Estatística II

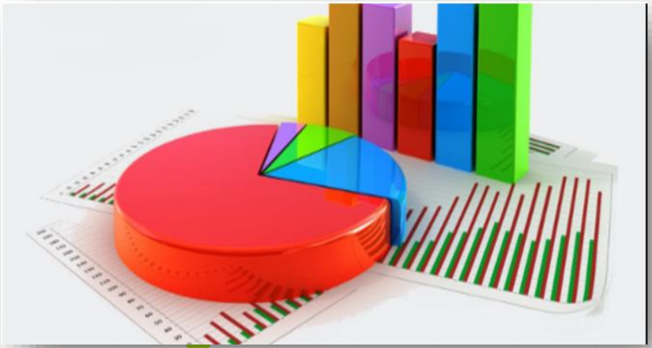
Licenciatura em Gestão do Desporto

2.º Ano/2.º Semestre

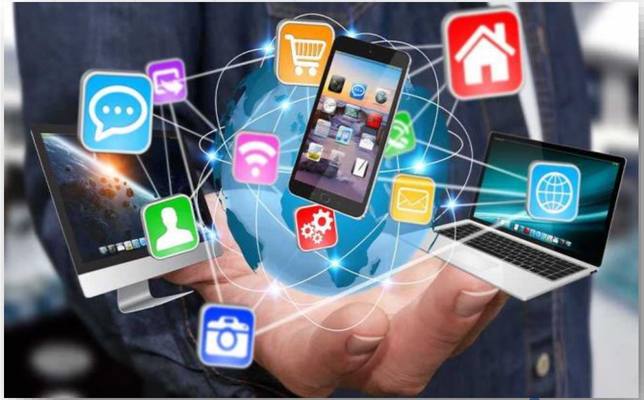
2024/2025

Aulas Teórico-Práticas N.º 9 e 10 (Semana 5)

Docente: Elisabete Fernandes
E-mail: efernandes@iseg.ulisboa.pt



https://doity.com.br/estatistica-aplicada-a-nutricao



https://basiccode.com.br/produto/informatica-basica/

Conteúdos Programáticos

Aulas Teórico-Práticas (Semanas 1 a 3)	Aulas Teórico-Práticas (Semanas 4 a 7)	Aulas Teórico-Práticas (Semanas 7 a 9)	Aulas Teórico-Práticas (Semanas 10 a 13)
• **Capítulo 1**: Revisões e Distribuições de Amostragem	• **Capítulo 2**: Estimação	• **Capítulo 3**: Testes de Hipóteses	• **Capítulo 4**: Modelo de Regressão Linear Múltipla

Material didático: Exercícios do Livro Murteira et al (2015), Formulário e Tabelas Estatísticas

Bibliografia: B. Murteira, C. Silva Ribeiro, J. Andrade e Silva, C. Pimenta e F. Pimenta; *Introdução à Estatística*, 2ª ed., Escolar Editora, 2015.

1 Propriedades dos Estimadores

Estimador Centrado, Consistente e Eficiente

Principais Propriedades dos Estimadores

Principais propriedades desejáveis nos estimadores:

- *Não enviesamento* – em termos médios, o estimador atinge o valor real do parâmetro;
- *Eficiência* – o estimador é mais eficiente quanto menor for a sua variância;
-
- *Consistência* – para n grande, o estimador deve ser aproximadamente igual ao parâmetro.

Estimador Centrado ou não Enviesado

Definição: Um estimador $\hat{\theta}$ diz-se **não enviesado** ou **centrado** do parâmetro θ se:

$$E(\hat{\theta}) = \theta.$$

Definição: Seja $\hat{\theta}$ um estimador do parâmetro θ. O **enviesamento** (Env) de $\hat{\theta}$ é dado por:

$$Env(\hat{\theta}) = E(\hat{\theta}) - \theta.$$

Viés ou Bias

ProbabilidadesEstatistica2019.pdf

Estimador Eficiente

Definição: Sejam $\hat{\theta}$ e $\tilde{\theta}$ dois estimadores centrados do parâmetro θ, baseados no mesmo número de observações. Então:

- $\hat{\theta}$ diz-se **mais eficiente** do que $\tilde{\theta}$ se: $Var(\hat{\theta}) < Var(\tilde{\theta})$
- A **eficiência relativa** do primeiro estimador relativamente ao segundo é dada por:

$$\text{Eficiência relativa} = \frac{Var(\hat{\theta})}{Var(\tilde{\theta})}.$$

$\hat{\theta}$ é um estimador eficiente de θ se ele satisfizer as seguintes condições: (i) $\hat{\theta}$ é não viesado; (ii) $Var(\hat{\theta}) \leq Var(\tilde{\theta})$, onde $\tilde{\theta}$ é qualquer outro estimador não viesado de θ.

Para comparar dois estimadores enviesados utiliza-se o critério do erro quadrático médio.

Microsoft Word - PROPRIEDADES DOS ESTIMADORES.docx (usp.br)

Definição: Seja $\hat{\theta}$ um estimador do parâmetro θ. O **erro quadrático médio** (EQM) de $\hat{\theta}$ é dado por:

$$EQM(\hat{\theta}) = Var(\hat{\theta}) + \left(Env(\hat{\theta})\right)^2.$$

Observação: Se $\hat{\theta}$ um estimador centrado para o parâmetro θ, então $Env(\hat{\theta}) = 0$ e, por conseguinte, $EQM(\hat{\theta}) = Var(\hat{\theta})$.

estimador viesado jamais será eficiente, por menor que seja o viés.

Microsoft Word - PROPRIEDADES DOS ESTIMADORES.docx (usp.br)

Definição: Sejam $\hat{\theta}$ e $\tilde{\theta}$ dois estimadores. Diz-se que $\hat{\theta}$ é "melhor" que $\tilde{\theta}$ se:

$$EQM(\hat{\theta}) < EQM(\tilde{\theta}).$$

ProbabilidadesEstatistica2019.pdf

Erro Quadrático Médio (EQM)

3. Erro quadrático médio

Chama-se erro quadrático médio do estimador $\mathbf{\Theta^*}$ a $\boldsymbol{EQM(\Theta^*) = E\left[(\Theta^* - \theta)^2\right]}$.

Esta propriedade é um dos critérios mais usados para comparar estimadores.

É muito fácil mostrar que

$$EQM(\Theta^*) = Var[\Theta^*] + [b_\theta(\Theta^*)]^2$$

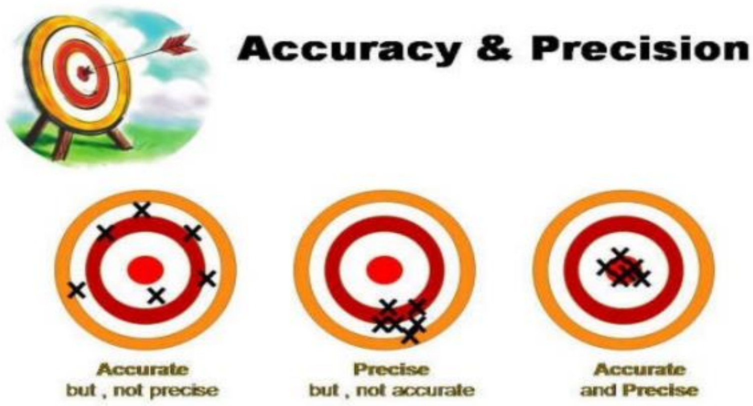
Se Θ^* é um estimador centrado de um parâmetro então o erro quadrático médio $\equiv$ variância, i.e.

$$E\left[(\Theta^* - \theta)^2\right] = Var(\Theta^*)$$

moduloI_aula3_4_Estimação.pdf

Exatidão vs Precisão

O termo exactidão (accuracy) refere-se à proximidade de uma medição ou estimativa ao verdadeiro valor.
O termo precisão ou variância (precision) refere-se ao “grau de concordância numa sucessão de medições”.



modulol_aula3_4_Estimação.pdf

Eficiência Relativa vs Eficiência Absoluta

Observações:

- A eficiência exige a existência de momentos de segunda ordem dos estimadores.
- A definição de eficiência apresenta dois conceitos diferentes:
  - O primeiro estabelece uma relação entre dois estimadores centrados para θ, sendo portanto uma eficiência ***relativa***.
  - O segundo é um conceito de eficiência ***absoluta*** na classe dos estimadores centrados para θ.
- Para obter estimadores mais eficientes recorre-se à **desigualdade de Fréchet-Cramér-Rao**.

Teorema 7.1 – Seja $(X_1, X_2, \ldots, X_n)$ uma amostra casual de população com função densidade (função probabilidade) $f(x \mid \theta)$, *satisfazendo certas condições de regularidade*, e seja $T = T(X_1, X_2, \ldots, X_n)$ um estimador centrado de θ. Então,

$$\operatorname{Var}(T) \geq \frac{1}{n\Im(\theta)},$$

onde,

$$\Im(\theta) = E\left\{\left[\frac{\partial \ln f(X \mid \theta)}{\partial \theta}\right]^2\right\} = -E\left\{\left[\frac{\partial^2 \ln f(X \mid \theta)}{\partial \theta^2}\right]\right\} \text{ (quantidade de informação de Fisher)}$$

Slides Andrade e Silva

Estimador Centrado de Variância Mínima

Conhecido o limite inferior dado pela desigualdade de Fréchet-Cramér-Rao, compara-se a variância do estimador centrado em análise com este limite:

- caso sejam iguais → não existe nenhum outro estimador centrado de variância inferior sendo o estimador em análise o mais eficiente;
- caso contrário → o quociente, $\left[n\Im(\theta)\right]^{-1} / \operatorname{Var}(T)$, fornece uma indicação sobre a eficiência relativa do estimador T face ao hipotético estimador de variância igual ao limite inferior da desigualdade.

Slides Andrade e Silva

Estimador Convergente ou Consistente

Definição: Um estimador $\hat{\theta}$ diz-se **consistente** quando para qualquer valor positivo δ, se verifica a condição:

$$\lim_{n\to\infty} P\left[|\hat{\theta} - \theta| < \delta\right] = 1.$$

Teorema: As condições:

- $\lim_{n\to+\infty} E(\hat{\theta}) = \theta;$
- $\lim_{n\to+\infty} Var(\hat{\theta}) = 0.$

são suficientes para que $\hat{\theta}$ seja estimador **consistente**.

ProbabilidadesEstatistica2019.pdf

Estimador dos Momentos vs EMV

Propriedades dos estimadores obtidos pelo método dos momentos

- Em condições bastante gerais, são consistentes e possuem distribuição aproximadamente normal quando a dimensão da amostra é muito grande (distribuição assintótica).

Propriedades dos estimadores obtidos pelo método da máxima verosimilhança

- Os estimadores de máxima verosimilhança não são necessariamente centrados.
- Em condições muito gerais eles são consistentes.
- Demonstra-se que, se existir estimador mais eficiente (na óptica do teorema de Fréchet-Cramér-Rao) ele é solução única da equação $dL/d\theta = 0$ e portanto estimador de máxima verosimilhança.
- Verificadas certas condições de regularidade, os estimadores da máxima verosimilhança seguem assintoticamente uma distribuição normal. Caso haja apenas 1 parâmetro desconhecido, tem-se

$$\sqrt{n\Im(\theta)}(\hat{\theta}-\theta) \overset{a}{\sim} N(0,1)$$

Slides Andrade e Silva

Propriedades dos Estimadores: Exercícios

2

Estimador Centrado, Eficiente e Convergente

Exercício Suplementar que não consta do livro Murteira et al (2015)

1. Suponha que uma variável aleatória X representa o número de avarias de um dispositivo durante um período de tempo e que obedece a uma lei de Poisson de parâmetro λ desconhecido. Para este parâmetro foram sugeridos dois estimadores:

$$\hat{\lambda} = \frac{X_1 + \ldots + X_n}{n} \quad \text{e} \quad \tilde{\lambda} = \frac{X_1 + X_n}{2}.$$

a) Compare-os quanto ao enviesamento.

b) Deduza a variância para cada um deles.

c) Qual dos dois estimadores é mais eficiente? Justifique a sua escolha.

d) Estude os dois estimadores quanto à consistência.

ProbabilidadesEstatistica_2019 (uevora.pt)

Exercício 1 a): Estimadores Centrados

Se $X \sim P(\lambda)$, então $E(X) = Var(X) = \lambda$.

Se $X_1, X_2, \dots, X_n$ é uma a. a. duma dada população $X \sim P(\lambda)$, então $X_i \sim P(\lambda), i = 1, \dots, n$.

a) $$E(\hat{\lambda}) = E\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right) = \frac{1}{n}E(X_1 + X_2 + \cdots + X_n) = \frac{1}{n}(E(X_1) + E(X_2) + \cdots + E(X_n))$$
$$= \frac{1}{n}(\lambda + \lambda + \cdots + \lambda) = \frac{1}{n}n\lambda = \lambda.$$

$$E(\tilde{\lambda}) = E\left(\frac{X_1 + X_n}{2}\right) = \frac{1}{2}(E(X_1) + E(X_n)) = \frac{1}{2}(\lambda + \lambda) = \lambda.$$

Portanto, ambos os estimadores são centrados.

ProbabilidadesEstatistica_2019 (uevora.pt)

VALOR ESPERADO, MOMENTOS E PARÂMETROS

Formulário

$$\mathrm{Var}(X) = E(X-\mu)^2 = E(X^2) - \mu^2;$$

$$\mathrm{Cov}(X,Y) = E\{(X-\mu_X)(Y-\mu_Y)\} = E(XY) - E(X)E(Y); \qquad \rho_{X,Y} = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}$$

$\mathrm{E}(aX + bY) = aE(X) + bE(Y)$ e $\mathrm{Var}(aX + bY) = a^2\mathrm{Var}(X) + b^2\mathrm{Var}(Y) + 2ab\,\mathrm{Cov}(X,Y)$ com a, b constantes

Exercício 1 b): Variâncias dos Estimadores

$$b)\ Var(\hat{\lambda}) = Var\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right)$$

$$= \frac{1}{n} Var(X_1 + X_2 + \cdots + X_n) \underset{\substack{X_i \\ indep.}}{=} \frac{1}{n^2}(Var[X_1] + Var[X_2] + \cdots + Var[X_n])$$

$$= \frac{1}{n^2}(\lambda + \lambda + \cdots + \lambda) = \frac{1}{n^2} n\lambda = \frac{\lambda}{n}.$$

$$Var(\tilde{\lambda}) = Var\left(\frac{X_1 + X_n}{2}\right) \underset{\substack{X_i \\ indep.}}{=} \frac{1}{2^2}(Var(X_1) + Var(X_n)) = \frac{1}{4}(\lambda + \lambda) = \frac{\lambda}{2}.$$

ProbabilidadesEstatistica_2019 (uevora.pt)

Exercício 1 c): Eficiência

c) Se $n = 1$, $Var(\hat{\lambda}) > Var(\tilde{\lambda})$ $\rightarrow$ $\tilde{\lambda}$ é mais eficiente do que $\hat{\lambda}$;
Se $n = 2$, $Var(\hat{\lambda}) = Var(\tilde{\lambda})$ $\rightarrow$ $\hat{\lambda}$ é tão eficiente como $\tilde{\lambda}$;
Se $n > 2$, $Var(\hat{\lambda}) < Var(\tilde{\lambda})$ $\rightarrow$ $\hat{\lambda}$ é mais eficiente do que $\tilde{\lambda}$

ProbabilidadesEstatistica_2019 (uevora.pt)

Exercício 1 d): Consistência

d) $\hat{\lambda}$ é um estimador consistente, pois $E(\hat{\lambda}) = \lambda$ e $Var(\hat{\lambda}) = \frac{\lambda}{n} \to 0$, quando $n \to \infty$.

$\tilde{\lambda}$ não é um estimador consistente: $E(\tilde{\lambda}) = \lambda$ mas $Var(\tilde{\lambda}) = \frac{\lambda}{2}$, seja qual for o valor de n.

ProbabilidadesEstatistica_2019 (uevora.pt)

Exercício Suplementar que não consta do livro Murteira et al (2015)

Exercício 6.4 [Estimadores, erro quadrático médio, eficiência relativa.]

Se (X_1, X_2, X_3) constitui uma amostra aleatória de dimensão 3 extraída de uma população normal com valor esperado μ e variância σ^2, qual a eficiência de $\frac{1}{4}(X_1 + 2X_2 + X_3)$ relativamente a $\bar{X}$?

Exercício 6.4: Eficiência

- **V.a. de interesse**

  $X \sim \text{Normal}(\mu, \sigma^2)$

- **Amostra aleatória**

  $\underline{X} = (X_1, X_2, X_3)$ é a.a. de dimensão 3 proveniente da população X, donde $X_i \sim_{i.i.d.} X$, $i = 1, 2, 3$.

- **Parâmetro desconhecido a estimar**

  $\mu = E(X)$

- **Estimador de $\mu = E(X)$**

  $\bar{X} = \frac{1}{3}\sum_{i=1}^{3} X_i$ [média da a.a. $\underline{X}$]

- **Erro quadrático médio de $\bar{X}$**

$$
\begin{aligned}
EQM_\mu(\bar{X}) &= V(\bar{X}) + \left[bias_\mu(\bar{X})\right]^2 \\
&= V(\bar{X}) + \left[E(\bar{X}) - \mu\right]^2 \\
&\overset{X_i \sim_{i.i.d.} X}{=} \frac{V(X)}{n} + [E(X) - E(X)]^2 \\
&= \frac{V(X)}{n} \\
&= \frac{\sigma^2}{3}
\end{aligned}
$$

Slides Professora Sofia Naique

Exercício 6.4: Eficiência

- **Outro estimador de** $\mu = E(X)$

  $T = \frac{1}{4}(X_1 + 2X_2 + X_3)$

- **Erro quadrático médio de** T

$$
\begin{aligned}
EQM_\mu(T) &= V(T) + \left[bias_\mu(T)\right]^2 \\
&= V(T) + \left[E(T) - \mu\right]^2 \\
&= V\left[\frac{1}{4}(X_1 + 2X_2 + X_3)\right] + \left\{E\left[\frac{1}{4}(X_1 + 2X_2 + X_3)\right] - E(X)\right\}^2 \\
&\overset{X_i\, indep}{=} \frac{1}{4^2}\left[V(X_1) + 4V(X_2) + V(X_3)\right] + \left\{\frac{1}{4}\left[E(X_1) + 2E(X_2) + E(X_3)\right] - E(X)\right\}^2 \\
&\overset{X_i \sim X}{=} \frac{1}{4^2}(1+4+1)\,V(X) + \left[\frac{1}{4}(1+2+1)\,E(X) - E(X)\right]^2 \\
&= \frac{3}{8}V(X) \\
&= \frac{3\sigma^2}{8}
\end{aligned}
$$

Exercício 6.4: Eficiência

- **Eficiência do estimador** $T = \frac{1}{4}(X_1 + 2X_2 + X_3)$ **relativamente a** $\bar{X}$

$$
\begin{aligned}
e_\mu(T, \bar{X}) &= \frac{EQM_\mu(\bar{X})}{EQM_\mu(T)} \\
&= \frac{\frac{\sigma^2}{3}}{\frac{3\sigma^2}{8}} \\
&= \frac{8}{9}
\end{aligned}
$$

- **Comentário**

Tendo em conta que

$$e_\mu(T, \bar{X}) = \frac{8}{9} < 1,$$

i.e.,

$$EQM_\mu(\bar{X}) < EQM_\mu(T),$$

pode afirmar-se que $T = \frac{1}{4}(X_1+2X_2+X_3)$ é um estimador menos eficiente que $\bar{X}$ no que respeita à estimação de $\mu = E(X)$.

Slides Professora Sofia Naique

Exercício Suplementar que não consta do livro Murteira et al (2015)

1ª) Sendo $X_1, \ldots, X_n$ é uma amostra aleatória de uma $U(0, \theta)$.

a) Encontre o estimador pelo método dos momentos de θ e o seu EQM.

Exercício 1 a): Método dos Momentos e EQM

- **UNIFORME (CONTÍNUA)** $X \sim U(\alpha, \beta)$, $(\alpha < \beta)$

$$f(x \mid \alpha, \beta) = \frac{1}{\beta - \alpha} \quad \alpha < x < \beta \quad ; \quad E(X) = \frac{\alpha + \beta}{2} \; ; \; \text{Var}(X) = \frac{(\beta - \alpha)^2}{12} \; ;$$

Formulário

Dado que $\mathbb{E}(X) = \frac{\theta}{2}$ e o primeiro momento amostral é $m_1 = \bar{X}$, temos que pelo método dos momentos,

$$\mathbb{E}(X) = m_1 \Rightarrow \frac{\hat{\theta}_{mm}}{2} = \bar{X} \Rightarrow \hat{\theta}_{mm} = 2\bar{X}.$$

Sabendo que o EQM de um estimador $\hat{\theta}$ de um parâmetro θ é definido como

$$EQM(\hat{\theta}) = Var(\hat{\theta}) + (\mathbb{E}(\hat{\theta}) - \theta)^2,$$

e dado que $Var(X) = \frac{\theta^2}{12}$, temos que para $\hat{\theta}_{mm} = 2\bar{X}$ o EQM é dado por,

$$\begin{aligned}
\mathbb{E}(2\bar{X}) &= \frac{2\sum_{i=1}^{n} \mathbb{E}(X_i)}{n} = \frac{2n\frac{\theta}{2}}{n} = \theta, \\
\text{Viés}(2\bar{X}) &= \mathbb{E}(2\bar{X}) - \theta = \theta - \theta = 0, \\
Var(2\bar{X}) &= \frac{4\sum_{i=1}^{n} Var(X_i)}{n^2} = \frac{4n\frac{\theta^2}{12}}{n^2} = \frac{\theta^2}{3n}, \\
EQM(2\bar{X}) &= Var(2\bar{X}) + (\text{Viés}(2\bar{X}))^2 = Var(2\bar{X}) = \frac{\theta^2}{3n}.
\end{aligned}$$

Exercícios sobre EQM,ENVUM,IF e LICR.pdf

Murteira et al (2015)
Capítulo 7

2. Pretende-se ter uma ideia da densidade do tráfego em certo local, entre as 8 e as 9 horas. Sabe-se que o número de veículos que aí passam durante esse período segue um processo de Poisson. Fizeram-se as seguintes 9 observações casuais:

$$(95, 100, 80, 70, 110, 98, 97, 90, 70).$$

a) Obtenha o estimador e a estimativa da máxima verosimilhança para o número médio de veículos que passam naquele local.
b) Mostre que se trata de um estimador consistente e que é o mais eficiente.
c) Obtenha a estimativa de máxima verosimilhança para a probabilidade de decorrerem pelo menos dois minutos sem passar qualquer veículo.
d) Mostre que $\Sigma_{i=1}^{n} X_i$ é estatística suficiente para a média da população.

Exercício 2

$X \equiv$ nº de veículos entre as 8 h e as 9 h (intervalo de 1 hora)

$X \sim Poi(\lambda) \quad (\lambda > 0) \qquad n = 9$

$f_X(x|\lambda) = \dfrac{e^{-\lambda}\lambda^{x}}{x!} \quad (x = 0, 1, 2, \dots) \qquad E(X) = \lambda$

$Var(X) = \lambda$

amostra $(n = 9)$:

$(x_1, x_2, \dots, x_9) = (95, 100, 80, 70, 110, 98, 97, 90, 70)$

Exercício 2 a)

a) $\lambda = ?$

$$L(\lambda) = \prod_{i=1}^{n} f_X(x_i \mid \lambda) = \prod_{i=1}^{n} \frac{e^{-\lambda} \lambda^{x_i}}{x_i!} =$$

$$= \frac{e^{-n\lambda} \lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!} \qquad (\lambda > 0)$$

$$\ell(\lambda) = \ln L(\lambda) = \ln\left(\frac{e^{-n\lambda} \lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!} \right) =$$

$$= -n\lambda + \sum_{i=1}^{n} x_i \ln(\lambda) - \sum_{i=1}^{n} \ln(x_i!)$$

$$\frac{d}{d\lambda} \ell(\lambda) = 0 \Leftrightarrow -n + \frac{\sum_{i=1}^{n} x_i}{\lambda} = 0 \Leftrightarrow$$

$$\Leftrightarrow \frac{\sum_{i=1}^{n} x_i}{\lambda} = n \Leftrightarrow \lambda = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x}$$

Exercício 2 a)

$$\frac{d^2}{d\lambda^2}\ell(\hat{\lambda}) = \frac{d}{d\lambda}\left(-n + \frac{\sum_{i=1}^{n} x_i}{\lambda}\right)\Bigg|_{\lambda=\bar{x}} = -\frac{\sum_{i=1}^{n} x_i}{\lambda^2}\Bigg|_{\lambda=\bar{x}}$$

$$= -\frac{n\bar{x}}{\bar{x}^2} = -\frac{n}{\bar{x}} < 0 \text{ porque } n > 0 \text{ e } \bar{x} > 0.$$

Conclusão:

- $\bar{X}$ é o estimador da M.V..
- $\bar{x} = \frac{95+100+80+70+110+98+97+90+70}{9} = \frac{810}{9} = 90$ é a estimativa da M.V..

Exercício 2 b)

b)

Um estimador $\hat{\lambda}$ é consistente para λ se:

$$\begin{cases} \lim\limits_{n \to +\infty} E(\hat{\lambda}) = \lambda \\ \lim\limits_{n \to +\infty} Var(\hat{\lambda}) = 0 \end{cases}$$

Neste caso temos $\hat{\lambda} = \bar{X}$

$$\lim_{n \to +\infty} E(\hat{\lambda}) = \lim_{n \to +\infty} E(\bar{X}) = \lim_{n \to +\infty} E(X) =$$

$$= \lim_{n \to +\infty} \lambda = \lambda$$

Exercício 2 b)

$$\lim_{n \to +\infty} Var(\hat{\lambda}) = \lim_{n \to +\infty} Var(\bar{X}) = \lim_{n \to +\infty} \left(\frac{Var(X)}{n} \right) =$$

$$= \lim_{n \to +\infty} \left(\frac{\lambda}{n} \right) = 0$$

Conclusão: $\hat{\lambda}$ é estimador consistente de λ.

Exercício 2 c)

c) $Y \equiv$ nº de veículos num intervalo de 2 minutos

$$\begin{array}{l} 60 \text{ min} - \lambda \\ 2 \text{ min} - \lambda_Y \end{array} \qquad \lambda_Y = \frac{2\lambda}{60} = \frac{\lambda}{30} \qquad Y \sim \text{Poi}\left(\frac{\lambda}{30}\right)$$

$$f_Y(0) = \frac{e^{-\frac{\lambda}{30}}\left(\frac{\lambda}{30}\right)^0}{0!} = e^{-\frac{\lambda}{30}} \rightarrow$$ função de λ (estritamente decrescente)

Como $e^{-\frac{\lambda}{30}}$ é função biunívoca de λ podemos usar a propriedade de invariância do estimador da M.V.:

$$\widehat{f_Y(0)} = \widehat{e^{-\frac{\lambda}{30}}} = e^{-\frac{\hat{\lambda}}{30}} = e^{-\frac{\bar{x}}{30}} = e^{-\frac{90}{30}} = e^{-3}$$

Exercício 2 c)

Resolução alternativa usando a distribuição exponencial:

$X \sim Poi(\lambda)$ nº de veículos por hora

$Y \sim ex(\lambda)$ tempo (em horas) até à chegada do 1º veículo (ou entre dois veículos consecutivos)

$$f_Y(y|\lambda) = \lambda e^{-\lambda y} \quad (y>0)$$

$$2 \text{ min} = \frac{2}{60} \text{ horas} = \frac{1}{30} \text{ horas}$$

Porque Y é V.A. contínua

$$P\left(Y \geqslant \frac{1}{30}\right) = 1 - P\left(Y < \frac{1}{30}\right) = 1 - F_Y\left(\frac{1}{30}\right) =$$

$$= 1 - \int_0^{\frac{1}{30}} \lambda e^{-\lambda y}\, dy = 1 - \left[-e^{-\lambda y}\right]_{y=0}^{y=\frac{1}{30}} =$$

$$= 1 - \left(-e^{-\frac{\lambda}{30}} + e^0\right) = 1 - \left(1 - e^{-\frac{\lambda}{30}}\right) =$$

$$= e^{-\lambda/30}$$ → Função estritamente decrescente de λ

Exercício 2 c)

Como $e^{-\lambda/30}$ é função biunívoca de λ podemos usar a propriedade de invariância do estimador da M.V.:

$$\widehat{P\left(Y \geq \frac{1}{30}\right)} = \widehat{e^{-\lambda/30}} = e^{-\hat{\lambda}/30} = e^{-\frac{90}{30}} =$$

$$= e^{-3} \approx 0.0498$$

<u>Nota</u>: Sabendo que $F_Y(y|\lambda) = 1 - e^{-\lambda y}$ podia-se fazer:

$$F_Y\left(\frac{1}{30}\right) = 1 - e^{-\frac{\lambda}{30}}$$

Murteira et al (2015)
Capítulo 7

11. Admite-se que o tempo de funcionamento, em milhares de horas, de determinada componente electrónica, X, tem distribuição $G(3;1/\theta)$, ou seja, a sua função densidade é dada por

$$f(x\,|\,\theta)=\frac{x^2e^{-x/\theta}}{2\theta^3}\quad(x>0),\ \theta>0.$$

Com base numa amostra casual de dimensão n, definiram-se os seguintes estimadores para o parâmetro θ:

$$T_1=\frac{\overline{X}}{3}\ \text{ e }\ T_2=\frac{n\overline{X}+1}{3n}.$$

a) Estude o enviesamento e a consistência destes estimadores.
b) Qual dos estimadores tem menor erro quadrático médio?
c) Observada uma amostra casual de 10 componentes, registou-se um tempo médio de funcionamento, por componente, de 1600 horas. Calcule o estimador e a estimativa da máxima verosimilhança para θ.

Exercício 11 a)

$X \equiv$ Tempo de funcionamento (em milhares de horas) da componente electrónica

$X \sim G(3, \frac{1}{\theta})$ $\qquad E(X) = \frac{3}{1/\theta} = 3\theta$

$Var(X) = \frac{3}{(1/\theta)^2} = 3\theta^2$

$$f_X(x|\theta) = \frac{x^2 e^{-x/\theta}}{2\theta^3} \quad (x>0), \ \theta > 0.$$

$T_1 = \frac{\bar{X}}{3}$ $\qquad T_2 = \frac{n\bar{X}+1}{3n}$

Exercício 11 a)

a) $E(T_1) = E\left(\frac{\bar{X}}{3}\right) = \frac{1}{3} E(\bar{X}) = \frac{1}{3} E(X) = \frac{1}{3} 3\theta = \theta$

T_1 é centrado.

$$Var(T_1) = Var\left(\frac{\bar{X}}{3}\right) = \frac{1}{9} Var(\bar{X}) = \frac{1}{9n} Var(X) =$$

$$= \frac{1}{9n} 3\theta^2 = \frac{1}{3n}\theta^2$$

$$\lim_{n \to +\infty} E(T_1) = \lim_{n \to +\infty} \theta = \theta$$

$$\lim_{n \to +\infty} Var(T_1) = \lim_{n \to +\infty} \frac{1}{3n}\theta^2 = 0$$

T_1 é consistente

Exercício 11 a)

$$E(T_2) = E\left(\frac{n\bar{X} + 1}{3n}\right) = \frac{n\,E(\bar{X}) + 1}{3n} =$$

$$= \frac{n\,E(X) + 1}{3n} = \frac{3n\theta + 1}{3n} = \theta + \frac{1}{3n}$$

T_2 é enviesado

$$\text{Enviesamento}(T_2) = E(T_2) - \theta = \frac{1}{3n}$$

$$Var(T_2) = Var\left(\frac{n\bar{X} + 1}{3n}\right) = \frac{n^2\,Var(\bar{X})}{9n^2} =$$

$$= \frac{Var(\bar{X})}{9} = \frac{Var(X)}{9n} = \frac{3}{9n}\theta^2 = \frac{1}{3n}\theta^2$$

$$\lim_{n \to +\infty} E(T_2) = \lim_{n \to +\infty}\left(\theta + \frac{1}{3n}\right) = \theta$$

$$\lim_{n \to +\infty} Var(T_2) = \lim_{n \to +\infty}\left(\frac{1}{3n}\theta^2\right) = 0$$

T_2 é consistente

Exercício 11 b)

b)

$$EQM(T_1) = Var(T_1) + Erro(T_1)^2 = Var(T_1) = \frac{\theta^2}{3n}$$

Porque T_1 é centrado

$$EQM(T_2) = Var(T_2) + Erro(T_2)^2 =$$

$$= \frac{\theta^2}{3n} + \left(\frac{1}{3n}\right)^2 > EQM(T_1) = \frac{\theta^2}{3n}$$

Conclusão: T_1 tem menor EQM.

Exercício 11 c)

c) $n = 10$ componentes $\quad \bar{x} = 1.6$ milhares de horas.

$$L(\theta) = \prod_{i=1}^{10} \frac{x_i^2 e^{-x_i/\theta}}{2\theta^3} = \frac{1}{2^{10}\theta^{30}} \left(\prod_{i=1}^{10} x_i^2 \right) e^{-\frac{1}{\theta}\sum_{i=1}^{10} x_i} \quad (\theta > 0)$$

$$\ell(\theta) = \ln L(\theta) = \ln\left(\frac{1}{2^{10}\theta^{30}}\right) + \sum_{i=1}^{10} \ln(x_i^2) - \frac{1}{\theta}\sum_{i=1}^{10} x_i \quad (\theta > 0)$$

$$= -10\ln(2) - 30\ln(\theta) + \sum_{i=1}^{10} 2\ln(x_i) - \frac{1}{\theta}\sum_{i=1}^{10} x_i \; (\theta > 0)$$

$$\frac{d}{d\theta}\ell(\theta) = -\frac{30}{\theta} - \left(-\frac{1}{\theta^2}\sum_{i=1}^{10} x_i\right) = 0 \iff$$

$$\iff \frac{30}{\theta} = \frac{1}{\theta^2}\sum_{i=1}^{10} x_i \Rightarrow \hat{\theta} = \sum_{i=1}^{10} x_i / 30 = \frac{\bar{x}}{3}$$

Exercício 11 c)

$$\frac{d^2}{d\theta^2}\,\ell(\theta) = \frac{d}{d\theta}\left(-\frac{30}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{10} x_i\right) =$$

$$= \frac{30}{\theta^2} - \frac{2\theta}{\theta^4}\sum_{i=1}^{10} x_i = \frac{30}{\theta^2} - \frac{2}{\theta^3}\sum_{i=1}^{10} x_i =$$

$$= \frac{30}{\theta^2} - \frac{20\,\bar{x}}{\theta^3}$$

$$\frac{d^2}{d\theta^2}\,\ell(\hat{\theta}) = \frac{30}{\left(\frac{\bar{x}}{3}\right)^2} - \frac{20\,\bar{x}}{\left(\frac{\bar{x}}{3}\right)^3} = \frac{270}{\bar{x}^2} - \frac{540}{\bar{x}^2} < 0$$

<u>Conclusão</u>: $\hat{\theta} = \frac{\bar{X}}{3}$ é o estimador da M.V.

$\hat{\theta} = \frac{\bar{x}}{3} = \frac{1.6}{3}$ é uma estimativa da M.V.

Murteira et al (2015)
Capítulo 7

14. Seja X uma variável aleatória com função densidade

$$f(x|\theta) = \frac{3\sqrt{x}}{2\theta} \exp\left\{-\frac{x^{3/2}}{\theta}\right\} \quad (x > 0)\text{, para } \theta > 0.$$

a) Mostre que

$$\hat{\theta} = \frac{\sum_{i=1}^{n} X_i^{3/2}}{n},$$

é o estimador da máxima verosimilhança para θ.

b) Sabendo que $X^{3/2}$ tem distribuição exponencial de média θ, estude o enviesamento e a consistência do estimador da alínea a).

Exercício 14 a)

$$f(x|\theta) = \frac{3\sqrt{x}}{2\theta} \exp\left\{-\frac{x^{\frac{3}{2}}}{\theta}\right\} \quad (x>0), \quad para \ \theta>0$$

a)

$$L(\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} \frac{3\sqrt{x_i}}{2\theta} \exp\left\{-\frac{x_i^{\frac{3}{2}}}{\theta}\right\} =$$

$$= \frac{3^n \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{2}}}{2^n \theta^n} \exp\left\{-\frac{1}{\theta}\sum_{i=1}^{n} x_i^{\frac{3}{2}}\right\}$$

$$l(\theta) = \log\{L(\theta)\} =$$

$$= n\log(3) + \frac{1}{2}\sum_{i=1}^{n}\log(x_i) - n\log(2) - n\log(\theta) - \frac{1}{\theta}\sum_{i=1}^{n} x_i^{\frac{3}{2}}$$

Exercício 14 a)

$$\frac{d}{d\theta}\{\ell(\theta)\} = -\frac{n}{\theta} + \frac{\sum_{i=1}^{n} x_i^{\frac{3}{2}}}{\theta^2} = 0 \Leftrightarrow$$

$$\Leftrightarrow \frac{n}{\theta} = \frac{\sum_{i=1}^{n} x_i^{\frac{3}{2}}}{\theta^2} \Leftrightarrow$$

$$\Rightarrow \hat{\theta} = \frac{\sum_{i=1}^{n} x_i^{\frac{3}{2}}}{n} \rightarrow \text{Candidato a EMV}$$

$$\frac{d^2}{d\theta^2}\ell(\theta) = \frac{d}{d\theta}\left(-\frac{n}{\theta} + \frac{\sum_{i=1}^{n} x_i^{\frac{3}{2}}}{\theta^2}\right)$$

$$= \frac{n}{\theta^2} - \frac{\sum_{i=1}^{n} x_i^{\frac{3}{2}}}{\theta^4} \cdot 2\theta =$$

$$= \frac{n}{\theta^2} - 2\frac{\sum_{i=1}^{n} x_i^{\frac{3}{2}}}{\theta^3}$$

Exercício 14 a)

$$\frac{d^2}{d\theta^2}\, l(\hat{\theta}) = \frac{n}{\left(\frac{\sum_{i=1}^{n} x_i^{3/2}}{n}\right)^2} - 2\,\frac{\sum_{i=1}^{n} x_i^{3/2}}{\left(\frac{\sum_{i=1}^{n} x_i^{3/2}}{n}\right)^3} =$$

$$= \frac{n^3}{\left(\sum_{i=1}^{n} x_i^{3/2}\right)^2} - \frac{2\,n^3}{\left(\sum_{i=1}^{n} x_i^{3/2}\right)^2} = -\frac{n^3}{\left(\sum_{i=1}^{n} x_i^{3/2}\right)^2} < 0 \text{ porque } n > 0$$

<u>Conclusão</u>: $\hat{\theta} = \frac{\sum_{i=1}^{n} x_i^{\frac{3}{2}}}{n}$ é o EMV, QED

Exercício 14 b)

b) $Y = X^{\frac{3}{2}} \sim Ex\left(\frac{1}{\theta}\right)$ $\qquad E(Y) = \theta$

$Var(Y) = \theta^2$

Um $\underbrace{\text{estimador}}_{\hat{\theta}}$ de θ é consistente se: $\lim\limits_{n \to +\infty} E(\hat{\theta}) = \theta$

$\lim\limits_{n \to +\infty} Var(\hat{\theta}) = 0$

$$E(\hat{\theta}) = E\left(\frac{\sum_{i=1}^{n} X_i^{3/2}}{n}\right) = \frac{1}{n} E\left(\sum_{i=1}^{n} X_i^{3/2}\right) =$$

$$= \frac{1}{n} \sum_{i=1}^{n} E(X_i^{3/2}) = \frac{1}{n} \sum_{i=1}^{n} E(Y) =$$

↳ porque $X_i^{3/2}$ são i.d. a Y $(i = 1, 2, \dots, n)$

$$= \frac{1}{n} \, n \, E(Y) = E(Y) = \theta$$

Conclusão: $\hat{\theta}$ é um estimador centrado em θ
(não enviesado)

Exercício 14 b)

$$\lim_{n \to +\infty} E(\hat{\theta}) = \lim_{n \to +\infty} (\theta) = \theta$$

$$Var(\hat{\theta}) = Var\left(\frac{\sum_{i=1}^{n} X_i^{3/2}}{n}\right) = \frac{1}{n^2} Var\left(\sum_{i=1}^{n} X_i^{3/2}\right) =$$

$$= \frac{1}{n^2} \sum_{i=1}^{n} Var(X_i^{\frac{3}{2}}) = \frac{1}{n^2}\, n \, Var(Y) = \frac{\theta^2}{n}$$

$$\lim_{n \to +\infty} Var(\hat{\theta}) = \lim_{n \to +\infty} \left(\frac{\theta^2}{n}\right) = 0$$

Como: $\lim_{n \to +\infty} E(\hat{\theta}) = \theta$ e $\lim_{n \to +\infty} Var(\hat{\theta}) = 0$ podemos

concluir que $\hat{\theta}$ é um estimador consistente.

Murteira et al (2015)
Capítulo 7

24. Classifique as afirmações apresentadas abaixo de verdadeiras ou falsas, justificando sucintamente.

a) Com base numa amostra casual de dimensão $n = 3$, o estimador $T = 0.5\,X_1 + 0.4\,X_2 + 0.1\,X_3$, sendo centrado é também o estimador mais eficiente para μ.

b) Comparar os Erros Quadráticos Médios de dois estimadores centrados (referentes ao mesmo parâmetro) equivale a comparar as respectivas variâncias.

c) Um estimador consistente é sempre centrado.

Exercício 24 a)

a) Falso. Nada garante que não há estimadores mais eficientes. Por exemplo:

$$Var(T) = 0.5^2 Var(X_1) + 0.4^2 Var(X_2) + 0.1^2 Var(X_3) =$$

$$= 0.5^2 \sigma_X^2 + 0.4^2 \sigma_X^2 + 0.1^2 \sigma_X^2 =$$

$$= \sigma_X^2 (0.5^2 + 0.4^2 + 0.1^2) \simeq 0.42 \sigma_X^2$$

$E(\bar{X}) = E(X) = \mu$ também é estimador centrado.

$$Var(\bar{X}) = \frac{\sigma_X^2}{3} \simeq 0.33 \sigma_X^2 < Var(T) \simeq 0.42 \sigma_X^2$$

Logo $\bar{X}$ é mais eficiente que T.

Exercício 24 b)

b) Verdade

$$EQM(T) = Var(T) + bias(T)^2$$

Se o estimador for centrado temos $bias(T) = 0$

logo $EQM(T) = Var(T)$.

Exercício 24 c)

c) Falso. Existem estimadores consistentes que são enviesados.

Exemplo: S^2

Demonstração no livro (página 365)

$$E(S^2) = \frac{n-1}{n}\sigma^2 \rightarrow S^2 \text{ é estimador enviesado de } \sigma^2$$

$$\lim_{n\to+\infty} E(S^2) = \lim_{n\to+\infty} \frac{n-1}{n}\sigma^2 = \sigma^2$$

$$\lim_{n\to+\infty} Var(S^2) = \lim_{n\to+\infty}\left(\frac{\mu_4-\mu_2^2}{n} - \frac{2(\mu_4-2\mu_2^2)}{n^2} + \frac{\mu_4-3\mu_2^2}{n^3}\right) =$$

$$= 0$$

Conclusão: S^2 é enviesado e consistente.

Obrigada!

Questões?