{x_i}{m})\right)^2$$

Derivando con respecto a σ y a m para hallar los valores que hacen máxima la función de verosimilitud obtenemos

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}\left(\ln(\frac{x_i}{m})\right)^2$$

$$\frac{\partial \ln L}{\partial m} = -\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left[2\ln(\frac{x_i}{m})(-\frac{1}{m})\right] = \frac{1}{m\sigma^2}\sum_{i=1}^{n}\left[\ln x_i - \ln m\right]$$

Si igualamos a cero esta última ecuación resulta

$$\sum_{i=1}^{n}\left[\ln x_i - \ln \hat{m}\right] = 0, \qquad \Rightarrow \qquad \hat{m} = e^{\frac{\sum \ln x_i}{n}}$$

De la primera ecuación, igualando igualmente a cero obtenemos

$$n\hat{\sigma}^2 = \sum_{i=1}^{n}\left(\ln(\frac{x_i}{\hat{m}})\right)^2, \qquad \Rightarrow \qquad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\ln x_i - \ln \hat{m}\right)^2$$