# CPE (SEGUNDO CURSO)

**PRÁCTICA 14** **SOLUCIONES** (**Curso 2024–2025**)

**1.–** Se sabe que un distanciómetro de alta precisión proporciona lecturas de una cierta distancia (aproximadamente entre 3 y 4 Km) distribuidas normalmente y de media igual a la distancia que se pretende medir. Aunque no se conoce la varianza, se sabe que ésta no excede a 4 $\text{cm}^2$.

a) ¿Cuál es el número mínimo de lecturas que habrá que tomar para asegurarse que el intervalo de confianza del 90 % sobre la media tenga longitud inferior a 1 cm?

b) Supongamos que un topógrafo realizó el número de lecturas adecuado según el apartado anterior y obtuvo $\bar{x} = 350089.78$ cm y $S^2 = 3.0\ \text{cm}^2$. ¿Cuál sería su estima por intervalo de la longitud deseada?

——————————SOLUCIÓN——————————

Sea $X$ la lectura del distanciómetro. Sabemos que $X \equiv N(m, \sigma^2),\ \sigma^2 \leq 4\ \text{cm}^2$

**a)** El intervalo de confianza del 90 % sobre la media es

$$\left(\overline{x} - 1.645\frac{\sigma}{\sqrt{n}},\ \overline{x} + 1.645\frac{\sigma}{\sqrt{n}}\right)$$

Para que este intervalo tenga una longitud inferior a 1 cm se debe verificar

$$2 \times 1.645\frac{\sigma}{\sqrt{n}} < 1 \ \Rightarrow\ \sqrt{n} > 3.29\sigma$$

Pero $\sigma \leq 2 \ \Rightarrow\ \sigma_{max} = 2$. Luego $\sqrt{n} > 6.58$. Es decir $n \geq 44$

**b)** Si utilizamos la varianza muestral el intervalo de confianza es ahora

$$IC_{90\,\%} \equiv \left(\overline{x} - a\frac{S^*}{\sqrt{n}},\ \overline{x} + a\frac{S^*}{\sqrt{n}}\right)$$

con $a = 1.684$ de las tablas de la $t$ de Student con 43 grados de libertad.

Como $S^2 = 3 \ \Rightarrow\ {S^*}^2 = \frac{n}{n-1}S^2 = 3.07 \ \Rightarrow\ S^* = 1.752$ y el intervalo pedido es

$$IC_{90\,\%} \equiv \left(350089.33,\ 350090.23\right)$$

Obsérvese que si hubiéramos supuesto $n = 44$ suficientemente grande y por lo tanto $\sigma \approx S^*$, el resultado hubiera sido prácticamente el mismo, ya que al aproximar por una normal en vez de una $t$ de Student el parámetro $a$ del intervalo hubiera sido 1.645 en lugar de 1.684.

**2.**– El tiempo de fallo de un componente sigue una distribución exponencial de parámetro desconocido. Estos componentes se utilizan en el ensamblaje de dos sistemas distintos; en el primero se ensamblan en serie $n_1$ componentes, y en el segundo, también en serie, $n_2$ componentes. Se ensayan ambos sistemas y se obtienen los tiempos hasta fallo de cada sistema, $t_1$ y $t_2$, respectivamente.

Se pide:

**a)** Con la información disponible determinar el estimador máximo verosímil de $\lambda$.

**b)** Calcular el intervalo de confianza por ambos lados del 90 % sobre $\lambda$, para el caso en que $n_1 = 10 = n_2$, y $t_1 = 400$ horas y $t_2 = 350$ horas.

*Nota: el montaje en serie supone que el sistema funciona únicamente cuando todos los componentes del sistema están funcionando correctamente.*

——————SOLUCIÓN——————

**a)** Si los sistemas están ensamblados en serie, su tiempo de vida será el del mínimo de sus componentes. Sean, por tanto, $T_1$ y $T_2$ las duraciones de ambos sistemas. Si $X_i$ es la vida de un componente cualquiera

$$T_1 = \min_{i=1,n_1} \{X_i\}, \qquad T_2 = \min_{j=1,n_2} \{X_j\}$$

Dado que todos los componentes se comportan de la misma forma y si suponemos que la duración de cada uno es independiente de la de los demás, podemos escribir, por ejemplo para $T_1$,

$$P[T_1 > t] = 1 - F_{T_1}(t) =$$
$$= P[\{X_1 > t\} \cap \{X_2 > t\} \cap \{X_3 > t\} \cap ... \cap \{X_{n_1} > t\}] = (1 - F_X(t))^{n_1}$$

Como $F_X(x) = 1 - e^{-\lambda x}, \;\; x \geq 0$, obtenemos que

$$F_{T_1}(t) = 1 - e^{-n_1 \lambda t}, \;\; t \geq 0$$

e igualmente

$$F_{T_2}(t) = 1 - e^{-n_2 \lambda t}, \;\; t \geq 0$$

Es decir

$$T_1 \equiv EX[n_1\lambda] \quad \text{y} \quad T_2 \equiv EX[n_2\lambda]$$

Hemos obtenido dos valores muestrales, $t_1$ y $t_2$. La función de verosimilitud de esta muestra es, por tanto,

$$L = f_{T_1}(t_1) f_{T_2}(t_2) = n_1 \lambda e^{-n_1 \lambda t_1} n_2 \lambda e^{-n_2 \lambda t_2} = n_1 n_2 \lambda^2 e^{-\lambda(n_1 t_1 + n_2 t_2)}$$

Calcularemos el logaritmo de esta función como paso previo para hallar el valor de $\lambda$ que la maximiza.

$$\ln L = \ln n_1 + \ln n_1 + 2 \ln \lambda - \lambda(n_1 t_1 + n_2 t_2)$$

Derivando con respecto a $\lambda$ e igualando a cero obtenemos

$$\frac{2}{\hat{\lambda}} - (n_1 t_1 + n_2 t_2) = 0 \quad \Rightarrow \quad \hat{\lambda} = \frac{2}{n_1 t_1 + n_2 t_2}$$

que es el estimador de máxima verosimilitud que estábamos buscando. En nuestro caso, con los datos muestrales, $\hat{\lambda} = 2.67 \times 10^{-4}$ horas$^{-1}$.

**b)** Veamos la distribución de las variables con las que tenemos que trabajar. Sea $Z_1 = n_1 T_1$. Como $T_1 \equiv EX[n_1\lambda]$, es obvio que $Z_1 \equiv EX[\lambda]$. Y de forma equivalente $Z_1 = n_2 T_2 \equiv EX[\lambda]$.

Por tanto, si llamamos $\theta = n_1 T_1 + n_2 T_2$ es evidente que $\theta \equiv \Gamma[2, \lambda]$

Tenemos que buscar una variable cuya distribución no dependa de $\lambda$. Como siempre en el caso de este tipo de distribuciones, si definimos $U = \lambda\theta$, obtenemos $U \equiv \Gamma[2, 1]$, distribución que no depende del parámetro. Si utilizamos un intervalo centrado en probabilidad, y teniendo en cuenta la distribución anterior, obtenemos

$P[U > a_1] = \frac{\alpha}{2} = 0.05 \quad \Longrightarrow \quad a_1 = 4.75$

$P[U \leq a_2] = \frac{\alpha}{2} = 0.05 \quad \Longrightarrow \quad a_2 = 0.35$

Pero

$$U = \lambda\theta = \lambda\frac{2}{\hat{\lambda}}$$

Y como $\hat{\lambda} = 2.67 \times 10^{-4}$ obtenemos el siguiente intervalo de confianza sobre el parámetro $\lambda$

$$IC_{90\,\%} = [0.467 \times 10^{-4}, 6.34 \times 10^{-4}]$$

---

**3.–** Una cierta población tiene una función de densidad que puede expresarse como $f_X(x) = \frac{2}{\theta^2}(\theta - x), \quad 0 < x < \theta$. Consideramos un estimador de $\theta$ dado por $T = 2x_1$, donde $x_1$ es una muestra de tamaño 1 de la población. Hallar un intervalo de confianza sobre $\theta$ $(1-\alpha = 0.95)$ utilizando la técnica que se desee. ¿Es $T$ un estimador de máxima verosimilitud de $\theta$?

SOLUCIÓN

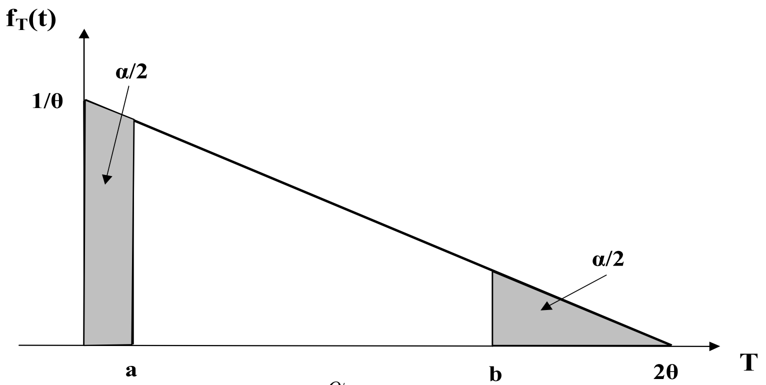
Veamos cuál es la distribución del estimador $\overline{\theta} = T = 2x_1$

$$f_T(t) = \frac{dx}{dt} f_X(x) = \frac{1}{2}\frac{2}{\theta^2}(\theta - \frac{t}{2}) = \frac{1}{\theta^2}(\theta - \frac{t}{2}) \text{ con } R_T = [0, 2\theta]$$

y

$$F_T(t) = \int_0^t \frac{1}{\theta^2}(\theta - \frac{t}{2})dt = \frac{1}{\theta^2}(\theta t - \frac{t^2}{4})$$

Queremos hallar los valores $a$ y $b$ tales que $P[a \leq T \leq b] = 1 - \alpha = 0.95$ (ver figura)




$$P[0 \leq T \leq a] = F_T(a) - F_T(0) = \frac{\alpha}{2}; \quad P[b \leq T \leq 2\theta] = F_T(2\theta) - F_T(b) = \frac{\alpha}{2}$$

luego

$$F_T(a) = \frac{1}{\theta^2}(\theta a - \frac{a^2}{4}) = 0.025 \implies a = 0.02516\theta$$

$$1 - F_T(b) = 1 - \frac{1}{\theta^2}(\theta b - \frac{b^2}{4}) = 0.025 \implies b = 1.68338\theta$$

Por lo tanto

$$0.95 = P[0.025516\theta \leq T \leq 1.68338\theta] = P[0.012258\theta \leq x_1 \leq 0.84419\theta] =$$

$$= P\left[\frac{1}{0.012258\theta} \geq \frac{1}{x_1} \geq \frac{1}{0.84419\theta}\right]$$

por lo que el intervalo de confianza del 95 % sobre $\theta$ es

$$IC_{95\,\%} = [1.1878x_1,\ 79.4913x_1]$$

Comprobemos ahora si $T = 2x_1$ es un estimador de máxima verosimilitud. La función de verosimilitud es, en este caso

$$L(\theta) = \frac{2}{\theta^2}(\theta - x_1)$$

Tomando logaritmos para calcular más fácilmente el máximo

$$\ln L(\theta) = \ln 2 - 2\ln\theta + \ln(\theta - x_1)$$

Derivando con respecto a $\theta$

$$\frac{\partial \ln L(\theta)}{\partial\theta} = -\frac{2}{\hat{\theta}} + \frac{1}{\hat{\theta} - x_1} = 0 \Rightarrow \hat{\theta} = 2\hat{\theta} - 2x_1 \Rightarrow \hat{\theta} = 2x_1$$

luego efectivamente $T = 2x_1$ es el estimador de máxima verosimilitud de $\theta$

---

**4.–** Para medir la vibración de una sonda de perforación de roca se utilizan dos medidores Doppler, de diferentes fabricantes, que utilizan distintos generadores de rayo laser. Sobre una sonda concreta se realizan 200 ensayos con uno de los medidores, obteniéndose una media de frecuencia $\bar{x}_1 = 3.5 \times 10^3\ Hz$ y una desviación típica de $S_1 = 0.2 \times 10^3\ Hz$. El resultado de 150 ensayos con el segundo medidor da como resultados $\bar{x}_2 = 3.4 \times 10^3\ Hz$ y $S_2 = 0.15 \times 10^3\ Hz$. Para comparar dichos medidores de acuerdo con la normativa se debe calcular un intervalo de confianza $(1 - \alpha = 0.99)$ sobre la diferencias de las frecuencias medias poblacionales medidas por ambos aparatos (los aparatos no miden directamente la frecuencia de vibración de la sonda, sino la detección de un cambio Doppler en la frecuencia de la luz coherente dispersada por un objetivo en movimiento, del cual se obtiene una medición resuelta en el tiempo de la velocidad del objetivo). Calcular dicho intervalo silas desviaciones típicas son las verdaderas $(S_1 = \sigma_1,\ \ S_2 = \sigma_2)$.

A la vista de los resultados y sin realizar más cálculos, ¿pueden considerarse iguales las medias proporcionadas por ambos aparatos?

——————SOLUCIÓN——————

Si las desviaciones típicas son las verdaderas utilizaremos como estadístico

$$\frac{\bar{x}_1 - \bar{x}_2 - (m_1 - m_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \equiv N(0,1)$$

Entonces, para $\alpha = 1\,\%$, el intervalo de confianza en la diferencias de medias es

$$IC_{99\,\%} = \left[(\bar{x}_1 - \bar{x}_2) - 2.576\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}, (\bar{x}_1 - \bar{x}_2) + 2.576\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right]$$

Sustituyendo valores

$$\boxed{IC_{99\,\%} = [0.0518 \times 10^3, 0.1482 \times 10^3]}$$

Dado que el valor $m_1 - m_2 = 0$ no está incluido en los intervalos de confianza calculados, y dado el valor bastante alto de los tamaños de muestra, podemos concluir que ambos aparatos dan resultados distintos para el mismo fenómeno físico.

---

**5.**– Se modela un sistema ecológico de manera que la especie animal dominante consume un tanto por uno, $X$, $\ 0 \leq X \leq 1$, de los recursos totales del sistema que denotamos $R$. Se ha supuesto que $X$ tiene la función de densidad de probabilidad:

$$f_X(x) = 2(1-x), \quad 0 \leq x \leq 1$$

**a)** Hallar la función de densidad de probabilidad de $W = XR$, el total de recursos consumido por la especie animal.

**b)** Suponiendo que se dispone de una muestra de $W$, $\ w_1, w_2, ...w_n$, determinar cómo se calcularía el estimador de máxima verosimilitud de $R$. Particularizar el resultado para el caso de una muestra de tamaño 1.

———————SOLUCIÓN———————

**a)** El total de los recursos consumidos por la especie animal es $W = XR$. Como esta relación es creciente en $X$,

$$f_W(w) = \left|\frac{dx}{dw}\right| f_X\left(\frac{w}{R}\right) = \frac{1}{R}2(1 - \frac{w}{R}) = \frac{2(R-w)}{R^2}, \quad 0 \leq w \leq R$$

y es obvio que

$$\int_0^R \frac{2(R-w)}{R^2} \; dw = \frac{2}{R^2}\left[Rw - \frac{w^2}{2}\right]_0^R = 1$$

**b)** La verosimilitud de una muestra de $W$, como función del parámetro $R$, es

$$L(R|w_1, w_2, ...w_n) = \prod_{i=1}^{n} \frac{2(R-w_i)}{R^2}$$

Tomando logaritmos

$$\ln L(R|w_1, w_2, ...w_n) = \sum_{i=1}^{n} \ln \frac{2(R-w_i)}{R^2} = \sum_{i=1}^{n} \left[\ln(2(R-w_i)) - 2\ln R\right]$$

Derivando con respecto a $R$

$$\frac{d\ln L}{dR} = \sum_{i=1}^{n} \left(\frac{1}{R-w_i}\right) - \frac{2n}{R}$$

y la ecuación a resolver para calcular el estimador de máxima verosimilitud de $R$ sería

$$\sum_{i=1}^{n} \left( \frac{1}{\hat{R} - w_i} \right) - \frac{2n}{\hat{R}} = 0$$

En general, esta ecuación se resolvería para un conjunto de valores $w_1, w_2, ...w_n$ mediante cualquier método numérico. Para el caso $n = 1$

$$\frac{1}{\hat{R} - w_1} - \frac{2}{\hat{R}} = 0 \implies \hat{R} = 2w_1$$