MAE0229 – Introdução à Probabilidade e à Estatística II

I semestre de 2018

Lista de exercícios 2 - Estimação I - C A S A (gabarito)

Exercício 1:

Numa empresa financeira, o lucro gerado por cada investimento realizado, no último ano, tem um valor médio de R$24600 e um desvio padrão de R$10800. Considere uma amostra aleatória de 100 investimentos. Determine:

(a) A distribuição da média amostral.

(b) A probabilidade da média amostral exceder R$30000.

(c) A probabilidade da média amostral estar entre R$21000 e R$25500.

(d) A distribuição do lucro total amostral.

(e) O 90^o percentil do lucro total amostral.

Solução:
Temos que, $X =$"lucro gerado por cada investimento realizado", com $\mu = 24600$, $\sigma = 10800$ e $n = 100$.

(a) Se $X \sim N(\mu, \sigma^2)$, então $\overline{X} \sim N(\mu, \sigma^2/n)$.
Neste caso, $X \sim N(24600, 10800^2)$, logo, $\overline{X} \sim N(24600, 1080^2)$.

(b) $P(\overline{X} > 30000)$

$$\begin{aligned} P(\overline{X} > 30000) &= P\left(\frac{\overline{X} - 24600}{1080} > \frac{30000 - 24600}{1080}\right) \\ &= P(Z > 5) \\ &= 1 - P(Z \leq 5) \approx 0. \end{aligned}$$

(c) $P(21000 < \overline{X} < 25500)$

$$\begin{aligned} P(21000 < \overline{X} < 25500) &= P\left(\frac{21000 - 24600}{1080} < \frac{\overline{X} - 24600}{1080} \leq \frac{25500 - 24600}{1080}\right) \\ &= P(-3,333 < Z < 0,370) \\ &= P(Z < 0,370) - P(Z < -3,333) \\ &= P(Z < 0,370) - (1 - P(Z < 3,333)) \\ &= 0,6443 - 1 + 0,9996 = 0,6439. \end{aligned}$$

(d) Seja $Y = \sum_{i=1}^{100} X_i =$"lucro total".

Assim, $Y \sim N(n\mu, n\sigma^2)$.

Neste caso, $Y \sim N(2460000, 108000^2)$.

(e) Qual é o valor de a tal que, $P(Y < a) = 0,9$.

$$\begin{aligned} P(Y < a) &= P\left(\frac{Y - 2460000}{108000} < \frac{a - 2460000}{108000}\right) \\ &= P\left(Z < \frac{a - 2460000}{108000}\right), \end{aligned}$$

lembrando que $P(Y < a) = 0,9$, temos que $P\left(Z < \frac{a-2460000}{108000}\right) = 0,9$, logo,

$$\frac{a - 2460000}{108000} = 1,285 \Rightarrow a = 108000(1,285) + 2460000 = 2598780.$$

Note que a distribuição da média amostral, $\overline{X}$, e do lucro total, Y, são exatas e não apróximadas pois a distribuição do lucro gerado por cada investimento, X é normal.

Exercício 2:

Numa pesquisa de mercado, desejamos estimar a proporção de pessoas que compram determinado sabonete.

(a) Que tamanho de amostra devemos recolher para que, com probabilidade 0,9 a proporção amostral não se desvie do verdadeiro valor por mais do que 0,05?

(b) Se tivermos informação adicional de que a aceitação do sabonete é no mínimo 0,8, qual deve ser então o tamanho da amostra?

(c) Decidimos recolher uma amostra de tamanho 81. Qual o erro máximo que cometemos com probabilidade 0,9?

(d) Para a amostra de tamanho 81, qual a probabilidade de que o erro máximo seja 0,08, quando p = 0,4?

Solução:

(a) O valor de n tal que, $\varepsilon = 0,05$ e $\gamma = P(-z \leq Z \leq z) = 0,9$.

Lembre que

$$n = \left(\frac{z}{\varepsilon}\right)^2 p(1-p),$$

como $\gamma = 0,9$, z é tal que, $A(z) = 0,95$, então, $z = 1,645$. Como não temos informação sobre o valor de p, toma–se o valor que fornece a maior variância, isto é, $p = 1/2$. Portanto,

$$\begin{aligned} n &= \left(\frac{1,645}{0,05}\right)^2 \frac{1}{2}\left(1-\frac{1}{2}\right) \\ &= \left(\frac{1,645}{0,05}\right)^2 \frac{1}{4} \\ &= 270,603 \approx 271. \end{aligned}$$

Devemos recolher uma amostra de tamanho 271 para que com probabilidade 0,9 a proporção amostral não se desvie do verdadeiro valor por mais do que 0,05.

(b) Neste caso, tem–se que $p > 0,8$, logo $p(1-p) < 0,8(0,2) = 0,16$ a qual é menor do que $0,25 = 0,5(0,5)$, toma–se este novo valor no cálculo do tamanho amostral.

$$
\begin{aligned}
n &= \left(\frac{z}{\varepsilon}\right)^2 p(1-p) \\
&= \left(\frac{1,645}{0,05}\right)^2 \frac{4}{5}\left(1-\frac{4}{5}\right) \\
&= \left(\frac{1,645}{0,05}\right)^2 \frac{4}{25} \\
&= 173,186 \approx 174.
\end{aligned}
$$

Com a informação adicional, $p > 0,8$, o tamanho de amostra que se deve recolher para que, com probabilidade 0,9 a proporção amostral não se desvie do verdadeiro valor por mais do que 0,05 é de 174.

(c) Sabendo que $n = 81$, $\gamma = P(-z \leq Z \leq z) = 0,9$, se vai considerar $p = 0,5$ pois com este valor se tem a maior variância possível, se quer conhecer o valor de ε. Lembrando que

$$
\begin{aligned}
\varepsilon &= z\sqrt{\frac{p(1-p)}{n}} \\
&= 1,645\sqrt{\frac{0,5(1-0,5)}{81}} \\
&= 1,645\sqrt{\frac{1}{4(81)}} \\
&= 0,0914.
\end{aligned}
$$

O erro máximo que cometemos com probabilidade 0,9 é de 0,0914.

(d) Sabendo que $n = 81$, $\varepsilon = 0,08$ e $p = 0,4$ se quer saber o valor de γ. Note que

$$
\begin{aligned}
z &= \frac{\varepsilon\sqrt{n}}{\sqrt{p(1-p)}} \\
&= \frac{0,08\sqrt{81}}{\sqrt{0,4(1-0,4)}} \\
&= \frac{9(0,08)}{\sqrt{0,4(0,6)}} \\
&= 1,470,
\end{aligned}
$$

logo,

$$\begin{aligned}
\gamma &= P(-1,470 \leq Z \leq 1,470) \\
&= P(Z \leq 1,470) - P(Z \leq -1,470) \\
&= P(Z \leq 1,470) - 1 + P(Z \leq 1,470) \\
&= 2P(Z \leq 1,470) - 1 \\
&= 2(0,9292) - 1 = 0,8584.
\end{aligned}$$

Para a amostra de tamanho 81, a probabilidade de que o erro máximo seja 0,08, quando $p = 0,4$ é de 0,8584.

Exercício 3:

Seja X uma variável aleatória com média μ e variância $\sigma^2 = 400$. Definimos a variável $EA = \overline{X} - \mu$ como sendo o erro amostral da média.

(a) Determine $E(EA)$ e $Var(EA)$.

(b) Que proporção das amostras de tamanho 25 terão erro amostral absoluto maior do que 2 unidades?

(c) E que proporção das amostras de tamanho 100?

(d) Neste último caso, qual o valor de d, tal que $P(|EA| > d) = 1\%$?

(e) Qual deve ser o tamanho da amostra para que 95% dos erros amostrais absolutos sejam inferiores a uma unidade?

Solução:

(a)

$$\begin{aligned}
E(EA) &= E(\overline{X} - \mu) \\
&= E(\overline{X}) - E(\mu) \\
&= E\left(\frac{1}{n}\sum X_i\right) - \mu \\
&= \frac{1}{n}\sum E(X_i) - \mu \\
&= \frac{1}{n}n\mu - \mu \\
&= \mu - \mu = 0.
\end{aligned}$$

$$\begin{aligned}
Var(EA) &= Var(\overline{X} - \mu) \\
&= Var(\overline{X}) + Var(\mu) \\
&= Var\left(\frac{1}{n}\sum X_i\right) \\
&= \frac{1}{n^2}\sum Var(X_i) \\
&= \frac{1}{n^2}n\sigma^2 \\
&= \frac{\sigma^2}{n}.
\end{aligned}$$

(b) Sabendo que $n = 25$, e que $|EA| > 2$ se quer saber $P(|EA| > 2)$.

$$\begin{aligned}
P(|EA| > 2) &= P(|\overline{X} - \mu| > 2) \\
&= P(\overline{X} - \mu > 2) + P(\overline{X} - \mu < -2) \\
&= P\left(\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} > \frac{2}{20/5}\right) + P\left(\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} < -\frac{2}{20/5}\right) \\
&= P\left(Z > \frac{1}{2}\right) + P\left(Z < -\frac{1}{2}\right) \\
&= 1 - P\left(Z < \frac{1}{2}\right) + 1 - P\left(Z < \frac{1}{2}\right) \\
&= 2 - 2P\left(Z < \frac{1}{2}\right) \\
&= 2 - 2(0,6915) = 0,617.
\end{aligned}$$

A proporção das amostras de tamanho 25 que terão erro amostral absoluto maior do que 2 unidades é de 0,617.

(c) Se $n = 100$, tem–se que

$$\begin{aligned}
P(|EA| > 2) &= P(|\overline{X} - \mu| > 2) \\
&= P(\overline{X} - \mu > 2) + P(\overline{X} - \mu < -2) \\
&= P\left(\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} > \frac{2}{20/10}\right) + P\left(\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} < -\frac{2}{20/10}\right) \\
&= P(Z > 1) + P(Z < -1) \\
&= 1 - P(Z \leq 1) + 1 - P(Z \leq 1) \\
&= 2 - 2P(Z \leq 1) \\
&= 2 - 2(0,8413) = 0,3174.
\end{aligned}$$

A proporção das amostras de tamanho 100 que terão erro amostral absoluto maior do que 2 unidades é de 0,3174.

(d) O valor de d, tal que $P(|EA| > d) = 0,01$.

$$\begin{aligned}
P(|EA| > d) &= P(EA > d) + P(EA < -d) \\
&= P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} > \frac{d}{20/10}\right) + P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} < -\frac{d}{20/10}\right) \\
&= P\left(Z > \frac{d}{2}\right) + P\left(Z < -\frac{d}{2}\right) \\
&= 1 - P\left(Z < \frac{d}{2}\right) + 1 - P\left(Z < \frac{d}{2}\right) \\
&= 2 - 2P\left(Z < \frac{d}{2}\right),
\end{aligned}$$

lembrando que $P(|EA| > d) = 0,01$, tem–se que

$$2 - 2P\left(Z < \frac{d}{2}\right) = 0,01 \Rightarrow P\left(Z < \frac{d}{2}\right) = \frac{2-0,01}{2} = 0,995,$$

assim,

$$\frac{d}{2} = 2,575 \Rightarrow d = 2(2,575) = 5,15.$$

(e) Dado que $P(|EA| < 1) = 0,95$ se quer conhecer o valor de n.

$$\begin{aligned}
P(|EA| < 1) &= P(-1 < EA < 1) \\
&= P\left(-\frac{1}{20/\sqrt{n}} < \frac{EA}{\sigma/\sqrt{n}} < \frac{1}{20/\sqrt{n}}\right) \\
&= P\left(-\frac{\sqrt{n}}{20} < Z < \frac{\sqrt{n}}{20}\right) \\
&= P\left(Z < \frac{\sqrt{n}}{20}\right) - P\left(Z < -\frac{\sqrt{n}}{20}\right) \\
&= P\left(Z < \frac{\sqrt{n}}{20}\right) - 1 + P\left(Z < \frac{\sqrt{n}}{20}\right) \\
&= 2P\left(Z < \frac{\sqrt{n}}{20}\right) - 1,
\end{aligned}$$

lembrando que, $P(|EA| < 1) = 0,95$, obtém–se que

$$P\left(Z < \frac{\sqrt{n}}{20}\right) = \frac{0,95+1}{2} = 0,975,$$

logo,

$$\frac{\sqrt{n}}{20} = 1,96 \Rightarrow n = (20(1,96))^2 = 1536,64 \approx 1537.$$

O tamanho da amostra para que 95% dos erros amostrais absolutos sejam inferiores a uma unidade deve ser de 1537.

Exercício 4:

A duração de um certo tipo de equipamento, em milhares de horas, pode ser bem representada por uma variável aleatória X, com $E(X) = 2\theta$ e $Var(X) = 2\theta^2$. Com base numa amostra aleatória simples de dimensão n, definiram-se os seguintes estimadores para θ:

$$T_1 = \frac{\overline{X}}{2} \qquad \text{e} \qquad T_2 = \frac{2 + \sum_{i=1}^{n} X_i}{2n}.$$

(a) Verifique se T_1 e T_2 são estimadores não viesados para θ.

(b) T_1 e T_2 são estimadores consistentes para θ?

Solução:

(a) Devemos verificar se $E(T_i) = \theta$ com $i = 1, 2$.

$$\begin{aligned} E(T_1) &= E\left(\frac{\overline{X}}{2}\right) \\ &= \frac{1}{2}E\left(\frac{1}{n}\sum X_i\right) \\ &= \frac{1}{2n}\sum E(X_i) \\ &= \frac{1}{2n}n(2\theta) \\ &= \frac{2n}{2n}\theta = \theta. \end{aligned}$$

e

$$\begin{aligned} E(T_2) &= E\left(\frac{2 + \sum X_i}{2n}\right) \\ &= \frac{1}{2n}(E(2) + E(\sum X_i)) \\ &= \frac{1}{2n}(2 + \sum E(X_i)) \\ &= \frac{1}{2n}(2 + n(2\theta)) \\ &= \frac{2}{2n}(1 + n\theta) = \frac{n\theta + 1}{n}. \end{aligned}$$

Logo, T_1 é um estimador não viesado para θ mas T_2 não é.

(b) Se deve verificar se $\lim_{n\to\infty} E(T_i) = \theta$ e $\lim_{n\to\infty} Var(T_i) = 0$.

$$\begin{aligned} Var(T_1) &= Var\left(\frac{\overline{X}}{2}\right) \\ &= \frac{1}{4}Var\left(\frac{1}{n}\sum X_i\right) \\ &= \frac{1}{4n^2}\sum Var(X_i) \\ &= \frac{1}{4n^2}n(2\theta^2) = \frac{\theta^2}{2n}, \end{aligned}$$

$$\begin{aligned} Var(T_2) &= Var\left(\frac{2+\sum X_i}{2n}\right) \\ &= \frac{1}{4n^2}Var(\sum X_i) \\ &= \frac{1}{4n^2}\sum Var(X_i) \\ &= \frac{n(2\theta^2)}{4n^2} = \frac{\theta^2}{2n}, \end{aligned}$$

Tomando o limite da esperança e da variância, tem–se que,

$$\lim_{n\to\infty} Var(T_1) = \lim_{n\to\infty} Var(T_2) = \lim_{n\to\infty} \frac{\theta^2}{2n} = 0,$$

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

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

como

$$\lim_{n\to\infty} E(T_i) = \theta \ \text{ e } \ \lim_{n\to\infty} Var(T_i) = 0, \quad i = 1, 2,$$

os estimadores T_1 e T_2 são consistentes para θ.

Exercício 5:

Suponha que dois economistas estimem μ (despesa familiar média com alimentação) com dois estimadores não viesados $\widehat{\mu}_1$ e $\widehat{\mu}_2$ (estatisticamente independentes). O segundo economista é menos cuidadoso que o primeiro, de maneira que o desvio padrão de $\widehat{\mu}_2$ é cinco vezes maior do que o de $\widehat{\mu}_1$. Solicitado a informar sobre como combinar $\widehat{\mu}_1$ e $\widehat{\mu}_2$ para obter uma estimativa global publicável, um grupo de estatísticos apresentou as seguintes propostas:

i) $\widehat{W}_1 = \frac{1}{2}(\widehat{\mu}_1 + \widehat{\mu}_2)$;

ii) $\widehat{W}_2 = \frac{1}{5}(4\widehat{\mu}_1 + \widehat{\mu}_2)$;

iii) $\widehat{W}_3 = \frac{1}{6}(5\widehat{\mu}_1 + \widehat{\mu}_2)$;

iv) $\widehat{W}_4 = \widehat{\mu}_1$.

(a) Quais dos estimadores são não viesados?

(b) Disponha esses estimadores por ordem decrescente de eficiência.

(c) Determine a combinação linear $a\widehat{\mu}_1 + (1-a)\widehat{\mu}_2$; $0 \le a \le 1$ mais eficiente possível como função de a.

Solução:

(a) Sabemos que $E(\widehat{\mu}_1) = E(\widehat{\mu}_2) = \mu$ pois os estimadores $\widehat{\mu}_1$ e $\widehat{\mu}_2$ são não viesados, logo

$$\begin{aligned}
E(\widehat{W}_1) &= E\left(\frac{1}{2}(\widehat{\mu}_1 + \widehat{\mu}_2)\right) \\
&= \frac{1}{2}(E(\widehat{\mu}_1) + E(\widehat{\mu}_2)) \\
&= \frac{1}{2}2\mu = \mu,
\end{aligned}$$

$$\begin{aligned}
E(\widehat{W}_2) &= E\left(\frac{1}{5}(4\widehat{\mu}_1 + \widehat{\mu}_1)\right) \\
&= \frac{1}{5}(E(4\widehat{\mu}_1) + E(\widehat{\mu}_2)) \\
&= \frac{1}{5}(4\mu + \mu) \\
&= \frac{5\mu}{5} = \mu,
\end{aligned}$$

$$
\begin{aligned}
E(\widehat{W}_3) &= E\left(\frac{1}{6}(5\widehat{\mu}_1 + \widehat{\mu}_2)\right)\\
&= \frac{1}{6}(E(5\widehat{\mu}_1) + E(\widehat{\mu}_2))\\
&= \frac{1}{6}(5\mu + \mu)\\
&= \frac{6\mu}{6} = \mu,
\end{aligned}
$$

$$
E(\widehat{W}_4) = E(\widehat{\mu}_1) = \mu.
$$

Todos os estimadores são não viesados.

(b) Sabemos que $\sigma_{\widehat{\mu}_2} = 5\sigma_{\widehat{\mu}_1}$, ou seja, $Var(\widehat{\mu}_2) = 25Var(\widehat{\mu}_1)$, logo,

$$
\begin{aligned}
Var(\widehat{W}_1) &= Var\left(\frac{1}{2}(\widehat{\mu}_1 + \widehat{\mu}_2)\right)\\
&= \frac{1}{4}(Var(\widehat{\mu}_1) + Var(\widehat{\mu}_2))\\
&= \frac{1}{4}(Var(\widehat{\mu}_1) + 25Var(\widehat{\mu}_1))\\
&= \frac{26}{4}Var(\widehat{\mu}_1) = 6,5Var(\widehat{\mu}_1),
\end{aligned}
$$

$$
\begin{aligned}
Var(\widehat{W}_2) &= Var\left(\frac{1}{5}(4\widehat{\mu}_1 + \widehat{\mu}_2)\right)\\
&= \frac{1}{25}(Var(4\widehat{\mu}_1) + Var(\widehat{\mu}_2))\\
&= \frac{1}{25}(16Var(\widehat{\mu}_1) + 25Var(\widehat{\mu}_2))\\
&= \frac{41}{25}Var(\widehat{\mu}_1) = 1,64Var(\widehat{\mu}_1),
\end{aligned}
$$

$$
\begin{aligned}
Var(\widehat{W}_3) &= Var\left(\frac{1}{6}(5\widehat{\mu}_1 + \widehat{\mu}_2)\right)\\
&= \frac{1}{36}(Var(5\widehat{\mu}_1) + Var(\widehat{\mu}_2))\\
&= \frac{1}{36}(25Var(\widehat{\mu}_1) + 25Var(\widehat{\mu}_1))\\
&= \frac{50}{36}Var(\widehat{\mu}_1) = 1,39Var(\widehat{\mu}_1),
\end{aligned}
$$

$$Var(\widehat{W}_4) = Var(\widehat{\mu}_1).$$

Portanto,

$$Var(\widehat{W}_4) < Var(\widehat{W}_3) < Var(\widehat{W}_2) < Var(\widehat{W}_1),$$

$\widehat{W}_4$ é o estimador mais eficiente.

(c) Sabemos que $E(a\widehat{\mu}_1 + (1-a)\widehat{\mu}_2) = \mu$ para qualquer valor de a, então o estimador mais eficiente será aquele que possui a menor variância. Assim, desejamos minimizar $Var(a\widehat{\mu}_1 + (1-a)\widehat{\mu}_2)$, lembrando que os estimadores são independentes, tem–se que,

$$\begin{aligned} Var(a\widehat{\mu}_1 + (1-a)\widehat{\mu}_2) &= Var(a\widehat{\mu}_1) + Var((1-a)\widehat{\mu}_2) \\ &= a^2Var(\widehat{\mu}_1) + (1-a)^2Var(\widehat{\mu}_2) \\ &= a^2Var(\widehat{\mu}_1) + (1-a)^2 25Var(\widehat{\mu}_1) \\ &= (a^2 + 25(1-a)^2)Var(\widehat{\mu}_1), \end{aligned}$$

assim,

$$\begin{aligned} \frac{\partial}{\partial a}Var(a\widehat{\mu}_1 + (1-a)\widehat{\mu}_2) &= \frac{\partial}{\partial a}(a^2 + 25(1-a)^2)Var(\widehat{\mu}_1) \\ &= Var(\widehat{\mu}_1)(2a + 50(1-a)(-1)) \\ &= Var(\widehat{\mu}_1)(2a - 50(1-a)) \\ &= Var(\widehat{\mu}_1)(52a - 50), \end{aligned}$$

$$Var(\widehat{\mu}_1)(52a - 50) = 0 \Rightarrow 52a - 50 = 0 \Rightarrow 52a = 50 \Rightarrow a = \frac{50}{52}$$

$$\frac{\partial^2}{\partial a^2}Var(a\widehat{\mu}_1 + (1-a)\widehat{\mu}_2) = Var(\widehat{\mu}_1)52 > 0,$$

logo, $a = \frac{50}{52}$ é ponto de mínimo. Portanto, $\frac{50}{52}\widehat{\mu}_1 + \left(1 - \frac{50}{52}\right)\widehat{\mu}_2 = \frac{50}{52}\widehat{\mu}_1 + \frac{2}{52}\widehat{\mu}_2$ é a combinação linear mais eficiente possível.

Exercício 6:

Uma variável aleatória tem distribuição Gama se a sua f.d.p. é

$$f(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, & x > 0, \alpha > 0 \text{ e } \lambda > 0 \\ 0, & c.c.; \end{cases}$$

onde α é o chamado parâmetro de forma e λ é o parâmetro de taxa e $\Gamma(\alpha) = \int_0^{+\infty} u^{\alpha-1} e^{-u} du, \alpha > 0$ é a função Gama.

Sejam X_i, $i = 1, \ldots, n$ as componentes de uma amostra aleatória desta distribuição representando o tempo entre duas ocorrências consecutivas de um abalo sísmico. Sabendo que $E(X) = \frac{\alpha}{\lambda}$ e $E(X^2) = \frac{\alpha(\alpha+1)}{\lambda^2}$, determine, pelo método dos momentos, os estimadores de α e λ.

Solução:

Seja $X =$“tempo entre duas ocorrências consecutivas de um abalo sísmico”,

$$E(X) = \frac{1}{n} \sum X_i \Rightarrow \frac{\alpha}{\lambda} = \overline{X} \Rightarrow \alpha = \lambda \overline{X},$$

$$\begin{aligned} E(X^2) = \frac{1}{n} \sum X_i^2 \Rightarrow \frac{\alpha(\alpha+1)}{\lambda^2} = \frac{1}{n} \sum X_i^2 \Rightarrow \frac{\lambda\overline{X}(\lambda\overline{X}+1)}{\lambda^2} &= \frac{1}{n} \sum X_i^2 \\ \Rightarrow \overline{X}(\lambda\overline{X}+1) = \frac{\lambda}{n} \sum X_i^2 \Rightarrow \lambda \overline{X}^2 - \frac{\lambda}{n} \sum X_i^2 &= -\overline{X} \\ \Rightarrow \widehat{\lambda} = \frac{-\overline{X}}{\overline{X}^2 - \frac{1}{n} \sum X_i^2} = \frac{\overline{X}}{\frac{1}{n} \sum X_i^2 - \overline{X}^2}, \end{aligned}$$

logo,

$$\widehat{\alpha} = \frac{\overline{X}^2}{\frac{1}{n} \sum X_i^2 - \overline{X}^2}.$$

Exercício 7:

Seja X uma v.a. com distribuição exponencial deslocada,

$$f(x) = \lambda \exp\{-\lambda(x-\tau)\} \text{ onde } x > \tau > 0 \text{ e } \lambda > 0.$$

Considere $X_1, X_2, \ldots, X_n$ uma amostra aleatória de tamanho n. Determine o estimador de máxima verossimilhança de λ, admitindo que τ é conhecido.

Solução:

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

$$\ell(\lambda) = \log L(\lambda) = n \log \lambda - \lambda \sum_{i=1}^{n}(x_i-\tau),$$

$$\frac{\partial \ell(\lambda)}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n}(x_i-\tau)$$

$$\begin{aligned} \frac{n}{\lambda} - \lambda\frac{\sum_{i=1}^{n}(x_i-\tau)}{\lambda} &= 0 \\ n &= \lambda \sum_{i=1}^{n}(x_i-\tau) \\ \widehat{\lambda} &= \frac{n}{\sum_{i=1}^{n}(x_i-\tau)}, \end{aligned}$$

$$\frac{\partial^2 \ell(\lambda)}{\partial \lambda^2} = -\frac{n}{\lambda^2} < 0,$$

logo, $\widehat{\lambda} = \frac{n}{\sum_{i=1}^{n}(x_i-\tau)}$ é ponto de máximo. Portanto, o estimador de máxima verossimilhança de λ é $\widehat{\lambda} = \frac{n}{\sum_{i=1}^{n}(X_i-\tau)}$. O estimador também pode ser escrito como $\widehat{\lambda} = \frac{1}{\overline{X}-\tau}$.