

CONTROL 3

MA3403 - PROBABILIDADES Y ESTADISTICA

PROFESOR: FERNANDO LEMA

AUXILIARES: MARTÍN CASTILLO & JOSÉ CERECEDA
16 DE NOVIEMBRE DE 2014

P1. a) Sea $X \sim U(0, \theta)$; $X_1, \dots, X_n$ una m.a.

1) **(2 puntos)** Se propone el estimador de θ dado por:

$$\hat{\theta} = k\bar{X}.$$

Determine k para minimizar el Error Cuadrático Medio.

R:

$$ECM = \mathbb{E}((\hat{\theta} - \theta)^2) = k^2 \mathbb{E}(\bar{X}^2) - 2k\theta \mathbb{E}(\bar{X}) + \theta^2.$$

Derivando e igualando a cero obtenemos:

$$k^* = \theta \cdot \frac{\mathbb{E}(\bar{X})}{\mathbb{E}(\bar{X}^2)} = \theta \frac{\frac{\theta}{2}}{\left(\frac{\theta}{2}\right)^2 + \frac{\theta^2}{12n}} = \frac{6}{3 + \frac{1}{n}}.$$

2) **(2 puntos)** Suponga $k = 2$. Muestre o justifique que, para n grande, $\hat{\theta} \sim N\left(\theta, \frac{\theta^2}{3n}\right)$. Use lo anterior para construir un Intervalo de Confianza para θ , con confianza $1 - \alpha$.

R: Como n es grande podemos usar TCL y entonces:

$$\frac{\sum X_i - n\mathbb{E}(X_i)}{\sqrt{n}\sqrt{\text{Var}(X_i)}} \sim N(0, 1) \Leftrightarrow \bar{X} \sim N\left(\mathbb{E}(X_i), \frac{\text{Var}(X_i)}{n}\right) \sim N\left(\frac{\theta}{2}, \frac{\theta^2}{12n}\right).$$

Luego,

$$\hat{\theta} = 2\bar{X} \sim N\left(\theta, \frac{\theta^2}{3n}\right).$$

Usando lo anterior queremos encontrar a tal que:

$$\mathbb{P}(\hat{\theta} - a \leq \theta \leq \hat{\theta} + a) = 1 - \alpha.$$

Pero desarrollando el lado izquierdo llegamos a:

$$\mathbb{P}(\hat{\theta} - a \leq \theta \leq \hat{\theta} + a) = 1 - 2 \left(1 - \Phi\left(\frac{a\sqrt{3n}}{\theta}\right)\right).$$

Así el a que buscamos debe cumplir con:

$$a = \frac{\theta}{\sqrt{3n}} \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \approx \frac{\hat{\theta}}{\sqrt{3n}} \Phi^{-1}\left(1 - \frac{\alpha}{2}\right).$$

- b) Sea $X \sim N(\mu_x, \sigma_x^2)$, $Y \sim N(\mu_y, \sigma_y^2)$ independientes y $(X_1, \dots, X_n), (Y_1, \dots, Y_m)$ m.a. de X e Y respectivamente. Considere σ_x^2, σ_y^2 conocidos; $\mu_x = \mu_y = \mu$ parámetro desconocido.

1) **(1 punto)** Determine el Estimador de Máxima Verosimilitud de μ .

R: Primero determinemos la Función Verosimilitud:

$$L(\mu, X_1, \dots, X_n, Y_1, \dots, Y_m) = \prod_{i=1}^n \prod_{k=1}^m f_X(X_i|\mu) \cdot f_Y(Y_k|\mu).$$

En donde:

$$f_X(X_i|\mu) = \frac{1}{\sqrt{2\pi}\sigma_x} \cdot e^{-\frac{(X_i-\mu)^2}{2\sigma_x^2}}, \quad f_Y(Y_k|\mu) = \frac{1}{\sqrt{2\pi}\sigma_y} \cdot e^{-\frac{(Y_k-\mu)^2}{2\sigma_y^2}}.$$

Luego, nos planteamos un problema de maximización equivalente dado por:

$$\max_{\mu} \ln \left((\sqrt{2\pi})^{n+m} \sigma_x^n \sigma_y^m L(\mu|\dots) \right) \Leftrightarrow \max_{\mu} - \sum_{i=1}^n \frac{(X_i - \mu)^2}{2\sigma_x^2} - \sum_{k=1}^m \frac{(Y_k - \mu)^2}{2\sigma_y^2}.$$

Luego derivando con respecto a μ , igualando a cero y luego despejando obtenemos el siguiente estimador:

$$\hat{\mu} = \frac{\sum_{i=1}^n \frac{X_i}{\sigma_x^2} + \sum_{k=1}^m \frac{Y_k}{\sigma_y^2}}{\frac{n}{\sigma_x^2} + \frac{m}{\sigma_y^2}}.$$

2) **(1 punto)** Analice los casos: $\frac{\sigma_x^2}{\sigma_y^2} \rightarrow 0$; $\frac{\sigma_x^2}{\sigma_y^2} \rightarrow \infty$; $\frac{\sigma_x^2}{\sigma_y^2} \rightarrow 1$.

- R:**
- Si $\frac{\sigma_x^2}{\sigma_y^2} \rightarrow 0$, entonces $\hat{\mu} = \bar{X}$.
 - Si $\frac{\sigma_x^2}{\sigma_y^2} \rightarrow \infty$, entonces $\hat{\mu} = \bar{Y}$.
 - Si $\frac{\sigma_x^2}{\sigma_y^2} \rightarrow 1$, entonces $\hat{\mu} = \frac{\sum X_i + \sum Y_k}{n+m}$.

P2. Considere una marca de ampollitas cuya duración es una v.a. $X \sim N(\mu, \sigma^2)$.

- a) **(2 puntos)** Considere $\sigma^2 = 16$. Se tomó una m.a. de tamaño $n = 25$ obteniéndose $\bar{X} = 170$.

Determine un Intervalo de Confianza para μ con confianza del 90%. Como cambia su respuesta si σ^2 es desconocido pero $\hat{\sigma}^2 = 16$.

R: Considerando el estadístico $Z = \frac{\bar{X} - \mu}{\sigma} \sqrt{n} \sim N(0, 1)$, imponemos

$$\begin{aligned}\mathbb{P}(-z \leq Z \leq z) &= 2\Phi(z) - 1 = 0,9 \\ \Rightarrow \Phi(z) &= ,95 \\ \Rightarrow z &= 1,65 \\ \Rightarrow \mathbb{P}(-z \leq Z \leq z) &= \mathbb{P}(-1,65 \leq Z \leq 1,65) \\ &= \mathbb{P}\left(-1,65 \leq \frac{\bar{X} - \mu}{\sigma} \sqrt{n} \leq 1,65\right) \\ &= \mathbb{P}\left(\bar{X} - \frac{1,65 \cdot \sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + \frac{1,65 \cdot \sigma}{\sqrt{n}}\right)\end{aligned}$$

luego $\mu \in [168,68; 171,32]$ con un 90 % de confianza.

Ahora poniéndose en el caso de σ desconocido, se tiene que $T = \frac{\bar{X} - \mu}{\hat{\sigma}} \sqrt{n} \sim t_{24}$ (sigue una t -student de 24 grados de libertad), entonces imponemos para T

$$\begin{aligned}\mathbb{P}(-z \leq T \leq z) &= 2\Phi_T(z) - 1 = 0,9 \\ \Rightarrow \Phi_T(z) &= ,95 \\ \Rightarrow z &= 1,71 \\ \Rightarrow \mathbb{P}(-z \leq T \leq z) &= \mathbb{P}(-1,71 \leq Z \leq 1,71)\end{aligned}$$

De igual manera que antes $\mu \in [168,63; 171,36]$ con un 90 % de confianza.

b) **(1 punto)** Considere $\sigma^2 = 16$. ¿Con que confianza se construyó el intervalo para μ : (169, 171)?

R: Como se trata de un intervalo simétrico, se tiene que $\mu \in (169, 171) = (\bar{X} - \frac{\sigma \cdot z}{\sqrt{n}}, \bar{X} + \frac{\sigma \cdot z}{\sqrt{n}})$

$$\begin{aligned}\frac{\sigma \cdot z}{\sqrt{n}} &= 1 \\ \Rightarrow z &= \frac{\sqrt{n}}{\sigma} = 5/4\end{aligned}$$

Luego la confianza del intervalo es $2\Phi(5/4) - 1 = 2 \cdot 0,89 - 1 = 0,78$

c) Otra marca de ampolletas tiene duración $Y \sim N(165, 36)$. Se toman dos ampolletas X y tres Y . Calcule

1) **(2 puntos)** $\mathbb{P}(\bar{Y} > \bar{X})$; asuma $X \sim N(170, 16)$.

R: Como Y_1, Y_2, Y_3, X_1, X_2 son independientes, se tiene que: $\bar{X} \sim N(170, \frac{16}{2})$ y $\bar{Y} \sim N(165, \frac{36}{3})$. Entonces $Z = \bar{Y} - \bar{X} \sim N(-5, 20)$. Así:

$$\mathbb{P}(\bar{Y} > \bar{X}) = \mathbb{P}(Z > 0) = \mathbb{P}\left(\frac{Z + 5}{\sqrt{20}} > \frac{5}{\sqrt{20}}\right)$$

En donde $\frac{Z+5}{\sqrt{20}} \sim N(0, 1)$ y entonces la probabilidad buscada es:

$$1 - \Phi\left(\frac{5}{\sqrt{20}}\right) \approx 0,133.$$

- 2) **(1 punto)** ¿Cuántas ampollitas Y deben probarse para encontrar, en promedio, una que dure menos de 155?

R: Primero calculemos

$$\begin{aligned}\mathbb{P}(Y < 155) &= \mathbb{P}(Y - 155 < 0) = \mathbb{P}(Y - 165 < -10) \\ &= \mathbb{P}\left(\frac{Y - 165}{6} < \frac{-10}{6}\right) \\ &= \Phi(-1,66) = 1 - \Phi(1,66) \\ &= 1 - 0,95 = 0,05 = p\end{aligned}$$

Ahora notemos la v.a. Z que define el experimento de probar hasta que ocurra $Y > 155$ es una geométrica, definida por

$Z = i$ si en las primeras $i - 1$ pruebas ocurre ($Y > 155$) en la prueba i se tiene ($Y < 155$), luego Z es una geométrica de parámetro p , luego buscamos en promedio la cantidad de pruebas, y esto es $\mathbb{E}(Z) = 1/p = 20$

- P3.** a) **(3 puntos)** Un Jugador compulsivo y supersticioso juega todos los fines de semana una cantidad de dinero aleatoria dada por: El lanza un dardo (en forma Uniforme) a un disco de radio 3. Si queda a menos de 1 unidad del centro apuesta 5 (UM). Si queda entre 1 unidad y 2 unidades, apuesta 2 (UM); en caso contrario no apuesta.

Calcule la probabilidad que en un año el jugador apueste más de 70 (UM).

R: Si notamos por X la v.a. que determina la cantidad apostada. Entonces:

$$\mu = \mathbb{E}(X) = \frac{11}{9}.$$

$$\sigma^2 = \text{Var}(X) = \frac{37}{9} - \left(\frac{11}{9}\right)^2.$$

Luego la idea es aprovechar que 52 es grande y usar TCL. Así,

$$\mathbb{P}\left(\sum_{i=1}^{52} X_i > 70\right) = \mathbb{P}\left(\underbrace{\frac{\sum X_i - n\mu}{\sigma\sqrt{n}}}_{\sim N(0,1)} > \frac{70 - 52\mu}{\sigma\sqrt{52}} = 0,55\right) = 1 - \Phi(0,55) \approx 0,292.$$

- b) Suponga que se debe estimar el área A de un rectángulo de lados x , y desconocidos ($A = xy$). Se dispone de n mediciones de x y de y respectivamente, en donde cada medición tiene asociado un error:

$$X_i = x + \epsilon_i, \quad \epsilon_i \sim N(0, 1), \quad \forall i = 1, \dots, n.$$

$$Y_i = y + \eta_i, \quad \eta_i \sim N(0, 1), \quad \forall i = 1, \dots, n.$$

Considere los estimadores basados en las muestras $(X_1, \dots, X_n)$, $(Y_1, \dots, Y_n)$ dados por:

$$\hat{A}_1 = \overline{X} \cdot \overline{Y},$$

$$\hat{A}_2 = \frac{1}{n} \sum_{i=1}^n X_i \cdot Y_i.$$

1) **(1.5 puntos)** Muestre que $\hat{A}_1$ y $\hat{A}_2$ son estimadores insesgados de A .

R: Antes que todo es importante notar que: $X_i \sim N(x, 1)$, $Y_i \sim N(y, 1)$, $\overline{X} \sim N\left(x, \frac{1}{n}\right)$ y $\overline{Y} \sim N\left(y, \frac{1}{n}\right)$.

Luego:

$$\mathbb{E}(\hat{A}_1) = \mathbb{E}(\overline{X})\mathbb{E}(\overline{Y}) = xy = A.$$

$$\mathbb{E}(\hat{A}_2) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}(X_i)\mathbb{E}(Y_i) = \frac{nxy}{n} = A.$$

2) **(1.5 puntos)** Si $Var(\hat{A}_2) = \frac{x^2+y^2+1}{n}$ (no la calcule). Calcule la varianza de $\hat{A}_1$ y decida cual estimador es mejor.

R:

$$Var(\hat{A}_1) = \mathbb{E}(\overline{X}^2 \overline{Y}^2) - \mathbb{E}(\overline{X} \overline{Y})^2$$

En donde:

■

$$\mathbb{E}(\overline{X} \overline{Y})^2 = (ab)^2.$$

■

$$\begin{aligned} \mathbb{E}(\overline{X}^2 \overline{Y}^2) &= \mathbb{E}(\overline{X}^2) \mathbb{E}(\overline{Y}^2) = \left(Var(\overline{X}) + \mathbb{E}(\overline{X}^2)\right) \left(Var(\overline{Y}) + \mathbb{E}(\overline{Y}^2)\right) \\ &= \left(x^2 + \frac{1}{n}\right) \left(y^2 + \frac{1}{n}\right). \end{aligned}$$

Juntando todo llegamos a que:

$$Var(\hat{A}_1) = \frac{x^2 + y^2 + \frac{1}{n}}{n} < Var(\hat{A}_2).$$

Por lo tanto $\hat{A}_1$ es mejor estimador que $\hat{A}_2$.