

# Probabilidade e Estatística

## Series de Problemas 4

08 de Maio de 2025

1. Admita que o número de programas examinados de modo independente até que se observe o primeiro programa que não compile é representado pela variável aleatória  $X$  com distribuição geométrica com parâmetro  $p$ , onde  $p$  é uma probabilidade desconhecida. Determine a estimativa de máxima verosimilhança de  $p$ , atendendo à amostra  $(4, 50, 5, 13, 30)$  proveniente da população  $X$ .

### Resolução:

Considere-se a variável aleatória  $X \sim \text{Geom}(p)$  que representa o número de programas examinados até observar o primeiro que não compile, incluindo este. A função de probabilidade da distribuição geométrica é:

$$P(X = x) = (1 - p)^{x-1}p, \quad x = 1, 2, 3, \dots$$

Para a amostra  $(x_1, x_2, x_3, x_4, x_5) = (4, 50, 5, 13, 30)$ , a função de verosimilhança é:

$$\begin{aligned} L(p) &= \prod_{i=1}^5 P(X = x_i) \\ &= \prod_{i=1}^5 (1 - p)^{x_i-1}p \\ &= p^5 \times (1 - p)^{\sum_{i=1}^5 (x_i-1)} \\ &= p^5 \times (1 - p)^{97} \end{aligned}$$

Para maximizar a verosimilhança, aplique-se o logaritmo natural:

$$\begin{aligned} \ell(p) &= \log L(p) \\ &= \log (p^5 \times (1 - p)^{97}) \\ &= 5 \log p + 97 \log(1 - p) \end{aligned}$$

Derive-se  $\ell(p)$  em relação a  $p$ :

$$\begin{aligned}\frac{d\ell}{dp} &= \frac{d}{dp} (5 \log p + 97 \log(1 - p)) \\ &= \frac{5}{p} + 97 \cdot \frac{1}{1 - p} \cdot (-1) \\ &= \frac{5}{p} - \frac{97}{1 - p}\end{aligned}$$

Igualando a derivada a zero para encontrar o valor de  $p$  que maximiza  $\ell(p)$ :

$$\begin{aligned}\frac{5}{p} - \frac{97}{1 - p} &= 0 \\ \frac{5}{p} &= \frac{97}{1 - p} \\ 5(1 - p) &= 97p \\ 5 - 5p &= 97p \\ 5 &= 102p \\ p &= \frac{5}{102} \approx 0.0490196\end{aligned}$$

**Resposta:** A estimativa de máxima verosimilhança de  $p$  é  $\frac{5}{102} \approx 0.0490196$ .

2. Tem-se assumido que o impacto hidrodinâmico (com valores medidos em unidades apropriadas),  $X$  do casco de um navio sobre uma onda em determinada região do globo possui função de densidade de probabilidade dada por

$$f(x) = \left| \frac{2x}{\lambda^2} \right| e^{-\frac{x^2}{\lambda^2}},$$

onde  $x > 0$  e o parâmetro  $\lambda$  é uma constante positiva desconhecida. Tendo por base uma amostra aleatória  $(X_1, X_2, \dots, X_{13})$  de  $X$  com  $n = 13$ , determine a estimativa de máxima verosimilhança do parâmetro  $\lambda$  para uma realização da amostra tal que  $\sum_{i=1}^{13} x_i^2 = 62$ .

**Resolução:**

Comece-se por observar que a função de densidade fornecida pode ser reescrita, já que  $x > 0$ , como:

$$f(x) = \frac{2x}{\lambda^2} e^{-\frac{x^2}{\lambda^2}}, \quad x > 0$$

Trata-se da distribuição de Rayleigh generalizada. A função de verossimilhança para uma amostra  $(x_1, x_2, \dots, x_n)$  é dada por:

$$\begin{aligned} L(\lambda) &= \prod_{i=1}^n \frac{2x_i}{\lambda^2} e^{-\frac{x_i^2}{\lambda^2}} \\ &= \left( \frac{2^n}{\lambda^{2n}} \prod_{i=1}^n x_i \right) \cdot e^{-\frac{1}{\lambda^2} \sum_{i=1}^n x_i^2} \end{aligned}$$

Passe-se à log-verossimilhança:

$$\begin{aligned} \ell(\lambda) &= \log L(\lambda) \\ &= n \log 2 - 2n \log \lambda + \sum_{i=1}^n \log x_i - \frac{1}{\lambda^2} \sum_{i=1}^n x_i^2 \end{aligned}$$

Derivando em relação a  $\lambda$  e igualando a zero:

$$\begin{aligned} \frac{d\ell}{d\lambda} &= -\frac{2n}{\lambda} + \frac{2}{\lambda^3} \sum_{i=1}^n x_i^2 = 0 \\ \Rightarrow -2n\lambda^2 + 2 \sum_{i=1}^n x_i^2 &= 0 \\ \Rightarrow \lambda^2 &= \frac{1}{n} \sum_{i=1}^n x_i^2 \\ \Rightarrow \hat{\lambda} &= \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2} \end{aligned}$$

Apliquemos agora à amostra fornecida:

$$\begin{aligned}n &= 13 \\ \sum_{i=1}^n x_i^2 &= 62 \\ \hat{\lambda} &= \sqrt{\frac{62}{13}} = \sqrt{4.7692} \approx 2.1839\end{aligned}$$

**Resposta:** A estimativa de máxima verosimilhança de  $\lambda$  é 2,18 (com duas casas decimais).

3. Considere a variável aleatória  $X \sim \text{Poi}(\lambda)$ , que modela o número de participações de sinistros automóveis a determinada seguradora num período de uma hora, e uma amostra aleatória  $(X_1, X_2, \dots, X_n)$  de  $X$ . Calcule a estimativa de máxima verosimilhança da probabilidade de ocorrerem mais de 3 participações de sinistros automóveis às seguradoras numa hora, sabendo que a concretização de uma amostra aleatória de dimensão 22 de  $X$  conduziu a

$$\sum_{i=1}^{22} x_i = 42.$$

- a) **0.126856**
- b) **0.873144**
- c) **0.002071**
- d) **0.983743**

**Resolução:**

Comece-se por recordar que a esperança da distribuição de Poisson é dada por  $E(X) = \lambda$ . Assim, a estimativa de máxima verosimilhança (EMV) de  $\lambda$ , com base na amostra de dimensão  $n = 22$  e soma dos valores observados igual a 42, é:

$$\hat{\lambda} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{42}{22} \approx 1.9091$$

Pretende-se estimar a probabilidade de ocorrerem mais de 3 sinistros numa hora, ou seja:

$$P(X > 3) = 1 - P(X \leq 3)$$

Com  $X \sim \text{Poi}(\hat{\lambda}) = \text{Poi}(1.9091)$ , utilizando a função `PoissonCD` da calculadora Casio fx-CG50, obtém-se:

$$P(X \leq 3) \approx 0.8731$$

Portanto:

$$P(X > 3) = 1 - 0.8731 = 0.1269$$

Arredondando com seis casas decimais, temos:

$$\boxed{0.126856}$$

**Resposta correta: a) 0.126856**

4. Admita que o tempo de vida em centenas de horas,  $X$ , de um novo tipo de lâmpadas de longa duração segue uma distribuição exponencial cujo parâmetro  $\lambda$  é uma constante desconhecida e positiva. Com o objetivo de estimar  $\lambda$ , registaram-se as durações de lâmpadas deste tipo, tendo-se obtido a seguinte amostra: (91, 81, 102, 72, 120). Calcule a estimativa de máxima verosimilhança da probabilidade de uma lâmpada desse tipo durar mais de 50 centenas de horas.

- a) **0.4922**
- b) **0.7504**
- c) **0.5848**
- d) **0.6758**

**Resolução:**

Começemos por lembrar que a função de densidade da distribuição

exponencial é dada por:

$$f(x; \lambda) = \lambda e^{-\lambda x}, \quad x > 0$$

A função de verosimilhança para uma amostra  $(x_1, x_2, \dots, x_n)$  é:

$$\begin{aligned} L(\lambda) &= \prod_{i=1}^n \lambda e^{-\lambda x_i} \\ &= \lambda^n e^{-\lambda \sum_{i=1}^n x_i} \end{aligned}$$

Passando à log-verosimilhança:

$$\begin{aligned} \ell(\lambda) &= \log L(\lambda) \\ &= n \log \lambda - \lambda \sum_{i=1}^n x_i \end{aligned}$$

Derivando em relação a  $\lambda$  e igualando a zero:

$$\begin{aligned} \frac{d\ell}{d\lambda} &= \frac{n}{\lambda} - \sum_{i=1}^n x_i = 0 \\ \Rightarrow \frac{n}{\lambda} &= \sum_{i=1}^n x_i \\ \Rightarrow \hat{\lambda} &= \frac{n}{\sum_{i=1}^n x_i} \end{aligned}$$

Note-se que esta é a expressão do estimador de máxima verosimilhança para o parâmetro  $\lambda$  da distribuição exponencial. Alternativamente, sabendo que  $E(X) = \frac{1}{\lambda}$ , temos  $\hat{\lambda} = \frac{1}{\bar{x}}$ .

Dada a amostra  $(91, 81, 102, 72, 120)$ , com  $n = 5$ , calcule-se:

$$\begin{aligned} \sum_{i=1}^n x_i &= 91 + 81 + 102 + 72 + 120 = 466 \\ \hat{\lambda} &= \frac{5}{466} \approx 0.0107296 \end{aligned}$$

Para uma variável aleatória exponencial, a probabilidade de ultrapassar um valor  $t$  é dada por:

$$P(X > t) = e^{-\lambda t}$$

Assim, a estimativa de máxima verosimilhança da probabilidade de uma lâmpada durar mais de 50 centenas de horas é:

$$\begin{aligned} P(X > 50) &= e^{-\hat{\lambda} \cdot 50} \\ &= e^{-0.0107296 \cdot 50} \\ &= e^{-0.53648} \\ &\approx 0.5848 \end{aligned}$$

**Resposta correta: c) 0.5848**

5. Admita que a proporção de potássio em certo fertilizante é representada pela variável aleatória  $X$  com função de densidade de probabilidade

$$f_X(x) = \begin{cases} \theta(1-x)^{\theta-1} & 0 < x < 1 \\ 0 & \text{caso contrário,} \end{cases}$$

onde  $\theta$  é um parâmetro positivo desconhecido. Determine a estimativa de máxima verosimilhança de  $P(X \leq 0,12)$  baseada na amostra  $(0.666, 0.117, 0.167, 0.57, 0.106, 0.095)$  proveniente da população  $X$ .

**Resolução:**

Observe-se que a função densidade corresponde a uma distribuição Beta(1,  $\theta$ ), também conhecida como distribuição de potência. A função de verosimilhança para a amostra  $(x_1, x_2, \dots, x_n)$  é dada por:

$$\begin{aligned} L(\theta) &= \prod_{i=1}^n \theta(1-x_i)^{\theta-1} \\ &= \theta^n \cdot \prod_{i=1}^n (1-x_i)^{\theta-1} \end{aligned}$$

Tomando o logaritmo para facilitar a maximização:

$$\begin{aligned}\ell(\theta) &= \log L(\theta) \\ &= n \log \theta + (\theta - 1) \sum_{i=1}^n \log(1 - x_i)\end{aligned}$$

Derivando em relação a  $\theta$  e igualando a zero:

$$\begin{aligned}\frac{d\ell}{d\theta} &= \frac{n}{\theta} + \sum_{i=1}^n \log(1 - x_i) = 0 \\ \Rightarrow \frac{n}{\theta} &= - \sum_{i=1}^n \log(1 - x_i) \\ \Rightarrow \hat{\theta} &= - \frac{n}{\sum_{i=1}^n \log(1 - x_i)}\end{aligned}$$

Com a amostra fornecida (0.666, 0.117, 0.167, 0.57, 0.106, 0.095), calcule-se:

$$\begin{aligned}\sum_{i=1}^n \log(1 - x_i) &= \log(1 - 0.666) + \log(1 - 0.117) + \log(1 - 0.167) \\ &\quad + \log(1 - 0.57) + \log(1 - 0.106) + \log(1 - 0.095) \\ &= \log(0.334) + \log(0.883) + \log(0.833) \\ &\quad + \log(0.430) + \log(0.894) + \log(0.905) \\ &\approx -1.0966 - 0.1244 - 0.1827 - 0.8440 - 0.1120 - 0.0998 \\ &= -2.4596\end{aligned}$$

Assim, a estimativa de máxima verosimilhança do parâmetro é:

$$\begin{aligned}\hat{\theta} &= - \frac{6}{-2.4596} \\ &\approx 2.4394\end{aligned}$$

Agora, para calcular  $P(X \leq 0.12)$  com o parâmetro estimado:

$$P(X \leq 0.12) = \int_0^{0.12} \hat{\theta}(1 - x)^{\hat{\theta}-1} dx$$



Utilizando a substituição  $u = 1 - x \Rightarrow du = -dx$ , mudando os limites:

$$\begin{aligned}x = 0 &\Rightarrow u = 1 \\x = 0.12 &\Rightarrow u = 0.88\end{aligned}$$

Com esta substituição, o integral torna-se:

$$\begin{aligned}P(X \leq 0.12) &= \int_{0.88}^1 \hat{\theta} u^{\hat{\theta}-1} (-du) \\&= \int_1^{0.88} -\hat{\theta} u^{\hat{\theta}-1} du \\&= \left[ -u^{\hat{\theta}} \right]_1^{0.88} \\&= -0.88^{\hat{\theta}} - (-1^{\hat{\theta}}) \\&= 1 - 0.88^{2.44} \\&\approx 1 - 0.7321 \\&= 0.2679\end{aligned}$$

**Resposta:** A estimativa de máxima verosimilhança de  $P(X \leq 0.12)$  é aproximadamente 0.2679.