Probabilidad y Estadística - Introducción a la Probabilidad y Estadística

1	2	3	4	Total
12 / 13	15 / 10	25	15 / 7	97

10 (diez)

Apellido y Nombre: Achával Berzero Tomás **Carrera:** LCC

Justifique claramente todas sus respuestas.

Ejercicio 1. Sean X e Y variables aleatorias con función de densidad de probabilidad conjunta dada por:

$$f(x,y) = \begin{cases} \frac{x}{5} + c.y, & \text{si } 0 < x < 1,\ 1 < y < 5 \\ 0, & \text{c.c.} \end{cases}$$

a) Encontrar el valor de c.

b) Encontrar las funciones de densidad marginal f_X y f_Y.

c) ¿Son X e Y independientes?.

d) Encontrar la $P(X+Y>3)$.

Ejercicio 2. Se seleccionaron aleatoriamente 10 paquetes de galletas rotuladas bajas en sodio de una marca particular. El promedio muestral y desviación estándar muestral (s_{n-1}) para la cantidad de sodio, obtenidas por cada 100 gr, fueron de 122,1 y 2,5 mg respectivamente. Suponga que la muestra proviene de una distribución normal.

a) Dar la estimación por máxima verosimilitud para:

✓ i) El contenido de sodio medio (μ) y el desvío estándar poblacional (σ), para esta marca de galleta.

✓ ii) El percentil 80 para la variable contenido de sodio para esta marca de galletas.

b) Hallar un intervalo de confianza del 99 % para el contenido medio de sodio (μ) para esta marca de galletas.

Ejercicio 3. El artículo "Limited Yield Estimation for Visual Defect Sources" (IEEE Trans. on Semiconductor Manuf., 1997: 17-23) reportó que, en un estudio de un proceso de inspección de obleas particular, 356 troqueles fueron examinados por una sonda de inspección y 201 de éstos pasaron la prueba. Suponiendo un proceso estable:

a) Dar un intervalo de confianza aproximado del 95 % para la proporción de todos los troqueles que pasan la prueba (p).

b) Determinar el menor tamaño de muestra necesario que deben seleccionarse para conseguir un intervalo de confianza de longitud a lo sumo 0.05 y de nivel de confianza 0.95, independientemente del valor de $\hat{p}$.

Ejercicio 4. Sean $X_1,\ldots,X_n$ m.a. con distribución $\mathcal{P}(\lambda)$. Considere los siguientes estimadores para λ:

$$\hat{\lambda}_1 = \frac{X_1+\ldots+X_{n-1}}{n-1} \quad \text{y} \quad \hat{\lambda}_2 = \frac{X_1+\ldots+X_n}{n}$$

a) ¿$\hat{\lambda}_1$ es insesgado para estimar λ?. ¿Y $\hat{\lambda}_2$?.

b) i) Encuentre el error estándar de los estimadores $\hat{\lambda}_1$ y $\hat{\lambda}_2$.
ii) ¿Cuál de los dos estimadores es mejor para estimar λ?.

Ayuda: Recordar que $E(X_i)=\lambda$ y $V(X_i)=\lambda$

16/11/2023 PARCIAL II $P_Y E$ - TEMA A

Tomás Achával Berzero
LCC- 45085146 - A.

EJERCICIO Nº 1:

X e Y v.a con F.d.p.c dada por

$$f(x,y) = \begin{cases} \frac{x}{5} + c\cdot y & 0<x<1 \ , \ 1<y<5 \\ 0 & c.c \end{cases}$$

a) Encontrar el valor de c.

NOTACIÓN $\int^{\mathbb{R}} = \int_{-\infty}^{\infty}$

Sabemos que $\int^{\mathbb{R}}\int^{\mathbb{R}} f(x,y)\,dy\,dx = 1$

Esto es,

$$1 \overset{\star}{=} \int_0^1 \int_1^5 \left(\frac{x}{5} + c\cdot y\right) dy\,dx$$ debido a que $f(x,y)=0$ para el resto de los valores.

$$\bullet \int_1^5 \left(\frac{x}{5} + c\cdot y\right) dy = \int_1^5 \frac{x}{5}\,dy + c\int_1^5 y\,dy = \frac{x}{5}\int_1^5 1\,dy + c\overbrace{\left(\frac{5^2}{2} - \frac{1^2}{2}\right)}^{\frac{24}{2}=12}$$

$$= \frac{x}{5}(5-1) + 12c = \frac{4x}{5} + 12c$$

Luego

$$\star = \int_0^1 \left(\frac{4}{5}x + 12c\right) dx = \frac{4}{5}\int_0^1 x\,dx + 12\int_0^1 c\,dx = \frac{4}{5}\left(\frac{1^2}{2} - \frac{0^2}{2}\right) + 12c$$
$$= \frac{4}{5}\cdot\frac{1}{2} + 12c = \frac{2}{5} + 12c$$

Por lo tanto

$$\frac{2}{5} + 12c = 1 \Rightarrow 12c = \frac{3}{5} \Rightarrow c = \frac{3}{60} \Rightarrow \boxed{c = \frac{1}{20} = 0.05}$$

Corroboramos:

$$\int_0^1 \int_1^5 \left(\frac{x}{5} + \frac{y}{20}\right) dy\,dx = \int_0^1 \left(\frac{4x}{5} + \frac{1}{20}\left(\frac{25-1}{2}\right)\right) dx = \int_0^1 \left(\frac{4x}{5} + \frac{12}{20}\right) dx = \frac{4}{5}\int_0^1 x\,dx + \frac{12}{20}\int_0^1 1\,dx$$
$$= \frac{4}{5}\cdot\frac{1}{2} + \frac{12}{20} = \frac{2}{5} + \frac{3}{5} = \boxed{1}$$

NOTA

b) Encontrar f_X y f_Y

$$f_X(x) = \int^{\mathbb{R}} f(x,y)\,dy = \int_1^5 \left(\frac{x}{5} + \frac{y}{20}\right) dy = \frac{x}{5}\int_1^5 1\,dy + \frac{1}{20}\int_1^5 y\,dy$$

$$= \frac{4x}{5} + \frac{1}{20}\left(\frac{5^2}{2} - \frac{1^2}{2}\right) = \frac{4x}{5} + \underbrace{\frac{3}{5}}_{\frac{12}{20}}$$

$$\boxed{f_X(x) = \frac{4x+3}{5}}$$ si $x \in [0,1]$, 0 c.c

$$f_Y(y) = \int^{\mathbb{R}} f(x,y)\,dx = \int_0^1 \left(\frac{x}{5} + \frac{y}{20}\right) dx$$

$$= \int_0^1 \frac{x}{5}\,dx + \int_0^1 \frac{y}{20}\,dx = \frac{1}{5}\int_0^1 x\,dx + \frac{y}{20}\int_0^1 1\,dx$$

$$= \frac{1}{5}\cdot\frac{1}{2} + \frac{y}{20} = \frac{1}{10} + \frac{y}{20} =$$

$$\boxed{f_Y(y) = \frac{y+2}{20}}$$ si $y \in [1,5]$, 0 c.c

c) Son X e Y independientes? Para ello,

$$f(x,y) \stackrel{?}{=} f_X(x)\cdot f_Y(y) \quad \forall x,y \in \mathbb{R}, \text{ en particular}$$

$$\frac{x}{5} + \frac{y}{20} \neq \frac{(4x+3)}{5}\cdot\frac{(y+2)}{20} \quad \text{si } x \in [0,1], y \in (1,5)$$

$$\frac{4x+y}{20} \neq \frac{4xy+8x+3y+6}{100}$$

$$\frac{20x+5y}{100} \neq \frac{4xy+8x+3y+6}{100} \quad \text{pues } 20x+5y \neq 4xy+8x+3y+6$$

pues $\underbrace{12x+2y-4xy}_{\text{no es una función constante en } x\,(0,1);\,y\,(1,5)} \neq 6$

Por lo tanto como $\exists x,y$ tq $f(x,y) \neq f_X(x) f_Y(y)$, las variables X e Y **NO** son independientes.

NOTA

HOJA Nº

FECHA

Tomás Achával Busso

LCC-45085146-A.

d) $P(X+Y>3)$

$$= P(Y>3-X) = \underline{P((0<X<1)\cap(3-X<Y<5))}$$

$$= \underline{\int_0^1 \int_{3-X}^5} f(x,y)\,dy\,dx = \int_0^1 \int_{3-X}^5 \frac{x}{5}+\frac{y}{20}\,dy\,dx$$

$$= \int_0^1 \frac{x}{5}\left(5-(3-x)\right) + \frac{1}{20}\int_{3-x}^5 y\,dy \quad dx$$

$$= \int_0^1 x - \frac{3x}{5} + \frac{x^2}{5} + \frac{1}{20}\left(\frac{5^2}{2} - \frac{(3-x)^2}{2}\right) dx$$

$$= \int_0^1 \frac{2x}{5} + \frac{x^2}{5} + \frac{25}{40} - \frac{3^2-6x+x^2}{2\cdot 20}\,dx$$

$$= \int_0^1 \frac{2x+x^2}{5} + \frac{5}{8} - \frac{9}{40} + \frac{3x}{20} - \frac{x^2}{40}\,dx$$

$$= \int_0^1 \frac{1}{40}\left(16x+8x^2+25-9+6x-x^2\right)dx \quad \checkmark$$

$$= \frac{1}{40}\int_0^1 22x+7x^2+16\,dx$$

$$= \frac{11+7/3+16}{40} = \boxed{0.7333} = P(X+Y>3)$$

NOTA

Ejercicio N°2

$n=10 \quad \bar{x}=122,1 \quad S_{n-1}=2,5$ Distribución Normal.

a) Estimación por máxima verosimilitud para:

i) El contenido medio μ y el desvío estandar σ.

$$\hat{\mu}_{MV}=\bar{x}=\boxed{122,1}$$

$$\hat{\sigma}_{MV}=S_n=\sqrt{\frac{n-1}{n}}S_{n-1}=\sqrt{\frac{9}{10}}\cdot 2,5=\boxed{2.3717}$$

ii) $\eta(0.8)$

Sabiendo que $\Phi\left(\frac{\eta(0.8)-\mu}{\sigma}\right)=0.8$, obtenemos

$$\frac{\eta(0.8)-\mu}{\sigma}\simeq 0.84$$

$$\eta(0.8)\simeq \overbrace{0.84\cdot\sigma+\mu}^{h(\sigma,\mu)}$$

Por lo tanto utilizando la invarianza de los estimadores de MV,

$$\widehat{\eta(0.8)}_{MV}=0,84\,\hat{\sigma}_{MV}+\hat{\mu}_{MV}$$

$$=0,84\cdot 2,3717+122,1\simeq\boxed{124.0922}$$

b) IC del 99% $\Rightarrow \alpha=1-0.99=0.01\Rightarrow \boxed{\frac{\alpha}{2}=0.005}$

Como tenemos distribución normal con

$n<40$ y σ desconocido, el intervalo para confianza de nivel $1-\alpha$ está dado por

$$\bar{x}\pm t_{\frac{\alpha}{2};n-1}\cdot\frac{S_{n-1}}{\sqrt{n}},$$

Es decir

$$122,1\pm t_{0,005;n-1}\cdot\frac{2,5}{\sqrt{10}}=122,1\pm 3,25\cdot 0,7906$$

$$=\boxed{[119,5306\,;\,124,6694]}$$

NOTA

HOJA N°

FECHA

EJERCICIO N°3

Tomás Aguilar Bosco
LCC- M5085146-A.

$n = 356$

PASARON PRUEBA $= 207$ $\Big\}\ \hat{p} = \frac{207}{356} \simeq 0.5646$

COMO SE CUMPLE QUE

"$n >> 0$" ✓

$n \cdot \hat{p} = 356 \cdot \frac{207}{356} = 207 \geq 10$ ✓

$n(1-\hat{p}) = 356 \cdot \frac{155}{356} = 155 \geq 10$ ✓

LUEGO

a) UN INTERVALO DE CONFIANZA $1-\alpha = 0.95$ PARA p ESTÁ DADO POR
$(\alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025)$

$$\underbrace{\hat{p} \pm z_{0.025}\sqrt{\frac{\hat{p}(1-\hat{p})}{356}}} = 0.5646 \pm 1.96 \cdot \sqrt{\frac{0.5646\,(0.4354)}{356}}$$

$$z_{0.025} = \mathcal{N}(0.975) = 1.96$$

HACIENDO LOS CÁLCULOS OBTENGO

$$IC = [0.5131\ ;\ 0.6161]$$

b) EL VALOR DE n TQ $L_{IC} \leq 0.05$ CON $\alpha = 0.05$ SIN IMPORTAR EL VALOR DE $\hat{p}$
ESTÁ DADO POR

$$\underbrace{n \geq \left(\frac{z_{0.025}}{L}\right)^2} \Rightarrow n \geq \left(\frac{1.96}{0.05}\right)^2 \Rightarrow n \geq 1536.64$$

$$\Rightarrow \boxed{n \geq 1537}$$

ESTO SE DEBE A QUE

$L_{IC} = 2 \cdot z_{0.025} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ Y $\hat{p}(1-\hat{p})$ ALCANZA SU MÁXIMO EN $\hat{p} = \frac{1}{2}$, OBTENIENDO ASÍ

$$L_{IC} = z_{0.025} \cdot \cancel{2}\sqrt{\frac{1/4}{n}} \leq 0.05 \iff n \geq \left(\frac{z_{0.025}}{0.05}\right)^2$$

NOTA

Ejercicio N° 4

$X_1, \ldots, X_n$ m.a. con distribución $P(\lambda)$

$$\hat{\lambda}_1 = \frac{X_1+\cdots+X_{n-1}}{n-1} \qquad \hat{\lambda}_2 = \frac{X_1+\cdots+X_n}{n}$$

a) Son insesgados para λ? Sabiendo que $E(X_i)=\lambda$ para todo $i=1,\ldots,n$

Para ello, debemos ver

$E(\hat{\lambda}_1)=\lambda$?

$$E(\hat{\lambda}_1) = E\left(\frac{\sum_{i=1}^{n-1} X_i}{n-1}\right) = \frac{1}{n-1} E\left(\sum_{i=1}^{n-1} X_i\right) = \frac{1}{n-1} \cdot \sum_{i=1}^{n-1} E(X_i) = \frac{1}{n-1}(n-1)\lambda = \lambda$$

$\therefore \hat{\lambda}_1$ es insesgado para λ.

$E(\hat{\lambda}_2)=\lambda$?

$$E(\hat{\lambda}_2) = E\left(\frac{\sum_{i=1}^{n} X_i}{n}\right) = \frac{1}{n} E\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n} \cdot n\lambda = \lambda$$

$\therefore \hat{\lambda}_2$ es insesgado para λ.

b) i) El error estándar de un estimador $\hat{\theta}$ está dado por $\sigma_{\hat{\theta}} = \sqrt{V(\hat{\theta})}$

i.1) $$V(\hat{\lambda}_1) = V\left(\frac{\sum_{i=1}^{n-1} X_i}{n-1}\right) = \left(\frac{1}{n-1}\right)^2 V\left(\sum_{i=1}^{n-1} X_i\right) \underset{\substack{X_i \text{ mutuamente} \\ \text{independientes}}}{=} \frac{1}{(n-1)^2} \cdot \sum_{i=1}^{n-1} V(X_i) = \boxed{\frac{n\lambda}{(n-1)^2}}$$

i.2) $$V(\hat{\lambda}_2) = V\left(\frac{\sum_{i=1}^{n} X_i}{n}\right) = \frac{1}{n^2} V\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) = \frac{n\lambda}{n^2} = \boxed{\frac{\lambda}{n}}$$

Luego

$$\sigma_{\hat{\lambda}_1} = \sqrt{\frac{n\lambda}{(n-1)^2}} = \boxed{\frac{\sqrt{n\lambda}}{n-1}} \quad \text{y} \quad \boxed{\sigma_{\hat{\lambda}_2} = \sqrt{\frac{\lambda}{n}}}$$

ii) Entre dos estimadores insesgados para λ, el mejor será aquel que tenga la menor varianza.

Notar que $V(\hat{\lambda}_2) \leq V(\hat{\lambda}_1) \iff \frac{\lambda}{n} \leq \frac{n\lambda}{(n-1)^2} \iff \lambda \leq \lambda \underbrace{\frac{n^2}{(n-1)^2}}_{\geq 1} \iff n^2 \geq (n-1)^2 \ \checkmark \ \forall_{n \in \mathbb{N}}$

Por lo tanto $\hat{\lambda}_2$ es el mejor estimador para λ.

OK. Bien de acuerdo a la varianza que obtuvo que está mal.