Estatística II, 2015/16 Exame, época de recurso 27/6/2016

Resolução (de uma das versão do exame – as resoluções das diferentes versões são semelhantes, com alteração dos dados em cada questão).

1 O tempo de vida, em anos, de certo artigo electrónico é uma variável aleatória (X), com função densidade $f_X(x) = (18x - 3x^2)/80, \quad 0 < x < 4$.

a) Determine a expressão da distribuição de X.

Resolução

$$F_X(x) = \int_{-\infty}^{x} f_X(u)du = \begin{cases} 0, & x < 0 \\ (9x^2 - x^3)/80\,, & 0 \le x < 4 \\ 1, & x \ge 4 \end{cases}$$

b) A marca oferece a seguinte garantia: se o artigo se avaria durante o primeiro ano, é substituído por um novo. Qual a probabilidade de um artigo ser substituído em garantia?

Res.

$$F_X(1) = (9 \times 1^2 - 1^3)/80 = 1/10$$

c) O custo de reparação por avaria, Y, relaciona-se com X através de $Y = \begin{cases} 0, & X \le 1 \\ 10X, & X > 1 \end{cases}$.

i. Calcule a probabilidade de o custo da reparação de um artigo ser inferior a 20 u.m..

Res.

Dada a relação entre Y e X, a condição $Y < 20$ equivale a $Y = 0 \vee 10 < Y < 20$. Donde,

$$\Pr(Y < 20) = \Pr(Y = 0) + \Pr(10 < Y < 20) = \Pr(X \le 1) + \Pr(10 < 10X < 20) =$$
$$\Pr(X \le 1) + \Pr(1 < X < 2) = F_X(1) + F_X(2) - F_X(1) = (9 \times 2^2 - 2^3)/80 = 7/20.$$

ii. Calcule a probabilidade de o custo da reparação ser inferior a 20 u.m., dado que o artigo já não está em garantia.

Res.

$$\Pr(Y < 20|X > 1) = \Pr(10 < 10X < 20)/\Pr(X > 1) = \Pr(1 < X < 2)/\Pr(X > 1) =$$
$$[F_X(2) - F_X(1)]/[1 - F_X(1)] = (7/20 - 1/10)/(1 - 1/10) = 5/18.$$

2 A função geradora de momentos da v.a. X é dada por $M(s) = (2e^s + e^{2s} + 2e^{3s})/5$. Calcule a média e variância de X.

Res.

$$E(X) = M'(0) = (2e^0 + e^{2\times 0} \times 2 + 2e^{3\times 0} \times 3)/5 = 7/5,$$
$$V(X) = E(X^2) - E(X)^2 = M''(0) - [M'(0)]^2 =$$
$$(2e^0 + e^{2\times 0} \times 4 + 2e^{3\times 0} \times 9)/5 - 49/25 = 71/25.$$

3 As vendas diárias de areia (em centenas de quilogramas) de um armazém de materiais de construção é uma v.a. normal com média 20 e desvio-padrão 2.

a) Sabendo que numa manhã já se vendeu uma tonelada de areia, qual a probabilidade de que, nesse dia, se venda mais de 2,5 toneladas?

Res.

$$X \sim \mathcal{N}(20;4) \Leftrightarrow Z = (X-20)/2 \sim \mathcal{N}(0;1)$$
$$\Pr(X > 25 | X > 10) = \Pr(X > 25)/\Pr(X > 10) =$$
$$\Pr[Z > (25-20)/2]/\Pr[Z > (10-20)/2] = [1-\Phi(2{,}5)]/[1-\Phi(-5)] \approx$$
$$(1-0{,}9938)/1 = 0{,}0062.$$

b) Qual a probabilidade de que em certo mês (20 dias úteis) as vendas ultrapassem 37 toneladas?

Res.

Seja o volume de vendas mensal,

$$Y = \sum_{i=1}^{20} X_i \sim \mathcal{N}(400;80),$$

em que $X_i \sim \mathcal{N}(20;4)$ denota o volume de vendas no dia i; seja $Z = (Y-400)/\sqrt{80}$.

$$\Pr(Y > 370) = \Pr\left(Z > -30/\sqrt{80}\right) \approx 1-\Phi(-3{,}35) \approx 0{,}9996.$$

(Nota: deve ter atenção às unidades de medida da variável.)

4 Considere uma amostra casual de dimensão n, de uma população $\mathcal{N}(0,\theta)$ (θ: variância).

a) Determine o estimador de máxima verosimilhança (EMV) de θ.

Res.

$$X \sim \mathcal{N}(0;\theta) \Leftrightarrow f_X(x) = \frac{1}{\sqrt{2\pi\theta}} \exp\left(-\frac{x^2}{2\theta}\right) \Leftrightarrow$$
$$\log f_X(x) = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log\theta - \frac{1}{2\theta}x^2$$

Função log-verosimilhança

$$l(\mu) = \sum_{i=1}^{n} \log f_X(x_i) = \sum_{i=1}^{n}\left[-\frac{1}{2}\log(2\pi) - \frac{1}{2}\log\theta - \frac{1}{2\theta}x_i^2\right] =$$
$$-\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\theta - \frac{1}{2\theta}\sum_{i=1}^{n} x_i^2$$

Condição de primeira ordem para obter o EMV

$$l'(\mu) = 0 \Leftrightarrow -\frac{n}{2\theta} + \frac{1}{2\theta^2}\sum_{i=1}^{n} x_i^2 = 0 \Leftrightarrow \frac{1}{\theta}\sum_{i=1}^{n} x_i^2 = n \Leftrightarrow \theta = \sum_{i=1}^{n} x_i^2/n.$$

EMV de θ:

$$\hat{\theta}_{MV} = \sum_{i=1}^{n} X_i^2/n.$$

b) Seja Q o quantil de ordem 0,95 da população. Escreva a expressão de Q como função de θ. Obtenha a expressão do EMV de Q.

Res.

$$Q\colon \Pr(X < Q) = 0{,}95 \Leftrightarrow \Pr\left[X/\sqrt{\theta} < Q/\sqrt{\theta}\right] = 0{,}95 \Leftrightarrow \Phi\left(Q/\sqrt{\theta}\right) = 0{,}95 \Leftrightarrow$$
$$Q/\sqrt{\theta} = 1{,}645 \Leftrightarrow Q = 1{,}645\sqrt{\theta};$$

Estimador MV de Q (dada a propriedade de invariância do estimador MV):

$$\hat{Q}_{MV} = 1{,}645\sqrt{\hat{\theta}_{MV}},$$

em que $\hat{\theta}_{MV}$ é dado na alínea anterior.

5 Admite-se que o número de sinistros por apólice de determinada carteira do ramo automóvel segue um processo de Poisson com média λ. Recolhida uma amostra casual de dimensão 100, observa-se $\bar{x} = 0{,}07$. Construa um intervalo de confiança (IC) a 95% para a média do número de acidentes por apólice. O IC é exacto? Porquê?
(Sugestão: recorde que a média e variância da distribuição Poisson são iguais.)

Res.

Variável número de sinistros: $X \sim \text{Po}(\lambda)$; variável fulcral: $Z = (\bar{X} - \lambda)/\sqrt{\bar{X}/n} \mathrel{\dot\sim} \mathcal{N}(0{,}1)$.
Donde, o IC para λ,

$$\left] \ \bar{x} - z_{\alpha/2}\sqrt{\bar{x}/n} \quad ; \quad \bar{x} + z_{\alpha/2}\sqrt{\bar{x}/n} \ \right[;$$

dado que $1 - \alpha = 0{,}95$, $z_{\alpha/2} = 1{,}96$, logo, o intervalo é $]0{,}018\ ;\ 0{,}122[$.
O IC não é exacto, porque a distribuição da variável fulcral é apenas assintótica (não é a distribuição exacta).

6 Recolhida uma amostra casual de 30 observações de uma população $\mathcal{N}(\mu, \sigma^2)$, obteve-se $\bar{x} = 1$ e $\sum_{i=1}^{30} x_i^2 = 320$. Teste, ao nível de 5%, a hipótese $H_0\colon \sigma^2 = 12$ contra a alternativa $H_1\colon \sigma^2 > 12$.

Res.

$H_0\colon \sigma^2 = 12$; $H_1\colon \sigma^2 > 12$. Teste unilateral direito $\qquad \alpha = 5\%$
Inferência a respeito de σ^2; estatística de teste: $(30-1){S'}^2/12 \sim \chi^2_{(30-1)}$, sob $H_0\colon \sigma^2 = 12$.
Rejeita-se H_0 ao nível de 5%, se
$29{s'}^2/12 > 42{,}56 \ \longleftarrow$ percentil 0,95 da distribuição $\chi^2(29)$ ou, de modo equivalente,

$${s'}^2 > 12 \times 42{,}56/29 = 17{,}611.$$

Obteve-se $\quad {s'}^2 = (320 - 30 \times 1^2)/29 = 10, \quad$ logo, aceita-se H_0.
Note-se que

$$\sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum_{i=1}^{n} (x_i^2 - 2\bar{x}x_i + \bar{x}^2) =$$

$$\sum_{i=1}^{n} x_i^2 - 2\bar{x}\sum_{i=1}^{n} x_i + n\bar{x}^2 = \sum_{i=1}^{n} x_i^2 - 2n\bar{x}^2 + n\bar{x}^2 = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2.$$