Ficha de Expectativa de Resposta - Edital 026/2019

Departamento de Estatística - UFRN

21/10/2020

Questões Discursivas

Questão 1: 2,0

Questão 2: 2,0

Questão 3: 2,0

Questão 4: 2,0

Questão 5: 2,0

Questão 1

Seja $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ o modelo de regressão linear simples. Encontre os estimadores $\widehat{\beta}_0$ e $\widehat{\beta}_1$ pelo método de mínimos quadrados.

Resposta

(0,66 ponto para definir os erros e sua soma de quadrados) Vamos encontrar os valores de β_0 e β_1 que minimizam a soma dos erros, dados por

$$\varepsilon_i = y_i - (\beta_0 + \beta_1 x_i),\ i = 1, \cdots, n$$

Obtemos então a soma dos quadrados dos erros:

$$\begin{aligned} \mathrm{SQ}(\beta_0, \beta_1) &= \sum_{i=1}^{n} \varepsilon_i^2 \\ &= \sum_{i=1}^{n} \{y_i - (\beta_0 + \beta_1 x_i)\}^2 \end{aligned}$$

(0,67 ponto para determinar que os estimadores são encontrados a partir da minimização de SQ) A solução de mínimos quadrados é a escolha de β_0 e β_1 que torna esta soma mínima. O problema se resume, então, a encontrar o mínimo de uma função de duas variáveis β_0 e β_1. Derivando $\mathrm{SQ}(\beta_0, \beta_1)$ em relação a β_0 e β_1 e igualando a zero, observamos que as soluções $\widehat{\beta}_0$ e $\widehat{\beta}_1$ devem satisfazer

$$\begin{aligned} n\widehat{\beta}_0 + \widehat{\beta}_1 \sum_{i=1}^{n} x_i &= \sum_{i=1}^{n} y_i \\ \widehat{\beta}_0 \sum_{i=1}^{n} x_i + \widehat{\beta}_1 \sum_{i=1}^{n} x_i^2 &= \sum_{i=1}^{n} x_i y_i \end{aligned}$$

(0,67 ponto para encontrar os estimadores e mostrar que são pontos de mínimo) Dessa forma,

$$
\begin{aligned}
\widehat{\beta}_0 &= \overline{y} - \widehat{\beta}_1 \overline{x} \\
\widehat{\beta}_1 &= \frac{\sum_{i=1}^{n} x_i y_i - n\overline{xy}}{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2},
\end{aligned}
$$

em que

$$
\begin{aligned}
\overline{x} &= \frac{1}{n}\sum_{i=1}^{n} x_i \\
\overline{y} &= \frac{1}{n}\sum_{i=1}^{n} y_i.
\end{aligned}
$$

Questão 2

Seja $X_1, X_2, \cdots, X_n$ uma amostra i.i.d. tal que $X_i \sim \text{Exponencial}(\lambda)$. Se $f(x) = \frac{1}{\lambda} e^{-x_i/\lambda}$, $x_i > 0$, encontre o estimador de máxima verossimilhança de λ.

Resposta

(0,5 pela definição da função de verossimilhança) A função de verossimilhança $L(\lambda|x)$ é dada por

$$
L(\lambda|x) = \prod_{i=1}^{n} \frac{1}{\lambda} e^{-x_i/\lambda}.
$$

(0,5 pela definição da função de log-verossimilhança) Assim, a log-verossimilhança é dada por

$$
\begin{aligned}
\ln L(\lambda|x) &= \left(\prod_{i=1}^{n} \frac{1}{\lambda} e^{-x_i/\lambda} \right) \\
&= \sum_{i=1}^{n} (-x_i/\lambda) - n\ln(\lambda).
\end{aligned}
$$

(0,5 pela diferenciação em relação a λ) Diferenciando com respeito a λ e igualando a zero, temos

$$
\frac{\partial \ln L(\lambda|x)}{\partial \lambda} = \sum_{i=1}^{n} \left(x_i/\lambda^2 \right) - \frac{n}{\lambda} = 0
$$

(0,5 pela conclusão, explicitando a fórmula de $\widehat{\lambda}$ e mostrando que é um ponto de máximo) Portanto,

$$
\widehat{\lambda} = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}.
$$

Tomando a segunda derivada de $L(\lambda|x)$, é possível mostrar que $\overline{X}$ é um ponto de máximo dessa função.

Questão 3

Sejam A e B eventos em um espaço amostral Ω. Mostre que $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Resposta

(0,66 por reescrever $A \cup B$) Podemos escrever $A \cup B = A \cup (B - A)$. Além disso, temos que $A \cap (B - A) = \emptyset$. Portanto,

$$P(A \cup B) = P(A) + P(B - A). \tag{1}$$

(0,67 por reescrever B) Também temos que $B = (B - A) \cup (A \cap B)$ com $(B - A) \cap (A \cap B) = \emptyset$. Assim,

$$P(B) = P(B - A) + P(A \cap B) \tag{2}$$

(0,67 ao combinar (1) e (2)) Combinando (1) e (2), temos que

$$\begin{aligned} P(A \cup B) &= P(A) + P(B - A) \\ &= P(A) + P(B) - P(A \cap B). \end{aligned}$$

Questão 4

Qual ou quais eventos em um espaço amostral Ω são independentes de si mesmos? Justifique rigorosamente a resposta.

Resposta

(0,5 ponto) Apresentar a definição de eventos independentes:

Sejam A e B eventos do espaço amostral Ω, os eventos A e B são independentes se $P(A \cap B) = P(A) \cdot P(B)$.

(0,5 ponto para a identificação de cada um dos dois eventos e 0,5 para a justificativa)

Identificar os eventos que são independentes de si mesmos

Seja A um evento qualquer do espaço amostral Ω, para que o evento A seja independente dele mesmo, a seguinte condição deve ser satisfeita: $P(A \cap A) = P(A) \cdot P(A)$

$P(A) = P(A) \cdot P(A)$

Como $P(A)$ está no intervalo $[0, 1]$, os únicos números que são iguais ao seu quadrado são 0 e 1. Assim, os eventos com probabilidade zero ou um são os eventos que são independentes de si mesmos.

Questão 5

Considere uma amostra aleatória $(X_1, \ldots, X_n)$ de $(X_i|\theta) \sim Unif(0, \theta)$, $i = 1, \ldots, n$, e θ desconhecido.

i. Qual a distribuição exata de $X_{(n)} = max\{X_1, \ldots, X_n\}$?
ii. Mostre que $X_{(n)}$ é um estimador consistente de θ.
iii. Encontre um estimador de θ pelo método dos momentos.
iv. O estimador de θ obtido pelo método dos momentos é consistente? Justifique.

Resposta

i) *(0,33 ponto)* Considere $Y = X_{(n)}$. Para $\theta > 0$, a f.d.a. de $Y|\theta$ é dada por:

$$F_Y(y|\theta) = P[Y \leq y|\theta] = P\left[\max\left\{X_1, \ldots, X_n\right\} \leq y|\theta\right] = \prod_{i=1}^{n} P\left[X_i \leq y|\theta\right] = \begin{cases} 0, & y \leq 0 \\ \frac{y^n}{\theta^n}, & 0 < y < \theta \\ 1, & y \geq \theta \end{cases}$$

Logo, a densidade de $Y|\theta$ é dada por:

$$f_Y(y|\theta) = \frac{d}{dy} F_Y(y|\theta) = \frac{ny^{n-1}}{\theta^n} 1_{(0,\theta)}(y)$$

ii) *(0,67 ponto)* Estudaremos a consistência analisando o comportamento assintótico de $\mathrm{E}[Y|\theta]$ e $\mathrm{Var}[Y|\theta]$, uma vez que a convergência em média quadrática implica na convergência em probabilidade. A esperança, a média quadrática e a variância de $Y|\theta$ são respectivamente dadas por:

$$\mathrm{E}[Y|\theta] = \int_{-\infty}^{\infty} y \cdot f_Y(y|\theta) dy = \frac{n}{\theta^n} \int_0^\theta y^n dy = \frac{n}{n+1} \cdot \frac{1}{\theta^n} \cdot y^{n+1} \bigg|_0^\theta = \frac{n}{n+1} \cdot \theta$$
$$\mathrm{E}\left[Y^2 \mid \theta\right] = \int_{-\infty}^{\infty} y^2 \cdot f_Y(y|\theta) dy = \frac{n}{\theta^n} \int_0^\theta y^{n+1} dy = \frac{n}{n+2} \cdot \frac{1}{\theta^n} \cdot y^{n+2} \bigg|_0^\theta = \frac{n}{n+2} \cdot \theta^2$$

e

$$\mathrm{Var}[Y|\theta] = \mathrm{E}\left[Y^2 \mid \theta\right] - \{\mathrm{E}[Y|\theta]\}^2 = \frac{n}{n+2} \cdot \theta^2 - \left\{\frac{n}{n+1} \cdot \theta\right\}^2 = \left\{\frac{n}{n+2} - \left(\frac{n}{n+1}\right)^2\right\} \cdot \theta^2$$

Como

$$\lim_{n\to\infty} \mathrm{E}[Y|\theta] = \theta \quad \text{e} \quad \lim_{n\to\infty} \mathrm{Var}[Y|\theta] = 0$$

concluímos que $Y = X_{(n)}$ é um estimador consistente de θ.

Alternativamente, recorde-se que a convergência em distribuição para uma constante implica na convergência em probabilidade para a mesma constante. Assim, outra maneira de demonstrar que $Y = X_{(n)}$ é um estimador consistente de θ é notando o seguinte:

$$0 < y < \theta \Rightarrow 0 < \frac{y}{\theta} < 1 \Rightarrow \lim_{n\to\infty} \left(\frac{y}{\theta}\right)^n = 0$$

Consequentemente,

$$\lim_{n\to\infty} F_Y(y|\theta) = \begin{cases} 0, & y < \theta \\ 1, & y \geq \theta \end{cases} \quad \Rightarrow \quad \lim_{n\to\infty} P[Y = \theta \mid \theta] = 1$$

(0,33 ponto) iii) O estimador de θ dado pelo método dos momentos é dado pela solução

$$\widehat{\mathrm{E}\left[X_i \mid \theta\right]} = \overline{X} \Rightarrow \frac{\widehat{\theta}}{2} = \overline{X} \Rightarrow \widehat{\theta} = 2\overline{X} = \frac{2}{n} \sum_{i=1}^{n} X_i$$

(0,67 ponto) iv) Estudaremos a consistência novamente com base no comportamento assintótico de $\mathrm{E}[\hat{\theta} \mid \theta]$e $\mathrm{Var}[\hat{\theta} \mid \theta]$. Note que não é necessário conhecer a densidade de $\hat{\theta} \mid \theta$ para obter ambas as estatísticas, pois $\hat{\theta} = 2\overline{X}$ é uma combinação linear de n v.a.'s condicionalmente i.i.d. dado θ. Aplicando propriedades de esperança e variância condicionais, tem-se:

$$\mathrm{E}[\widehat{\theta} \mid \theta] = \frac{2}{n} \sum_{i=1}^{n} \mathrm{E}\left[X_i \mid \theta\right] = \frac{2}{n} \cdot n \cdot \frac{\theta}{2} = \theta$$

e

$$\mathrm{Var}[\hat{\theta} \mid \theta] = \frac{4}{n^2} \sum_{i=1}^{n} \mathrm{Var}\left[X_i \mid \theta\right] = \frac{4}{n^2} \cdot n \cdot \frac{\theta^2}{12} = \frac{\theta^2}{3n}$$

Como

$$\lim_{n \to \infty} \mathrm{E}[\hat{\theta} \mid \theta] = \theta \quad \text{e} \quad \lim_{n \to \infty} \mathrm{Var}[\hat{\theta} \mid \theta] = 0$$

concluímos que $\widehat{\theta} = 2\overline{X}$ é um estimador consistente de θ.

Alternativamente, observando-se que $\widehat{\theta}$ é um estimador centrado em θ (*i.e.*, $\mathrm{E}[\hat{\theta} \mid \theta] = \theta$), podemos verificar a consistência ou convergência em probabilidade para θ por meio da desigualdade de Bienaymé-Chebyshev. De antemão, note o seguinte:

$$\mathrm{E}\left[\{\widehat{\theta} - \theta\}^2 \mid \theta\right] = \mathrm{E}\left[\{\hat{\theta} - \mathrm{E}[\hat{\theta} \mid \theta]\}^2 \mid \theta\right] = \mathrm{Var}[\hat{\theta} \mid \theta]$$

Então, para $\epsilon > 0$, tem-se:

$$P[|\widehat{\theta} - \theta| \geq \epsilon \mid \theta] = P\left[\{\widehat{\theta} - \theta\}^2 \geq \epsilon^2 \mid \theta\right] \leq \frac{\mathrm{E}\left[\{\widehat{\theta} - \theta\}^2 \mid \theta\right]}{\epsilon^2} = \frac{\mathrm{Var}[\hat{\theta} \mid \theta]}{\epsilon^2} = \frac{\theta^2}{3n\epsilon^2} \to 0$$

MARCUS ALEXANDRE NUNES PRESIDENTE	CARLA ALMEIDA VIVACQUA 1º EXAMINADOR	PAULO CANAS RODRIGUES 2º EXAMINADOR