Estatística II, Exame época especial, 2018/07/23 RESOLUÇÃO

1 Numa sessão de formação uma pessoa tenta desempenhar duas vezes seguidas certa tarefa, podendo, em cada tentativa, ter êxito ou falhar. A probabilidade de falhar a primeira tentativa é igual a 0,25. Se falhar a primeira, a probabilidade de ser bem-sucedida na segunda é de 0,5. Se for bem-sucedida na primeira, a probabilidade de falhar na segunda é 0,1. Qual a probabilidade de falhar a segunda tentativa?

Resolução

Fj: "falhar na tentativa j", $j = 1{,}2$. $\Pr(F1) = 0{,}25$

$\Pr(F2^c|F1) = 0{,}5 = 1 - \Pr(F2|F1)$ $\Pr(F2|F1^c) = 0{,}1$

$$\Pr(F2) = \Pr[F2 \wedge (F1 \vee F1^c)] = \Pr[(F1 \wedge F2) \vee (F1^c \wedge F2)] =$$

$$\Pr(F1 \wedge F2) + \Pr(F1^c \wedge F2) = \Pr(F2|F1)\Pr(F1) + \Pr(F2|F1^c)\Pr(F1^c) =$$

$$(1 - 0{,}5) \times 0{,}25 + 0{,}1 \times (1 - 0{,}25) = 0{,}2.$$

2 Classifique, justificando, cada uma das seguintes proposições como verdadeira ou falsa.

a) Seja f a função probabilidade de uma v.a. discreta; então $0 \leq f(x) \leq 1, \forall x \in \mathcal{R}$.

Res.

Verdadeira: se X é discreta, $f(x) = \Pr(X = x)$, logo, $0 \leq f(x) \leq 1, \forall x \in \mathcal{R}$.

b) Seja f a função densidade de uma v.a. contínua; então $0 \leq f(x) \leq 1, \forall x \in \mathcal{R}$.

Res.

Falsa: se X é contínua apenas se tem, em geral, $f(x) \geq 0, \forall x \in \mathcal{R}$.

c) Seja F a função de distribuição de uma v. a. contínua; então $0 \leq F(x) \leq 1, \forall x \in \mathcal{R}$.

Res.

Verdadeira: $F(x) = \Pr(X \leq x)$, logo, $0 \leq F(x) \leq 1, \forall x \in \mathcal{R}$.

d) Seja F a função distribuição de uma v.a. contínua e simétrica em relação à origem. Então, $F(-x) = 1 - F(x), \forall x \in \mathcal{R}$.

Res.

Verdadeira: se X é contínua, $F(-x) = \Pr(X \leq -x) = \Pr(X \geq x) = \Pr(X > x) = 1 - \Pr\Pr(X \leq x) = 1 - F(x), \forall x \in \mathcal{R}$.

3 Considere a função densidade da v.a. X, $f_X(x) = \begin{cases} x, & 0 < x < 1, \\ 1/2\,, & 1 \leq x < 2. \end{cases}$

a) Calcule $E(X)$ e $V(X)$.

Res.

$$E(X) = \int_0^1 x \cdot x dx + \int_1^2 x(1/2)dx = [x^3/3]_0^1 + [x^2/4]_1^2 = 1/3 + (1 - 1/4) = 13/12.$$

$$V(X) = E(X^2) - E(X)^2 = \int_0^1 x^2 \cdot x dx + \int_1^2 x^2(1/2)dx - (13/12)^2 =$$

$$[x^4/4]_0^1 + [x^3/6]_1^2 - (13/12)^2 = 1/4 + (8/6 - 1/6) - (13/12)^2 = 35/144.$$

b) Calcule a mediana de X.

Res.

$$Me\text{:}\ \Pr(X \leq Me) = 0{,}5 \Leftrightarrow \int_0^{Me} f_X(x)dx = 0{,}5 \Leftrightarrow \int_0^{Me} x dx = 0{,}5 \Leftrightarrow$$

$$[x^2/2]_0^{Me} = Me^2/2 = 0{,}5 \Leftrightarrow Me = 1.$$

c) Calcule a variância da variável $U = \begin{cases} -1, & X < 1/2\,, \\ 1, & X \geq 1/2\,. \end{cases}$

Res.

Note-se que $U^2 = 1$ para qualquer valor de X; donde, $E(U^2) = 1$.

$$V(U) = E(U^2) - E(U)^2 = 1 - [(-1)\Pr(U = -1) + 1\Pr(U = 1)]^2 =$$

$$1 - [\Pr(X \geq 1/2) - \Pr(X < 1/2)]^2 = 1 - \left(\int_{1/2}^{1} xdx + \int_{1}^{2} dx/2 - \int_{0}^{1/2} xdx\right)^2 =$$

$$1 - (3/8 + 1/2 - 1/8)^2 = 7/16.$$

4 Sejam $Y \sim \mathcal{N}(\mu, 1)$ e $(Y_1, \dots, Y_n)$ uma amostra casual desta população. Obtenha o estimador da máxima verosimilhança de μ.

Res.

$$Y \sim \mathcal{N}(\mu, 1) \equiv \exp[-(y - \mu)^2/2]/\sqrt{2\pi},$$

$$l(\mu) = \sum_{i=1}^{n} \log\left\{\exp[-(y_i - \mu)^2/2]/\sqrt{2\pi}\right\} = -\sum_{i=1}^{n} (y_i - \mu)^2/2 - n\log(2\pi)/2 \Rightarrow$$

$$l'(\mu) = 0 \Leftrightarrow \sum_{i=1}^{n} (y_i - \mu) = 0 \Leftrightarrow n\mu = \sum_{i=1}^{n} y_i \Leftrightarrow \mu = \sum_{i=1}^{n} y_i/n = \bar{y};$$

$$l''(\mu) = -n < 0, \forall\mu.$$

Estimador MV $\qquad \hat{\mu}_{MV} = \bar{Y}$.

5 Sejam $Y \sim \mathcal{N}(\mu, \sigma^2)$ e $(Y_1, \dots, Y_4)$ uma amostra casual desta população. O estimador $T = (Y_1 + Y_2 + 2Y_3)/4$ é cêntrico para μ? T é mais eficiente do que $\bar{Y}$? Justifique mediante cálculos adequados.

Res.

Centricidade de T: $E(T) = [E(Y_1) + E(Y_2) + 2E(Y_3)]/4 = (\mu + \mu + 2\mu)/4 = \mu$.

Eficiência:

$$V(T) = [V(Y_1) + V(Y_2) + 4V(Y_3)]/16 = (\sigma^2 + \sigma^2 + 4\sigma^2)/16 = 3\sigma^2/8$$

$$V(\bar{Y}) = [V(Y_1) + \cdots + V(Y_4)]/16 = 4\sigma^2/16 = 2\sigma^2/8 < V(T)\text{: } T \text{ não é mais eficiente que } \bar{Y}.$$

6 Sejam $X \sim \mathcal{N}(\mu, \sigma^2)$ e $(X_1, \dots, X_{16})$ uma amostra casual desta população, de que se obteve $\bar{x} = 10$ e $s = 3{,}872$.

a) Construa um intervalo de confiança a 95% para $V(X)$.

Res.

Variável fulcral $\qquad 16S^2/\sigma^2 \sim \chi^2_{15}$

$$\Pr\left(\underbrace{\xi_{0,025}}_{6,262} < 16S^2/\sigma^2 < \underbrace{\xi_{0,975}}_{27,488}\right) = 0{,}95 \quad \rightarrow$$

$$IC =]16 \times 3{,}872^2/27{,}488\ ;\ 16 \times 3{,}872^2/6{,}262[\approx]8{,}727\ ;\ 38{,}307[.$$

b) Suponha agora que a população tem variância 36. Qual a dimensão mínima da amostra, de modo que a amplitude de um intervalo de confiança a 95% para $E(X)$ não exceda 6,5?

Res.

Variável fulcral: $(\bar{X} - \mu)/\left(6/\sqrt{n}\right) \sim \mathcal{N}(0{,}1)$ $\qquad$ Extremos do IC para μ: $10 \pm \underbrace{\xi_{0,975}}_{1,96} \times 6/\sqrt{n}$

Amplitude do IC $\qquad 2 \times 1{,}96 \times 6/\sqrt{n} \leq 6{,}5 \Leftrightarrow n \geq 13{,}093 \Rightarrow n \geq 14.$

7 Indique quais das seguintes condições são hipóteses estatísticas (não é necessário justificar).

A $E(X) = 5$.

B $\bar{x} = 5$.

C $\Pr(X > 2) = 0{,}2$.

D $4 < \sigma^2 < 9$.

E $s' < 1$.

Res.

A C D.

8 Seja $X \sim \mathcal{N}(\mu, 2500)$ e o teste da hipótese $H_0: \mu = 250$ contra $H_1: \mu = 200$, cujo critério de rejeição se define como $\bar{x} < 230$. Calcule, para uma amostra de dimensão 16, o nível de significância e a potência deste teste.

Res.

Nível de significância

$$\alpha = \Pr(\bar{X} < 230 | \mu = 250) =$$
$$\Pr\left[(\bar{X} - 250)/\left(50/\sqrt{16}\right) < (230 - 250)/\left(50/\sqrt{16}\right) | \mu = 250\right] =$$
$$\Phi(-1{,}6) = 1 - 0{,}9452 = 0{,}0548.$$

Potência

$$1 - \beta = \Pr(\bar{X} < 230 | \mu = 200) =$$
$$\Pr\left[(\bar{X} - 200)/\left(50/\sqrt{16}\right) < (230 - 200)/\left(50/\sqrt{16}\right) | \mu = 200\right] = \Phi(2{,}4) = 0{,}9918.$$