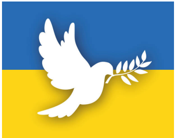
Estatística II

Aulas 13, 14 & 15

Nelson Faustino [1]

[1]**Faculdade de Economia (FEUC)**
Universidade de Coimbra – Portugal

nelson@fe.uc.pt



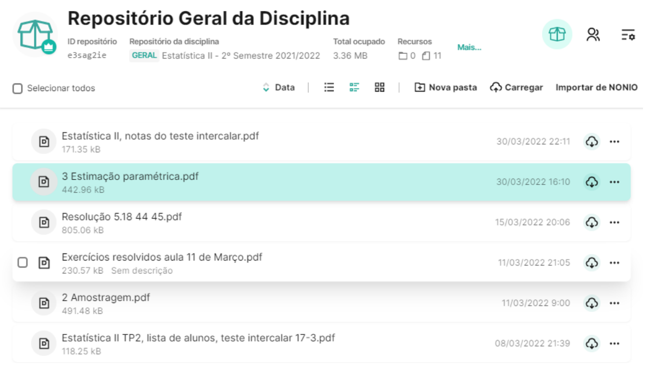
Conteúdos

Métodos de Estimação por Pontos

Leituras Recomendadas

Complemento dos Slides

- **Páginas 399 a 426** de [Murteira et al. (2015), Capítulo 7].
- **Páginas 1 a 15** do ficheiro `3 Estimação paramétrica.pdf`, disponível no **Repositório Geral da Disciplina (UC Student).**



Método dos Momentos

Vide [Murteira et al. (2015), pp. 402–403]

Aplicação do Método dos Momentos

A partir de uma amostra casual $(X_1, X_2, \ldots, X_n)$ de $X \sim F$:

- **Passo 1:** Identificar o número de parâmetros a estimar ($\theta_1, \theta_2, \ldots, \theta_m -$desconhecidos) e calcular os momentos de ordem k, $E(X^k)$ $(k = 1, 2, \ldots, m)$;
- **Passo 2:** Resolver sistema de m equações a m incógnitas em ordem a $\theta_1, \theta_2, \ldots, \theta_m$:

$$\left\{\begin{array}{l} E(X) = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} x_i \\ \ldots \\ E(X^m) = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} x_i^m \end{array}\right. \Leftrightarrow \left\{\begin{array}{l} \underbrace{\theta_1 = \phi_1(x_1, x_2, \ldots, x_n)}_{\text{estimador } \widehat{\theta}_1 = \phi_1(X_1, X_2, \ldots, X_n)} \\ \ldots \\ \underbrace{\theta_m = \phi_m(x_1, x_2, \ldots, x_n)}_{\text{estimador } \widehat{\theta}_m = \phi_m(X_1, X_2, \ldots, X_n)} \end{array}\right.$$

Alguns exemplos de estimadores que iremos considerar a partir de uma amostra casual $(X_1, X_2, \ldots, X_n)$ de $X \sim F$.

- **Média da amostra:** $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$
  - **Estimativa:** $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$;
  - **eq. método dos momentos:** $E(X) = \overline{x}$.
- **Variância da amostra:** $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$
  - **Estimativa:** $s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{x}^2$;
  - **eq. método dos momentos:** $V(X) = s^2$.

Métodos dos Momentos

Um (1) parâmetro a estimar

$E(X) = \overline{x}$

- **Bernoulli** $X \sim B(1,\theta)$
  - $E(X) = \theta$
  - $\theta = \overline{x}$
  - $\widehat{\theta} = \overline{X}$
- **Geométrica** $X \sim BN(1,\theta)$
  - $E(X) = \frac{1}{\theta}$
  - $\theta = \frac{1}{\overline{x}}$
  - $\widehat{\theta} = \frac{1}{\overline{X}}$
- **Poisson** $X \sim Po(\lambda)$
  - $E(X) = \lambda$
  - $\lambda = \overline{x}$
  - $\widehat{\lambda} = \overline{X}$
- **Exponencial** $X \sim Ex(\lambda)$
  - $E(X) = \frac{1}{\lambda}$
  - $\lambda = \frac{1}{\overline{x}}$
  - $\widehat{\lambda} = \frac{1}{\overline{X}}$

Método dos Momentos

Duas (2) equações a duas (2) incógnitas

Estimativa para a Média da Amostra ($\overline{X}$)

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Estimativa para Variância da Amostra (S^2)

$$s^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{x}^2$$

Equivalência do Sistema de Equações

$$\begin{cases} E(X) = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} x_i \\[2ex] E(X^2) = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} x_i^2 \end{cases} \Leftrightarrow \begin{cases} E(X) = \overline{x} \\[2ex] E(X^2) - E(X)^2 = s^2 \end{cases} \Leftrightarrow \begin{cases} E(X) = \overline{x} \\[2ex] V(X) = s^2 \end{cases}$$

Obs: $V(X) = E(X^2) - E(X)^2$.

Métodos dos Momentos

Dois (2) parâmetros a estimar

Distribuição Normal $N(\mu, \sigma^2)$

$$\begin{cases} E(X) = \overline{x} \\ V(X) = s^2 \end{cases} \Longleftrightarrow \begin{cases} \mu = \overline{x} \\ \sigma^2 = s^2 \end{cases} \Longrightarrow \begin{cases} \widehat{\mu} = \overline{X} \\ \widehat{\sigma^2} = S^2 \end{cases}$$

Distribuição Gama $G(\alpha, \lambda)$

$$\begin{cases} E(X) = \overline{x} \\ V(X) = s^2 \end{cases} \Leftrightarrow \begin{cases} \dfrac{\alpha}{\lambda} = \overline{x} \\ \dfrac{\alpha}{\lambda^2} = s^2 \end{cases} \Leftrightarrow \begin{cases} \alpha = \dfrac{\overline{x}^2}{s^2} \\ \lambda = \dfrac{\overline{x}}{s^2} \end{cases} \Longrightarrow \begin{cases} \widehat{\alpha} = \dfrac{\overline{X}^2}{S^2} \\ \widehat{\lambda} = \dfrac{\overline{X}}{S^2} \end{cases}$$

Método da Máxima Verossimilhança

Vide [Murteira et al. (2015), pp. 403–412]

Aplicação do Método da Máxima Verossimilhança

A partir de uma amostra casual $(X_1, X_2, \ldots, X_n)$ de $X \sim F$:

- **Passo 1:** Identificar o número de parâmetros a estimar ($\theta_1, \theta_2, \ldots, \theta_m -$desconhecidos) e construir a função de máxima verossimilhança $L(\theta_1, \theta_2, \ldots, \theta_m) = \prod_{i=1}^{n} f(x_i|\theta_1, \theta_2, \ldots, \theta_m)$.
- **Passo 2:** Maximizar a função de log-verossimilhança $\ell(\theta_1, \theta_2, \ldots, \theta_m) = \ln L(\theta_1, \theta_2, \ldots, \theta_m)$:
  - **Condições de primeira ordem:** $\theta_1, \theta_2, \ldots, \theta_m$ pontos estacionários de $\ell(\theta_1, \theta_2, \ldots, \theta_m)$ (i.e. $\frac{\partial \ell}{\partial \theta_i} = 0, \ i = 1, 2, \ldots, m$);
  - **Condições de segunda ordem:** $\widehat{\theta}_j = \phi_j(X_1, X_2, \ldots, X_m)$ estimadores de máxima verossimilhança $(j = 1, 2, \ldots, m)$ se $(\theta_1, \theta_2, \ldots, \theta_m)$ é máximo local de $\ell(\theta_1, \theta_2, \ldots, \theta_m)$ **(2^a derivada/matriz Hessiana)**.

Função de Verossimilhança

$L(\lambda) = \prod_{i=1}^{n} f(x_i|\lambda)$ a partir de $f(x|\lambda) = \lambda e^{-\lambda x}$.

Cálculo $L(\lambda)$ para $X \sim Ex(\lambda)$

$$
\begin{aligned}
L(\lambda) &= \prod_{i=1}^{n} f(x_i|\lambda) \\
\text{[Multiplicação dos termos } e^{-\lambda x_i}\text{]} &= \prod_{i=1}^{n} \underbrace{\lambda}_{\text{constante}} \quad \underbrace{e^{-\lambda x_i}}_{\text{exponencial}} \\
\text{"dá-se a mesma base e} &= \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i} \\
\text{somam-se os expoentes."} &= \lambda^n e^{-n\lambda \overline{x}}.
\end{aligned}
$$

Maximização $\ell(\lambda)$ para $X \sim Ex(\lambda)$

$$L(\lambda) = \lambda^n e^{-n\lambda\overline{x}}$$

$$\overset{\ell(\lambda)=\ln L(\lambda)}{\Longrightarrow} \ell(\lambda) = n\ln(\lambda) - n\lambda\overline{x}$$

Os cálculos abaixo permitem-nos concluir que $\widehat{\lambda} = \frac{1}{\overline{X}}$ é o **estimador de máxima verossimilhança**:

- **Condições de primeira ordem:**
  $$\ell'(\lambda) = 0 \Longleftrightarrow \frac{n}{\lambda} - n\overline{x} = 0 \Longleftrightarrow \lambda = \frac{1}{\overline{x}}.$$
- **Condições de segunda ordem:**
  $$\ell''\left(\frac{1}{\overline{x}}\right) = \left[-\frac{n}{\lambda^2}\right]_{\lambda=\frac{1}{\overline{x}}} = -n\overline{x}^2 < 0.$$

Método da Máxima Verossimilhança

Distribuição Normal (cf. [Murteira et al. (2015), Exemplo 7.15])

Função de log-Verossimilhança

$$\ell(\mu,\sigma^2) = -\frac{n}{2}\ln(2\pi) - -\frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

Cálculo pontos estacionários (variáveis μ e σ^2)

$$\begin{cases} \dfrac{\partial \ell}{\partial \mu} = 0 \\[2ex] \dfrac{\partial \ell}{\partial \sigma^2} = 0 \end{cases} \Leftrightarrow \begin{cases} \dfrac{\sum_{i=1}^{n}(x_i-\mu)}{\sigma^2} = 0 \\[2ex] -\dfrac{n}{2\sigma^2} + \dfrac{\sum_{i=1}^{n}(x_i-\mu)^2}{2(\sigma^2)^2} = 0 \end{cases} \Leftrightarrow \begin{cases} \mu = \overline{x} \\[2ex] \sigma^2 = s^2 \end{cases}$$

Matriz Hessiana de $\ell(\mu, \sigma^2)$

$$\begin{aligned} H_\ell(\mu,\sigma^2) &= \begin{bmatrix} \dfrac{\partial^2 \ell}{\partial \mu^2} & \dfrac{\partial^2 \ell}{\partial \mu \partial \sigma^2} \\[2ex] \dfrac{\partial^2 \ell}{\partial \sigma^2 \partial \mu} & \dfrac{\partial^2 \ell}{\partial (\sigma^2)^2} \end{bmatrix} \\ &= \begin{bmatrix} -\dfrac{n}{\sigma^2} & -\dfrac{n\overline{x} - n\mu}{(\sigma^2)^2} \\[2ex] -\dfrac{n\overline{x} - n\mu}{(\sigma^2)^2} & \dfrac{n}{2(\sigma^2)^2} - \dfrac{\sum_{i=1}^n (x_i - \mu)^2}{(\sigma^2)^3} \end{bmatrix}. \end{aligned}$$

Matriz Hessiana de $\ell(\mu, \sigma^2)$ calculada para $\mu = \overline{x}$ e $\sigma^2 = s^2$

$$H_\ell(\overline{x}, s^2) = \begin{bmatrix} -\dfrac{n}{s^2} & 0 \\ 0 & \dfrac{n}{2(s^2)^2} - \dfrac{2ns^2}{(s^2)^3} \end{bmatrix} = \mathrm{diag}\left(-\dfrac{n}{s^2}, -\dfrac{n}{2s^2}\right).$$

Matriz Hessiana de $\ell(\mu, \sigma^2)$ calculada para $\mu = \overline{x}$ e $\sigma^2 = s^2$

Os cálculos do sinal dos menores da matriz Hessiana, $H_\ell(\overline{x}, s^2)$, permitem-nos concluir que $\widehat{\mu} = \overline{X}$ e $\widehat{\sigma^2} = S^2$ são os **estimadores de máxima verossimilhança**:

- **Menor Δ_1:** $\Delta_1 = -\frac{n}{s^2} < 0$.
- **Menor Δ_2:** $\Delta_2 = \det H_\ell(\overline{x}, s^2) > 0$.

Método dos Momentos vs. Método da Máxima Verossimilhança

Exemplos de [Murteira et al. (2015), pp. 407–409]

Exemplos

- **Exemplo 7.9 –** Para uma amostra casual de uma população de Bernoulli, é verificado que o estimador de máxima verossimilhança **coincide** com o estimador determinado pelo método dos momentos;
- **Exemplo 7.10 –** Dá-nos um exemplo claro de que os **estimadores** determinados por ambos os métodos não são necessariamente iguais;
- **Exemplo 7.12 –** Mostra-nos que a estimativa de máxima verossimilhança de $X_{(n)} = \max(X_i)$ de $X \sim U(0,\theta)$, $x_{(n)}$, não é um zero da da derivada da **função de log-verossimilhança** ($\ell'(x_{(n)}) \neq 0$ mas $\ell(x_{(n)})) > \ell(x_{(n)} + \delta)$, para todo o $\delta > 0$).

Propriedade da Invariância

$\tau(\widehat{\theta})$ é um **estimador de máxima verossimilhança** de $\tau(\theta)$, desde que:

1. $\widehat{\theta}$ seja um **estimador de máxima verossimilhança** de θ;
2. $\theta \mapsto \tau(\theta)$ seja uma **função biunívoca** – i.e. admite inversa.

Exemplos

A propriedade anterior permite-nos concluir que:

- **Populações $Ex(\lambda)$:** $\widehat{\tau} = \overline{x}$ estimador de máxima verossimilhança para a **média da população**, $\tau = E(X)$ (cf. [Murteira et al. (2015), Exemplo 7.14]);
- **Populações $N(\mu, \sigma^2)$:** $\widehat{\sigma} = S$ é um estimador de máxima verossimilhança para o **desvio padrão da população** (cf. [Murteira et al. (2015), Exemplo 7.15]).

Conteúdos

Estimadores Cêntricos/Enviesados

[Murteira et al. (2015), pp. 412–413]

Estimador Centrado ou Cêntrico

$T = T(X_1, X_2, \ldots, X_n)$ é um **estimador centrado/cêntrico** de θ se

$$E(T) = \theta.$$

Enviesamento de um Estimador

Se $T = T(X_1, X_2, \ldots, X_n)$ é um estimador de θ, a

$$Env(T) = E(T) - \theta$$

ou $B(T) = E(T) - \theta$ designamos por **enviesamento** de T.

Observações:

- Um estimador T diz-se **enviesado** sempre que $Env(T) \neq 0$;
- Iremos designar os **estimadores cêntricos/centrados** como sendo estimadores **não enviesados**, em virtude de $Env(T) = 0$.

.

Estimadores Cêntricos/Enviesados

Média e Variância de uma Amostra Casual

cf. [Murteira et al. (2015), Teorema 6.1 & Teorema 6.2]

Se $(X_1, X_2, \ldots, X_n)$ é uma **amostra casual** de uma população para a qual existem **média** $\mu = E(X_i)$ e **variância** $\sigma^2 = V(X_i)$ $(i = 1, 2, \ldots, n)$, tem-se:

- **Valor esperado da média da amostra:** $E(\overline{X}) = \mu$;
- **Valor esperado da variância da amostra:** $E(S^2) = \dfrac{n-1}{n}\sigma^2$.

Observações:

Os resultados acima permitem-nos concluir que:

1. $T = \overline{X}$ é um estimador cêntrico/centrado para a **média da população**;
2. $T = S^2$ não é um estimador cêntrico/centrado para a **variância da população**.

Variância Corrigida da Amostra

$$
\begin{aligned}
(S')^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2 \\
&= \frac{n}{n-1}\,\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2 \\
&= \frac{n}{n-1}S^2 \\
\Longrightarrow E((S')^2) &= \frac{n}{n-1}\overbrace{E(S^2)}^{=\frac{n-1}{n}\sigma^2} \\
&= \frac{\cancel{n}}{\cancel{n-1}}\,\frac{\cancel{n-1}}{\cancel{n}}\sigma^2 = \sigma^2
\end{aligned}
$$

Estimadores Cêntricos/Enviesados

Variância [Corrigida] de uma Amostra Casual

Observações:

1. S^2 – estimador T_2 de [Murteira et al. (2015), p. 451, ex. 18.]) – diz-se estimador enviesado, uma vez que $Env(S^2) = E(S^2) - \sigma^2 = \frac{n-1}{n}\sigma^2 - \sigma^2 = -\frac{\sigma^2}{n}$
2. Já $(S')^2$ – estimador T_3 de [Murteira et al. (2015), p. 451, ex. 18.]) – é um estimador cêntrico/centrado para a **variância da população**, uma vez que $Env((S')^2) = 0$.
3. Podemos ainda verificar que $T = S^2 + \overline{X}^2$ é um estimador **é centrado/não enviesado** para $\sigma^2 + \mu^2$ (verificar que coincide com T de [Murteira et al. (2015), p. 449, ex. 9.b)]). **Cálculos no próximo slide**.

Estimadores Cêntricos/Enviesados

Estimador $T = S^2 + \overline{X}^2$

Aplicação directa de [Murteira et al. (2015), Teorema 6.1 & Teorema 6.2]:

Estimador $T = S^2 + \overline{X}^2$

$$\begin{aligned} E(\overline{X}^2) &= V(\overline{X}) + E(\overline{X})^2 \\ &= \frac{V(X)}{n} + E(X)^2 \\ &= \frac{\sigma^2}{n} + \mu^2 \\ \Longrightarrow E(S^2 + \overline{X}^2) &= E(S^2) + E(\overline{X}^2) \\ &= \frac{n - \cancel{1}}{n}\sigma^2 + \cancel{\frac{\sigma^2}{n}} + \mu^2 \\ &= \sigma^2 + \mu^2. \end{aligned}$$

Eficiência de um Estimador

[Murteira et al. (2015), pp. 414–416]

Eficiência Relativa (cf. [Murteira et al. (2015), Definição 7.1])

Sejam T e T' dois **estimador centrados/cêntricos** de θ. Dizemos que T é mais eficiente que T' se

$$V(T) \leq V(T').$$

Eficiência Absoluta (cf. [Murteira et al. (2015), Teorema 7.1])

Seja $(X_1, X_2, \ldots, X_n)$ uma amostra casual de uma população com função densidade de probabilidade $f(x|\theta)$ ($\theta-$desconhecido). Então é sempre válida a desigualdade

$$V(T) \geq \frac{1}{n\mathcal{I}(\theta)},$$

onde $\mathcal{I}(\theta) = -E\left[\frac{\partial^2 \ln f(X|\theta)}{\partial \theta^2}\right]$ designa a **quantidade de informação de Fischer**.

Eficiência de um Estimador

Eficiência Absoluta vs. Eficiência Relativa

Estimador mais eficiente

Os resultados anteriores permitem-nos concluir que T é o **estimador mais eficiente** – i.e. o estimador com menor variância – se, e só se

$$V(T) = \frac{1}{n\mathcal{I}(\theta)}.$$

Verificação:

Verifique se os **estimadores de máxima verossimilhança** abaixo são eficientes.

- $X \sim B(1, \theta)$: $T = \overline{X}$;
- $X \sim N(\mu, \sigma^2)$ (σ^2-**conhecido**): $T = \overline{X}$;
- $X \sim Ex(\lambda)$: $T = \frac{1}{\overline{X}}$.

Formulário

Quantidade de Informação de Fischer

Para verificar se $V(T) = \dfrac{1}{n\mathcal{I}(\theta)}$**:**

Quadro 7.1 de [Murteira et al. (2015), p. 415]

- $X \sim B(k,\theta)$ **(**$k-$**conhecido):** $\mathcal{I}(\theta) = \dfrac{k}{\theta(1-\theta)}$
- $X \sim BN(r,\theta)$ **(**$r-$**conhecido):** $\mathcal{I}(\theta) = \dfrac{r}{\theta^2(1-\theta)}$
- $X \sim Po(\lambda)$**:** $\mathcal{I}(\lambda) = \dfrac{1}{\lambda}$
- $X \sim N(\mu,\sigma^2)$ **(**σ^2-**conhecido):** $\mathcal{I}(\mu) = \dfrac{1}{\sigma^2}$
- $X \sim N(\mu,\sigma^2)$ **(**$\mu-$**conhecido):** $\mathcal{I}(\sigma^2) = \dfrac{1}{2\sigma^4}$
- $X \sim Ex(\lambda)$**:** $\mathcal{I}(\lambda) = \dfrac{1}{\lambda^2}$

Erro Quadrático Médio (EQM)

[Murteira et al. (2015), pp. 417–418]

Definição do EQM

Se $T = T(X_1, X_2, \ldots, X_n)$ é um **estimador** de θ, então

$$EQM(T) = E((T-\theta)^2).$$

dá-nos o **erro quadrático médio (EQM)** de T.

Fórmula para cálculo do EQM

Se $T = T(X_1, X_2, \ldots, X_n)$ é um estimador de θ, então

$$EQM(T) = V(T) + Env(T)^2$$

ou $EQM(T) = V(T) + B(T)^2$.

Observação: $EQM(T) = V(T)$ no caso de T ser um **estimador cêntrico/centrado (ou não enviesado)**.

Estimador Consistente

[Murteira et al. (2015), pp. 419–4422]

Estimador Consistente em Média Quadrática

$T_n = T(X_1, X_2, \ldots, X_n)$ diz-se **consistente em média quadrática** para θ, e escreve-se $T_n \xrightarrow{mq} \theta$, se e só se

$$\lim_{n\to+\infty} EQM(T_n) = 0 \iff \lim_{n\to+\infty} E(T_n) = \theta \quad \wedge \lim_{n\to+\infty} V(T_n) = 0.$$

Estimador [Simplesmente] Consistente

$T_n = T(X_1, X_2, \ldots, X_n)$ **[simplesmente] consistente** de θ, e escreve-se $T_n \xrightarrow{p} \theta$, se e só se, para todo o $\varepsilon > 0$

$$\lim_{n\to+\infty} P(\theta - \varepsilon < T_n < \theta + \varepsilon) = 1$$

Estimador [Simplesmente] Consistente (condição suficiente)

$$T_n \xrightarrow{mq} \theta \Longrightarrow T_n \xrightarrow{p} \theta$$

Conteúdos

Exercícios Resolvidos

Exercício 1., página 447

Formulação 1.a)

Temos de considerar equação $E(X) = \overline{x}$, sendo:

- $E(X) = \frac{1-\theta}{\theta}$;
- $1000\overline{x} = \sum_{i=1}^{1000} x_i = 980 \Longrightarrow \overline{x} = 0.98.$

Resolução 1.a)

Da igualdade $\frac{1-\theta}{\theta} = \frac{1}{\theta} - 1$ segue que

$$\begin{aligned} E(X) = \overline{x} \quad &\Longleftrightarrow \quad \frac{1}{\theta} - 1 = 0.98 \\ &\Longleftrightarrow \quad \frac{1}{\theta} = 1.98 \\ &\Longleftrightarrow \quad \theta = \frac{1}{1.98} = 0.(50) \end{aligned}$$

Formulação 1.b)

Precisamos de construir a **função de verossimilhança** $L(\theta)=\prod_{i=1}^{1000} f(x_i|\theta)$ a partir de $f(x|\theta)=\theta(1-\theta)^x$.

Cálculo $L(\theta)$ em 1.b)

$$
\begin{aligned}
L(\theta) &= \prod_{i=1}^{1000} f(x_i|\theta) \\
&= \prod_{i=1}^{1000} \underbrace{\theta}_{\text{constante}} \quad \underbrace{(1-\theta)^{x_i}}_{\text{exponencial}} \\
&= \theta^{1000}(1-\theta)^{\sum_{i=1}^{1000} x_i} = \theta^{1000}(1-\theta)^{1000\overline{x}}.
\end{aligned}
$$

Maximização $\ell(\theta)$ em 1.b)

$$L(\theta) = \theta^{1000}(1-\theta)^{1000\overline{x}}$$

$$\overset{\ell(\theta)=\ln L(\theta)}{\Longrightarrow} \ell(\theta) = 1000\ln(\theta) + 1000\overline{x}\ln(1-\theta)$$

- **Condições de primeira ordem:**
  Verificar que $\ell'(\theta) = \frac{1000}{\theta} - \frac{1000\overline{x}}{1-\theta}$ **se anula** (i.e. $\ell'(\theta) = 0$) quando $\theta = \frac{1}{1+\overline{x}}$.
- **Condições de segunda ordem:**
  Verificar que $\ell''(\theta) = -\frac{1000}{\theta^2} - \frac{1000\overline{x}}{(1-\theta)^2}$ **é negativa** quando $\theta = \frac{1}{1+\overline{x}}$ (i.e $\ell''\left(\frac{1}{1+\overline{x}}\right) < 0$).

Solução 1.b)

$\widehat{\theta} = \frac{1}{1+\overline{X}}$ é o **estimador de máxima verossimilhança**, uma vez que a estimativa $\theta = \frac{1}{1+\overline{x}}$ corresponde a um máximo local da função de log-verossimilhança $\ell(\theta)$.

Solução 1.c)

Sendo $\mu = \frac{1-\theta}{\theta}$ a média da população ($E(X) = \frac{1-\theta}{\theta}$), e $\theta = \frac{1}{1+\overline{x}}$ a estimativa obtida pelo método da máxima verossimilhança, segue que

$$\mu = \frac{1 - \frac{1}{1+\overline{x}}}{\frac{1}{1+\overline{x}}} = \frac{\frac{(\cancel{1}+\overline{x})-\cancel{1}}{\cancel{1+\overline{x}}}}{\frac{1}{\cancel{1+\overline{x}}}} = \overline{x}.$$

Assim, de $\sum_{i=1}^{1000} x_i = 980$, concluímos que $\overline{x} = 0.98$ é a **estimativa de máxima verossimilhança** para $E(X)$.

Formulação 1.d)

Precisamos de construir a **função de verossimilhança** $L(\mu) = \prod_{i=1}^{1000} f(x_i|\theta(\mu))$ a partir de $\mu = \frac{1-\theta}{\theta}$.

Cálculo $L(\theta)$ em 1.b)

$$\mu = \frac{1-\theta}{\theta} \iff \theta = \frac{1}{1+\mu}$$

$$\begin{aligned} \Rightarrow L(\mu) &= \prod_{i=1}^{1000} f\left(x_i|\frac{1}{1+\mu}\right) \\ &= \left(\frac{1}{1+\mu}\right)^{1000}\left(1-\frac{1}{1+\mu}\right)^{1000\overline{x}} \\ &= (1+\mu)^{-1000-1000\overline{x}}\mu^{1000\overline{x}}. \end{aligned}$$

Maximização $\ell(\theta)$ em 1.d)

$$L(\mu) = (1+\mu)^{-1000-1000\overline{x}}\mu^{1000\overline{x}}$$

$$\overset{\ell(\mu)=\ln L(\mu)}{\Longrightarrow} \ell(\mu) = -1000(1+\overline{x})\ln(1+\mu) + 1000\overline{x}\ln(\mu)$$

- **Condições de primeira ordem:**
  Verificar que $\ell'(\mu) = 0$ quando $\mu = \overline{x}$.
- **Condições de segunda ordem:** Verificar que

$$\ell''(\overline{x}) = \left[\frac{1000(1+\overline{x})}{(1+\mu)^2} - \frac{1000\overline{x}}{\mu^2}\right]_{\mu=\overline{x}} = \frac{1000}{1+\overline{x}} - \frac{1000}{\overline{x}} < 0.$$

Formulação 3. via método dos momentos

Temos de considerar equações $E(X) = \overline{x}$ e $V(X) = s^2$, sendo:

- $\overline{x} = \frac{1}{320}\sum_{i=1}^{320} x_i = \frac{22.2}{320} = 0.069375$;
- $s^2 = \frac{1}{320}\sum_{i=1}^{320} x_i^2 - \overline{x}^2 = \frac{535.8}{320} - \left(\frac{22.2}{320}\right)^2 = 1.669562$.

Resolução sistema de equações

De $E(X) = \alpha\theta$ & $V(X) = \alpha\theta^2$ segue que

$$\begin{cases} E(X) = \overline{x} \\ V(X) = s^2 \end{cases} \Leftrightarrow \begin{cases} \alpha\theta = \overline{x} \\ \alpha\theta^2 = s^2 \end{cases} \Leftrightarrow \begin{cases} \alpha\theta = \overline{x} \\ \overline{x}\theta = s^2 \end{cases} \Leftrightarrow \begin{cases} \alpha = \frac{\overline{x}^2}{s^2} \\ \theta = \frac{s^2}{\overline{x}} \end{cases}$$

Exercícios Resolvidos

Exercício 8., página 448

Formulação 8.

Temos de estimar μ e σ^2 de $N(\mu, \sigma^2)$ a partir de $\overline{x}$ e s^2, respectivamente, para podermos calcular $P(X < 83)$.

Cálculo $\overline{x}$ e s^2

De $\sum_{i=1}^{10} x_i = 846$ e $\sum_{i=1}^{10} x_i^2 = 71607$ segue que

- $\overline{x} = \frac{1}{10} \sum_{i=1}^{10} x_i = \frac{846}{10} = 84.6$;
- $s^2 = \frac{1}{10} \sum_{i=1}^{10} x_i^2 - \overline{x}^2 = \frac{71607}{10} - \left(\frac{846}{10}\right)^2 = 3.54$.

Resolução 8.

Sendo $\overline{x} = 84.6$ e $s^2 = 3.54$ segue que

$$P(X < 83) = \Phi\left(\frac{83-\mu}{\sigma}\right) \approx \Phi\left(\frac{83-\overline{x}}{s}\right) = 1 - \Phi\left(\frac{1.6}{\sqrt{3.54}}\right)$$

Resolução 8. em *R*

```
> mean<-84.6
> sd<-sqrt(3.54)
> pnorm(83,mean,sd)
[1] 0.197554
> pnorm((83-mean)/sd)
[1] 0.197554
> 1-pnorm(1.6/sqrt(3.54))
[1] 0.197554
```

Cálculo $L(\theta)$ em 10.a)

$$
\begin{aligned}
L(\theta) &= \prod_{i=1}^{n} f(x_i|\theta) \\
&= \prod_{i=1}^{n} \underbrace{\frac{1}{\theta}}_{\text{constante}} \quad \underbrace{e^{-\frac{x_i-2}{\theta}}}_{\text{exponencial}} \\
&= \frac{1}{\theta^n} e^{-\frac{1}{\theta}\sum_{i=1}^{n}(x_i-2)} = \frac{1}{\theta^n} e^{-\frac{n(\overline{x}-2)}{\theta}}.
\end{aligned}
$$

Maximização $\ell(\theta) = \ln L(\theta)$ em 10.a)

$$L(\theta) = \frac{1}{\theta^n} e^{-\frac{n(\overline{x}-2)}{\theta}}.$$

$$\stackrel{\ell(\theta)=\ln L(\theta)}{\Longrightarrow} \ell(\theta) = -n\ln(\theta) - \frac{n(\overline{x}-2)}{\theta}$$

Para concluir:

- **Condições de primeira ordem:** Verificar que $\ell'(\theta) = 0 \Longleftrightarrow \theta = \overline{x} - 2.$
- **Condições de segunda ordem:** Verificar que $\ell''(\overline{x} - 2) < 0.$

Consistência de $\widehat{\theta} = \overline{X} - 2$ em 10.a)

Comecemos por observar que a igualdade

$$E(X) = \overbrace{\theta + 2}^{=\mu} \implies E(\widehat{\theta}) = E(\overline{X} - 2) = \overbrace{E(\overline{X})}^{=E(X)} - 2 = \theta$$

permite-nos concluir que o estimador $\widehat{\theta} = \overline{X} - 2$ **é centrado** e, por conseguinte $\lim\limits_{n\to+\infty} E(\widehat{\theta}) = \theta \iff \lim\limits_{n\to+\infty} Env(\widehat{\theta}) = 0$. Para concluir exercício:

- **Propriedades da variância:** $V(\widehat{\theta}) = V(\overline{X} - 2) = V(\overline{X}) = \frac{V(X)}{n} = \frac{\theta^2}{n}$.
- **Mostrar que $\widehat{\theta} \xrightarrow{p} \theta$:** Verificar que $\lim\limits_{n\to+\infty} EQM(\widehat{\theta}) = 0$.

Resolução 10.b)

Sendo

$$\overline{x}-2=\frac{\sum_{i=1}^{50}x_i}{50}-2=\frac{350}{50}-2=5$$

a estimativa de máxima verossimilhança de $\widehat{\theta}=\overline{X}-2$, segue da mudança de variável $x=y+2$ na f.d.p que $Y\sim Ex\left(\frac{1}{\theta}\right)$. Logo:

$$P(X>5)=P(Y+2>5)=P(Y>3)=e^{-\frac{3}{\theta}}\approx e^{-\frac{3}{5}}.$$

Resolução 10.b) em *R*

```
> theta<-5
> 1-pexp(3,1/theta)
[1] 0.5488116
```

Para resolução Exercício 18.

Precisamos de considerar os **três (3) estimadores auxiliares**, já abordados nos slides Semana6.pdf:

- $$U_1 = \sum_{i=1}^{n} \frac{(X_i - \mu)^2}{\sigma^2} \sim \chi^2_n$$
  $$\Longrightarrow E(U_1) = n \ \wedge \ V(U_1) = 2n;$$
- $$U_2 = \frac{nS^2}{\sigma^2} \sim \chi^2_{n-1}$$
  $$\Longrightarrow E(U_2) = n-1 \ \wedge \ V(U_2) = 2(n-1);$$
- $$U_3 = \frac{(n-1)(S')^2}{\sigma^2} \sim \chi^2_{n-1}$$
  $$\Longrightarrow E(U_3) = n-1 \ \wedge \ V(U_3) = 2(n-1).$$

Exercício 18.(a)– Enviesamento e Consistência de T_1

$T_1 = \frac{\sigma^2}{n} U_1$:

- $E(T_1) = \frac{\sigma^2}{n} \underbrace{E(U_1)}_{=n} = \sigma^2$ **(centrado relativamente a σ^2)**;
- $V(T_1) = \left(\frac{\sigma^2}{n}\right)^2 \underbrace{V(U_1)}_{=2n} = \frac{2\sigma^4}{n} \xrightarrow{n \to +\infty} 0$

  (consistente pois é centrado e $\lim_{n \to +\infty} V(T_1) = 0$).

Exercício 18.(a)– Enviesamento e Consistência de T_2

$$T_2 = \frac{\sigma^2}{n} U_2 :$$

- $E(T_2) = \frac{\sigma^2}{n} \underbrace{E(U_2)}_{=n-1} = \frac{n-1}{n}\sigma^2$

  (é enviesado relativamente a σ^2);

- $V(T_2) = \left(\frac{\sigma^2}{n}\right)^2 \underbrace{V(U_2)}_{=2(n-1)} = \frac{2(n-1)}{n^2}\sigma^4 \overset{n\to+\infty}{\longrightarrow} 0$

  (consistente pois $\lim\limits_{n\to+\infty} E(T_2) = \sigma^2$ e $\lim\limits_{n\to+\infty} V(T_2) = 0$).

Exercício 18.(a)– Enviesamento e Consistência de T_3

$$T_3 = \frac{\sigma^2}{n-1} U_3 :$$

- $E(T_3) = \frac{\sigma^2}{n-1} \underbrace{E(U_3)}_{=n-1} = \sigma^2$ **(é centrado relativamente a σ^2);**
- $V(T_3) = \left(\frac{\sigma^2}{n-1}\right)^2 \underbrace{V(U_3)}_{=2(n-1)} = \frac{2\sigma^4}{n-1} \xrightarrow{n\to+\infty} 0$

  (consistente pois é centrado e $\lim_{n\to+\infty} V(T_3) = 0$).

Exercício 18.(a)– Eficiência

- O estudo da eficiência reduz-se apenas a T_1 e T_3 **(os únicos que são centrados)**;
- De $V(T_1) = \frac{2\sigma^4}{n}$ & $V(T_3) = \frac{2\sigma^4}{n-1}$ segue que

$$\frac{V(T_1)}{V(T_3)} = \frac{\frac{\cancel{2\sigma^4}}{n}}{\frac{\cancel{2\sigma^4}}{n-1}} = \frac{n-1}{n} < 1 \Longrightarrow V(T_1) < V(T_3).$$

(T_1 é o estimador mais eficiente em termos relativos).

- De facto, T_1 **também é o estimador mais eficiente em termos absolutos**, uma vez que $V(T_1) = \frac{1}{n\mathcal{I}(\sigma^2)}$ (cf. [Murteira et al. (2015), p. 415]).

Exercício 18.(b)– Erro Quadrático Médio (EQM)

- De T_1 e T_3 **(os únicos que são centrados)**, já sabemos pela **fórmula do EQM** que $EQM(T_1) < EQM(T_3)$.
- De $E(T_2) = \frac{n-1}{n}\sigma^2$ & $V(T_2) = \frac{2(n-1)}{n^2}\sigma^4$ segue que

$$EQM(T_2) = V(T_2) + Env(T_2)^2 = \frac{2(n-1)}{n^2}\sigma^4 + \left(\frac{\sigma^2}{n}\right)^2 = \frac{2n-1}{n^2}\sigma^4$$

- **Comparando com** $EQM(T_1) = \frac{2\sigma^4}{n}$:

$$\frac{EQM(T_2)}{EQM(T_1)} = \frac{\frac{2n-1}{n^2}\cancel{\sigma^4}}{\frac{2\cancel{\sigma^4}}{n}} = \frac{2n^2-n}{2n^2} < 1$$

$$\Longrightarrow EQM(T_2) < EQM(T_1) < EQM(T_3).$$

Conteúdos

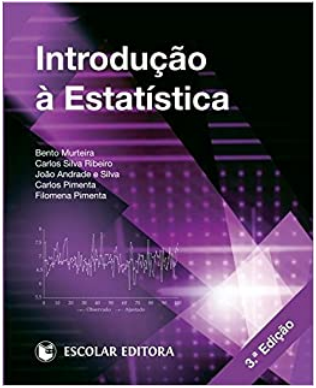
Exercícios Propostos

Complemento da Aula 13 & Aula 14

Método dos Momentos/Máxima Verossimilhança, Enviesamento e Consistência

- **págs. 447–451:**
  2. a), b), c), 6., 9., 16.

Murteira, B., C. Silva Ribeiro, J. Andrade e Silva, C. Pimenta, F. Pimenta (2015), **Introdução à Estatística, 3ª Edição**', Escolar Editora.

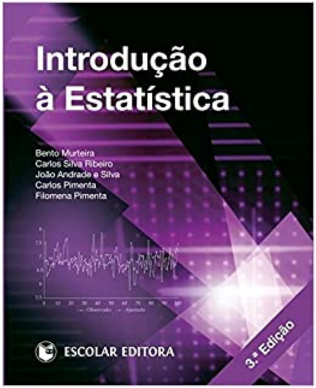


Exercícios Propostos

Complemento da Aula 15

Propriedades dos Estimadores

- **págs. 451–452:**
  17., 19. a),b), 20., 21., 23.



Exercício 2., p. 447

a) Verificar que para $X \sim Po(\lambda)$ que a função de verossimilhança é dada por $\prod_{i=1}^{n} f(x_i|\lambda) = e^{-n\lambda} \dfrac{\lambda^{n\overline{x}}}{\prod_{i=1}^{n} x_i!}$. A partir daqui:

- Mostrar que $\widehat{\mu} = \overline{X}$;
- Determinar a **estimativa de máxima verossimilhança**, $\overline{x}$, a partir de $(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9) = (95, 100, 80, 70, 110, 98, 97, 90, 70)$.

b) Verificar que $\lim_{n\to+\infty} E(\overline{X}) = \lambda$ e $\lim_{n\to+\infty} V(\overline{X}) = 0$ para **demonstrar consistência**. Para **demonstrar a eficiência**, tem de verificar que $V(\overline{X}) = \frac{1}{n\mathcal{I}(\theta)}$, onde $\mathcal{I}(\theta)$ denota a informação de Fischer de $X \sim Po(\lambda)$ (cf. [Murteira et al. (2015), pp. 414–415]).

Exercício 2., p. 447

c) Sendo $X \sim Po(\lambda)$ descreve o número de ocorrências num intervalo de 1 horas, terá de considerar $X(t) \sim Po(\lambda t)$ para descrever o **número de ocorrência num intervalo de 2 minutos** $\left(t = \frac{2}{60} = \frac{1}{30}\right)$. A partir daqui:

- Verificar que $\tau(\lambda) = P(X(t) = 0)$ coincide com $e^{-\frac{\lambda}{30}}$;
- Verificar que $e^{-\frac{\overline{x}}{30}}$ nos dá a **estimativa de máxima verossimilhança** (cf. [Murteira et al. (2015), p. 410]).

d) **Ignorar.**

Exercício 6., p. 448

a) Determinar $\widehat{\theta}$ a partir da equação $E(X) = \overline{x}$.

b) Para $(x_1, x_2, x_3, x_4, x_5, x_6) = (1, 0, 2, 3, 2, 0)$, determinar a **função de máxima verossimilhança** a partir da igualdade $L(\theta) = f(1|\theta)f(0|\theta)f(2|\theta)f(3|\theta)f(2|\theta)f(0|\theta)$.

Exercício 9., p. 449

a) $f(x|\theta)$ corresponde à função densidade de probabilidade de $X \sim N(0, \theta)$.

b) Note que $X \sim N(0, \theta)$, obtemos $E(X_i^2) = V(X_i) + E(X_i) = \theta$. Adicionalmente, para verificar que o estimador é **não enviesado** ($Env(T) = 0$) e **consistente** ($\lim_{n \to +\infty} EQM(T) = 0$), teremos que recorrer às propriedades do próximo slide.

c) **Ignorar.**

Variáveis Aleatórias Independentes

Média e Variância

Média da População

$$E(b) = b$$

$$E(aX+b) = aE(X)+b \; (a, b \in \mathbb{R})$$

Variância da população

$$V(b) = 0$$

$$V(aX+b) = a^2 V(X) \; (a, b \in \mathbb{R})$$

Média v.a.'s independentes

$$Y = \sum_{i=1}^{n} a_i X_i^2 \overset{\text{linearidade}}{\Longrightarrow} E(Y) = \sum_{i=1}^{n} a_i E(X_i^2)$$

Variância v.a.'s independentes

$$Y = \sum_{i=1}^{n} a_i X_i^2 \overset{\text{independência das v.a.'s}}{\Longrightarrow} V(Y) = \sum_{i=1}^{n} a_i^2 V(X_i^2)$$

Teste os seus conhecimentos

Distribuição Geométrica

Distribuição Geométrica/Pascal $BN(1,\theta)$

1. Determine o **estimador de máxima verossimilhança** para $X \sim BN(1,\theta)$ a partir de $f(x|\theta) = \theta(1-\theta)^{x-1}$ $(x = 1, 2, \ldots)$;
2. Verifique se o **estimador de máxima verossimilhança** – determinado anteriormente – é o **estimador mais eficiente**;
3. Diga, justificando, qual o **estimador de máxima verossimilhança** para a média da população $X \sim BN(1,\theta)$, $\mu = E(X)$.
4. Estude $\widehat{\mu}$ de $\mu = E(X)$ – determinado anteriormente – quando ao **enviesamento e consistência**.

Cálculo de Média e Variância [corrigida] da Amostra

Implementação em R

Comandos R

- **length** – número de elementos de `> x`.
- **med** – estimativa para a média da amostra, $\overline{x}$.
- **var** – estimativa para variância corrigida da amostra, $(s')^2$.

Calcular média e variância corrigida de uma população $N(5,9)$

- `> x<-rnorm(23,5,9)` para **gerar aleatoriamente** uma amostra de dimensão 23 de $N(5,9)$;
- **Estimativa para média da amostra:** `> mean(x)`
- **Estimativa para variância corrigida da amostra:** `> var(x)`
- **Estimativa para variância corrigida da amostra:**
  `> (length(x)-1)/length(x)*var(x)` ou
  `> mean((x-mean(x))^2)`.

Máximo e Mínimo da Amostra

vide [Murteira et al. (2015), pp. 362-363]

Mínimo da Amostra

$$X_{(1)} = \min_{1 \leq i \leq n} (X_i)$$

Máximo da Amostra

$$X_{(n)} = \max_{1 \leq i \leq n} (X_i)$$

Ordenação Mínimo da Amostra

$$X_{(1)} \geq x \overset{\text{ordenação}}{\Longleftrightarrow} X_1 \geq x \ \wedge X_2 \geq x \wedge \ldots \wedge X_n \geq x$$

Ordenação Máximo da Amostra

$$X_{(n)} \geq x \overset{\text{ordenação}}{\Longleftrightarrow} X_1 \leq x \ \wedge X_2 \leq x \wedge \ldots \wedge X_n \leq x$$

Teste os seus conhecimentos

Distribuição Uniforme

Distribuição Uniforme $U(a, b)$

Use o **método dos momentos** para mostrar que os estimadores $\widehat{a}$ e $\widehat{b}$ dos parâmetros a e b, respectivamente, são dados por

$$\widehat{a} = \overline{X} - \sqrt{3}S \text{ e } \widehat{b} = \overline{X} + \sqrt{3}S.$$

Distribuição Uniforme $U(a, b)$ em R

Verifique se as estimativas de máxima verossimilhança obtidas a partir dos comandos estimativa aproximada de $a = 2$ e $b = 5$, a partir de uma amostra casual de dimensão $n = 326$.

```
> x<-runif(326,2,5)
> media<-mean(x)
> variancia<-mean((x-media)^2)
> media-sqrt(3)*sqrt(variancia)
> media-sqrt(3)*sqrt(variancia)
```

Teste os seus conhecimentos

Distribuição Uniforme

Distribuição Uniforme $U(a,b)$

Verifique que:

- $L(a,b) = \frac{1}{(b-a)^n}\ (a \leq x_{(1)} < x_{(n)} \leq b)$ é a **função de verossimilhança** de $X \sim U(a,b)$;
- Para a **função de log-verossimilhança**, $\ell(a,b) = \ln L(a,b)$, se tem $\ell(x_{(1)}, x_{(n)}) \geq \ell(c,d)$ para todos os c, d tais que $c \leq x_{(1)} < x_{(n)} \leq d$.

Distribuição Uniforme $U(a,b)$ em R

A partir da amostra de $X \sim U(a,b)$, gerada aleatoriamente a partir do comando `> x<-runif(326,2,5)` **(slide anterior)**, verifique se os comandos abaixo nos permitem obter estimativas mais precisas de $a = 2$ e $b = 5$.

```
> min(x)
> max(x)
```

Distribuição Exponencial Negativa

Função de distribuição et al.

As fórmulas abaixo poderão vir a ser úteis para resolver as questões do slide seguinte:

Função de Densidade ($X \sim Ex(\lambda)$)

$$f(x|\lambda) = \begin{cases} 0 & (x \leq 0) \\ \lambda e^{-\lambda x} & (x > 0) \end{cases}$$

Quantil de ordem α (q_α)

$$F(q_\alpha) = \alpha \quad \Leftrightarrow \quad e^{-\lambda q_\alpha} = 1 - \alpha$$
$$\Leftrightarrow \quad q_\alpha = \frac{-\ln(1-\alpha)}{\lambda}$$

Teste os seus conhecimentos

Verdadeiros (V) ou Falsos (F)

Diga, justificando, se as afirmações abaixo são (V) ou (F)

a) $\widehat{\theta} = \frac{\overline{X}}{5}$ é um estimador de θ obtido pelo **método dos momentos** para a população $X \sim B(5, \theta)$.

b) $\widehat{\theta} = \ln(2)\overline{X}$ é o **estimador de máxima verossimilhança** para a mediana $Me = q_{0.5}$ de uma população $X \sim Ex(\lambda)$.

c) $\widehat{\theta} = -\ln(0.25)\overline{X}$ é o **estimador de máxima verossimilhança** para o primeiro quartil $q_{0.25}$, de uma população $X \sim Ex(\lambda)$.

d) $3\overline{X}$ é o **estimador mais eficiente** para $V(X)$ de $X \sim Po(3\lambda)$.

e) Se T é um **estimador enviesado** de θ, então a desigualdade $EQM(T) \geq V(T)$ é verdadeira.

f) Se $EQM(T) = V(T)$, então T é um **estimador centrado**.

g) Se T é um **estimador consistente**, então T é um estimador **não enviesado**.