

PAUTA CONTROL 3

MA3403 - PROBABILIDADES Y ESTADISTICA

PROFESOR: FERNANDO LEMA

AUXILIARES: MARTÍN CASTILLO & JOSÉ CERECEDA

23 DE JUNIO DE 2014

P1. a) La demanda diaria de pan (en miles de Kg) en una manaderia es v.a. con densidad

$$f_X(x) = \begin{cases} 2x & x \in (0, 1) \\ 0 & \sim \end{cases}$$

Calcule la probabilidad que la demanda total en un mes (30 dias) supere los 18000 Kg.

Solución: Primero calculando $\mathbb{E}(X)$

$$\begin{aligned} \mathbb{E}(X) &= \int_{\Omega_X} x f_X(x) dx \\ &= \int_0^1 x \cdot 2x dx \\ &= \frac{2}{3} x^3 \Big|_0^1 = \frac{2}{3} = 0,67 \end{aligned}$$

Y calculando $\mathbb{E}(X^2)$ para obtener $\mathbb{V}(X)$

$$\begin{aligned} \mathbb{E}(X^2) &= \int_{\Omega_X} x^2 f_X(x) dx \\ &= \int_0^1 x^2 \cdot 2x dx \\ &= \frac{2}{4} x^4 \Big|_0^1 = \frac{1}{2} \end{aligned}$$

luego $\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = 1/2 - 4/9 = 0,06$, ahora lo que deseamos calcular es

$$\mathbb{P}\left(\sum_{i=1}^{30} X_i \geq 18\right) = \mathbb{P}\left(\frac{\bar{X} - \mu}{\sigma} \sqrt{30} \geq \frac{\frac{18}{30} - \mu}{\sigma} \sqrt{30}\right)$$

en virtud de TCL se tiene que $Z = \frac{\sum_{i=1}^{30} X_i - \mu}{\sigma} \sqrt{30} \sim N(0, 1)$, luego

$$\begin{aligned} \mathbb{P}\left(\sum_{i=1}^{30} X_i \geq 18\right) &= \mathbb{P}\left(Z \geq \frac{\frac{18}{30} - 0,67}{0,24} \sqrt{30}\right) \\ &= \mathbb{P}(Z \geq -1,6) \\ &= 1 - \Phi(-1,6) \\ &= 1 - 0,05 = 0,95 \end{aligned}$$

b) Suponga ahora que la demanda diaria se modela por una v.a. con densidad

$$f_X(x) = \begin{cases} \frac{1}{\theta} x^{\frac{1}{\theta}-1} & x \in (0, 1) \\ 0 & \sim \end{cases}$$

Sea $X_1, \dots, X_n$ una m.a.s.

- 1) Determine el estimador $\theta(\hat{\theta})$ por el método de los momentos.

Solución:

Primero se tiene que

$$\begin{aligned}\mathbb{E}(X) &= \int_0^1 x f_X(x) dx \\ &= \int_0^1 x \frac{1}{\theta} x^{\frac{1}{\theta}-1} dx \\ &= \frac{1}{\theta} \int_0^1 x^{\frac{1}{\theta}} dx \\ &= \frac{1}{\theta} \cdot \frac{1}{1/\theta + 1} x^{1/\theta+1} \Big|_0^1 \\ &= \frac{1}{1+\theta}\end{aligned}$$

Luego de acuerdo al método de los momentos, aproximamos $\mathbb{E}(X)$ por $\bar{X}$, por lo que se tiene

$$\frac{1}{1+\hat{\theta}} = \bar{X} \Rightarrow \hat{\theta} = \frac{1}{\bar{X}} - 1$$

- 2) Determine el EMV $\theta(\hat{\theta})$ y muestre que para n grande $\hat{\theta} \sim N\left(\theta, \frac{\theta^2}{n}\right)$

Solución:

Definiendo la función de verosimilitud se tiene que

$$\begin{aligned}L(X_1, \dots, X_n; \theta) &= \prod_{i=1}^n \{\theta^{-1} x_i^{1/\theta-1}\} \\ &= \frac{1}{\theta^n} \prod_{i=1}^n \{x_i^{1/\theta-1}\} \\ \log(L) &= \log\left(\frac{1}{\theta^n} \prod_{i=1}^n \{x_i^{1/\theta-1}\}\right) \\ &= -n \log(\theta) + \left(\frac{1}{\theta} - 1\right) \sum_{i=1}^n \log(x_i) \\ \frac{\partial \log(L)}{\partial \theta} &= \frac{-n}{\theta} + \frac{-1}{\theta^2} \sum_{i=1}^n \log(x_i) \quad ! = 0 \\ \hat{\theta} &= -\frac{\sum_{i=1}^n \log(x_i)}{n}\end{aligned}$$

Por otro lado por TCL,

$$\frac{\hat{\theta} - \mathbb{E}(\log(x_i))}{\sqrt{\frac{\text{Var}(\log(x_i))}{n}}} \sim N(0, 1)$$

Luego para n grande se tiene $\hat{\theta} \sim N\left(\mathbb{E}(\log(x_i)), \frac{\text{Var}(\log(x_i))}{n}\right) = N\left(\theta, \frac{\theta^2}{n}\right)$. Pues $\log(X_i)$ distribuye como una exponencial de parámetro $\frac{1}{\theta}$.

- c) Usando b) 2) construya un I.C. para θ con una confianza del 90 %.

Solución:

Por la parte anterior se tiene que $\frac{\hat{\theta}-\theta}{\theta}\sqrt{n} \sim N(0,1)$, por lo que se plantea la siguiente ecuación

$$\begin{aligned} 2\Phi(z) - 1 &= \mathbb{P}\left(-z \leq \frac{\hat{\theta}-\theta}{\theta}\sqrt{n} \leq z\right) \\ &= \mathbb{P}\left(\frac{-z}{\sqrt{n}} + 1 \leq \frac{\hat{\theta}}{\theta} \leq \frac{z}{\sqrt{n}} + 1\right) \\ &= \mathbb{P}\left(\frac{\sqrt{n}}{\sqrt{n}-z} \geq \frac{\theta}{\hat{\theta}} \geq \frac{\sqrt{n}}{\sqrt{n}+z}\right) \\ &= \mathbb{P}\left(\frac{\sqrt{n}}{\sqrt{n}-z}\hat{\theta} \geq \theta \geq \hat{\theta}\frac{\sqrt{n}}{\sqrt{n}+z}\right) \end{aligned}$$

Luego con un z tal que $2\Phi(z) - 1 = 0,9$, el intervalo aleatorio $\left(\frac{\sqrt{n}}{\sqrt{n}+z}\hat{\theta}, \frac{\sqrt{n}}{\sqrt{n}-z}\hat{\theta}\right)$ tiene probabilidad 0.9 de contener a θ .

P2. Suponga que las notas de hombres (H) y mujeres (M) son v.a. independientes tales que $H \sim N(40; 1^2)$, $M \sim N(45; 0,5^2)$

- a) Calcule la probabilidad $\mathbb{P}(M > H)$ **Solución:**
Lo que se busca es la siguiente probabilidad

$$\begin{aligned} \mathbb{P}(M > H) &= \mathbb{P}(M - H > 0) \\ Z = (M - H) &\sim N(45 - 40; 1^2 + 0,5^2) = N(5, 1,25) \\ &= \mathbb{P}\left(\frac{Z - 5}{\sqrt{1,25}} > \frac{-5}{\sqrt{1,25}}\right) \\ &= 1 - \Phi(-4,47) \sim 1 \end{aligned}$$

- b) Se eligen 3 hombres y 2 mujeres (de forma independiente) calcule la probabilidad que su nota promedio sea mayor a 4.

Solución:

Se busca la siguiente probabilidad $\mathbb{P}\left(\frac{3H+2M}{5} \geq 4\right)$, pero se tiene que

$$\frac{3H+2M}{5} \sim N\left(\frac{3\mu_H + 2\mu_M}{5}; \frac{3\sigma_H^2 + 2\sigma_M^2}{5}\right)$$

- c) Considere ahora $H \sim N(\mu, \sigma^2)$. Se toma una m.a. de 9 hombres obteniéndose el intervalo de confianza simétrico (3.7;4.7) ¿Con qué confianza se construyó dicho intervalo?

Solución:

Se tiene que

$$2\Phi(z) - 1 = \mathbb{P}\left(\bar{X} - \frac{z\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + \frac{z\sigma}{\sqrt{n}}\right)$$

como es un intervalo simétrico se tiene que $\bar{X} = 4,2$ entonces $\frac{z\sigma}{\sqrt{n}} = 0,5$, entonces $z = 1,5$, luego la confianza de este intervalo es $2\Phi(1,5) - 1 = 2 \cdot 0,93 - 1 = 0,86$

- d) Suponga que de la lista de nombres y notas se pierden los nombres y a través de la nota se desea determinar el sexo de la persona. Se propone el siguiente procedimiento: se fija C y si la nota de una persona es mayor a C , entonces decimos que es mujer, y si es menor a C , decimos que es hombre.

Determine C que cumpla

$$\mathbb{P}(H > C) = \mathbb{P}(M < C)$$

¿Cómo interpreta esta igualdad?

Solución:

Tenemos que $Z = \frac{H-40}{1} = \frac{M-45}{0,5} \sim N(0, 1)$. Entonces c tiene que cumplir que,

$$\mathbb{P}(H > c) = \mathbb{P}(M < c) \iff \mathbb{P}(Z > c - 40) = \mathbb{P}(Z < 2c - 90).$$

Se pueden tomar 2 caminos:

- Camino 1: Dejar expresada la condición:

$$\Phi(c - 40) + \Phi(2c - 90) = 1.$$

- Camino 2: Ocupar la simetría de la Normal para encontrar una solución particular:
Imponer:

$$-(2c - 90) = c - 40 \iff c = \frac{130}{3} = 43,333$$

Encontrar c de esta manera nos permite asegurar que la probabilidad de equivocarnos al discriminar sea equiprobable tanto para Hombres como Mujeres.

- P3.** a) En una reciente encuesta hecha a 100 empresarios, el 55 % se mostró contrario a la reforma tributaria. Determine un I.C. del 95 % para la proporción de empresarios que esta en contra de la reforma tributaria. ¿Puede decir que la mayoría de los empresarios está en contra? ¿Con que confianza puede decir que la mayoría de los empresarios está en contra?

Solución:

Para esto definimos la variable aleatoria $X_i = \begin{cases} p & \text{en contra} \\ 1 - p & \text{a favor} \end{cases}$, por lo que tenemos que

$\mathbb{E}(X) = p$ y $\mathbb{V}(X) = p(1 - p)$, como no se conoce la proporción a priori, aproximamos

$\mathbb{V}(X) = \hat{p}(1 - \hat{p}) = 0,55 \cdot 0,45 = 0,25$, luego consideramos que $\frac{\hat{p} - p}{\sqrt{\hat{p}(1 - \hat{p})}} \sqrt{n} \sim N(0, 1)$

$$2\Phi(z) - 1 = \mathbb{P}\left(-z \leq \frac{\hat{p} - p}{\sqrt{\hat{p}(1 - \hat{p})}} \sqrt{n} \leq z\right)$$

imponiendo $\Phi(z) = 0,975 \Rightarrow z = 1,96$ entonces el intervalo $(0,55 - 1,96 \cdot \frac{0,5}{10}; 0,55 + 1,96 \cdot \frac{0,5}{10}) = (0,452, 0,648)$ contiene a p con probabilidad 0,95, por lo que no se puede decir que la mayoría de los empresarios están en contra.

Ahora para responder la segunda pregunta, imponemos $0,05 = z \cdot \frac{0,5}{10}$ (esto para que nuestro intervalo tenga como mínimo a 0,5) entonces $z = 10$ y el intervalo aleatorio $(0,5; 0,6)$ contiene a p con probabilidad $2\Phi(1) - 1 = 2 \cdot 0,84 - 1 = 0,68$

- b) Analice el siguiente argumento: Sea X v.a. tal que $X \sim U(0, 1)$, y definimos $X_1 = X$; $X_i = 0$ para $i \in \{2, 3, \dots\}$, por lo que podemos escribir $X = \sum_{i \geq 1} X_i$, entonces $X \sim U(0, 1)$ ó $X \sim$

$N(\mathbb{E}(X), \mathbb{V}(X))$

Solución:

Una de las hipótesis del TCL es que la varianza $\mathbb{V}(\sum X_i) \rightarrow \infty$ con $n \rightarrow \infty$, pero en este caso se tiene que $\mathbb{V}(\sum X_i) = \mathbb{V}(X) = \frac{1}{12}$, por lo que este experimento no cumple con las hipótesis, luego $X \sim U(0, 1)$

- c) Considere una v.a. X que se distribuye como una v.a. uniforme en $(-\alpha, \alpha)$, sea una m.a.s. $X_1, \dots, X_n$

- 1) Obtenga el estimador de los momentos para α .

Solución:

Primero observar que $\mathbb{E}(X) = 0$, por lo que no sirve el primer momento para estimar α , luego se tiene que $\mathbb{V}(X) = \frac{1}{3}\alpha^2$, luego aproximamos $\mathbb{V}(X)$ por $\frac{1}{n} \sum X_i^2$ entonces $\hat{\alpha} =$

$$\sqrt{\frac{3}{n} \sum X_i^2}.$$

- 2) Obtenga el estimador de máxima verosimilitud para α .

Solución:

Se tiene que la función densidad de la Uniforme es $f_X(x_i) = \frac{1}{2\alpha}$, luego nuestra función de verosimilitud esta dado por

$$L(X_1, \dots, X_n; \alpha) = \frac{1}{(2\alpha)^n}$$

para $-\alpha \leq X_i \leq \alpha \quad \forall i \in \{1, \dots, n\}$, que es equivalente a $\max\{|X_i|\} \leq \alpha$ luego como deseamos maximizar L , se tiene que entre más pequeño sea nuestro α , más grande será L y en virtud de lo anterior, el minimo valor que puede tomar α coincide con el $\max\{|X_i|\}$ luego, $\hat{\alpha} = \max\{|X_i|\}$