MAE0229 – Introdução à Probabilidade e à Estatística II

I semestre de 2018

Lista de exercícios 3 - Estimação II - C A S A (gabarito)

Exercício 1:

Sabe-se que o tempo de viagem de um local A na zona norte de São Paulo até a USP segue uma distribuição normal com desvio padrão 9 minutos. Em 200 dias anotou-se o tempo gasto para vir desse ponto A até à USP. A média amostral dos tempos foi de 45 minutos.

(a) Estime o tempo médio de viagem do ponto A até a USP através de um intervalo com coeficiente de confiança igual a 97%.

(b) Se formos estimar este tempo médio pelo intervalo $[45-0,6; 45+0,6]$ minutos, qual será o valor do coeficiente de confiança desta estimativa?

Solução:

Seja X :"Tempo de origem de um local A na zona norte de São Paulo até a USP". $X \sim N(\cdot; 9^2)$, do exercício $n = 200$, $\overline{x} = 45$ minutos.

(a) como o desvio padrão é conhecido, o intervalo de confiança para a média é dado por

$$IC(\mu, \gamma\%) = \left[\overline{x} - z\frac{\sigma}{\sqrt{n}} \ ; \ \overline{x} + z\frac{\sigma}{\sqrt{n}}\right],$$

com z é tal que, $P(-z \leq Z \leq z) = \gamma$. Assim, como $\gamma = 0,97$ então z é tal que $P(-z \leq Z \leq z) = 0,97$, logo, z é tal que $P(Z \leq z) = 0,985$, portanto $z = 2,17$. Logo,

$$\begin{aligned} IC(\mu, 97\%) &= \left[45 - 2,17\frac{9}{\sqrt{200}} \ ; \ 45 + 2,17\frac{9}{\sqrt{200}}\right] \\ &= [45 - 1,381 \ ; \ 45 + 1,381] \\ &= [43,619 \ ; \ 46,381] \end{aligned}$$

(b) Queremos saber o valor de γ, logo

$$IC(\mu, \gamma\%) = [45 - 0,6 \ ; \ 45 + 0,6],$$

então,

$$0,6 = z\frac{\sigma}{\sqrt{n}} = z\frac{9}{\sqrt{200}} \Rightarrow z = \frac{0,6}{9}\sqrt{200} = 0,943.$$

Lembrando que $\gamma = P(-z \leq Z \leq z)$, tem–se que

$$
\begin{aligned}
\gamma &= P(-0,943 \leq Z \leq 0,943) \\
&= P(Z \leq 0,943) - 1 + P(Z \leq 0,943) \\
&= 2P(Z \leq 0,943) - 1 \\
&= 2(0,8264) - 1 = 0,6528.
\end{aligned}
$$

O valor do coeficiente de confiança é de $65,28\%$. Uma outra forma de resolver é

$$z = 0,943 \text{ é tal que } P(Z \leq 0,943) = 0,8264 = \frac{1+\gamma}{2},$$

logo $\gamma = 0,6528$.

Exercício 2:

Em um estudo de subsídios de empréstimos para estudantes, o Departamento de Educação relatou que aqueles que tomam empréstimos da Stanford Loan com quatro anos de prazo, terão uma dívida média de $US\$12.168$ (USA Today, 5 de abril de 1995). Considere que essa quantia média de endividamento está baseada numa amostra de 480 empréstimos de estudantes, e que na graduação o desvio padrão da população para a quantia emprestada seja de $US\$2.200$.

(a) Construa intervalos de confiança a 90%, 95% e 99% para a quantia média devida pelos estudantes.

(b) Compare a amplitude dos intervalos e comente os resultados.

Solução:

Seja X : "dividida gerada pelo empréstimo da Stanford Loan com quatro anos de prazo". Note que não sabemos qual a distribuição de X, mas como n é grande e σ é conhecido, então $\overline{X} \sim N(\mu, \sigma^2/n)$. Do exercício tem–se que $\overline{x} = 12168$, $\sigma = 2200$ e $n = 480$.

(a) Como o desvio padrão é conhecido

$$IC(\mu; \gamma\%) = \left[\overline{X} - z\frac{\sigma}{\sqrt{n}} \ ; \ \overline{X} + z\frac{\sigma}{\sqrt{n}}\right],$$

com $\gamma = 0,9$ tem–se que $P(-z \leq Z \leq z) = 0,9$, logo, $P(Z \leq z) = 0,95$, então $z = 1,645$. Assim,

$$\begin{aligned} IC(\mu, 90\%) &= \left[12168 - 1,645\frac{2200}{\sqrt{480}} \ ; \ 12168 + 1,645\frac{2200}{\sqrt{480}}\right] \\ &= [12168 - 165,184 \ ; \ 12168 + 165,184] \\ &= [12002,82 \ ; \ 12333,18] \end{aligned}$$

Considerando $\gamma = 0,95$, tem–se que $P(-z \leq Z \leq z) = 0,95$, logo, $P(Z \leq z) = 0,975$, então $z = 1,96$. Assim,

$$\begin{aligned} IC(\mu, 95\%) &= \left[12168 - 1,96\frac{2200}{\sqrt{480}} \ ; \ 12168 + 1,96\frac{2200}{\sqrt{480}}\right] \\ &= [12168 - 196,815 \ ; \ 12168 + 196,815] \\ &= [11971,19 \ ; \ 12364,81] \end{aligned}$$

Considerando $\gamma = 0,99$, tem–se que $P(-z \leq Z \leq z) = 0,99$, logo, $P(Z \leq z) = 0,995$, então $z = 2,575$. Assim,

$$
\begin{aligned}
IC(\mu, 99\%) &= \left[12168 - 2,575\frac{2200}{\sqrt{480}}\ ;\ 12168 + 2,575\frac{2200}{\sqrt{480}}\right] \\
&= [12168 - 258,5707\ ;\ 12168 + 258,5707] \\
&= [11909,43\ ;\ 12426,57]
\end{aligned}
$$

(b) A amplitude dos intervalos pode ser calculada como $LS - LI$ ou $2z\sigma/\sqrt{n}$.

1. $\gamma = 90\%, \Rightarrow LS - LI = 12333,18 - 12002,82 = 330,36,$
2. $\gamma = 95\%, \Rightarrow LS - LI = 12364,81 - 11971,19 = 393,62,$
3. $\gamma = 99\%, \Rightarrow LS - LI = 12426,57 - 11909,43 = 517,14.$

A medida que o coeficiente de confiança aumenta também aumenta a amplitude do intervalo de confiança.

Exercício 3:

O nível de poluição do ar de determinada cidade (medido em concentração de monóxido de carbono no ar) distribui-se normalmente. Recolheram-se os seguintes valores da referida concentração em 10 dias diferentes (em ppm): 0.09, 0.33, 0.01, 0.25, 0.20, 0.05, 0.03, 0.18, 0.13, 0.24. Com base nesta amostra determine um intervalo de confiança a 99% para a concentração média de monóxido de carbono na atmosfera.

Solução:

Seja X : "nível de poluição do ar de determinada cidade (medida em concentração de monóxido de carbono no ar)", $X \sim N(\mu, \sigma^2)$. Dos dados da amostra se calcula

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{10} x_i = \frac{0,09 + \cdots + 0,24}{10} = \frac{1,51}{10} = 0,151,$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{10}(x_i - \overline{x})^2 = \frac{(0,09-0,151)^2 + \cdots + (0,24-0,151)^2}{9} = \frac{0,10189}{9} = 0,011,$$

$$s = \sqrt{s^2} = \sqrt{0,011} = 0,106.$$

Como o desvio padrão não é conhecido e foi estimado dos dados e n é pequeno

$$IC(\mu, \gamma\%) = \left[\overline{X} - t\frac{s}{\sqrt{n}}\ ;\ \overline{X} + t\frac{s}{\sqrt{n}}\right],$$

em que $\gamma = P(-t \leq T \leq t)$ com T tendo distribuição t–Student com $n-1$ graus de liberdade, como $\gamma = 0,99$, então t é tal que $P(T \leq t) = 0,995$, ver tabela da t–student do livro do Morettin: $p = 0,01$ e portanto, para 9 graus de liberdade obtemos $t = 3,250$. Logo,

$$\begin{aligned} IC(\mu, 99\%) &= \left[0,151 - 3,250\frac{0,106}{\sqrt{10}}\ ;\ 0,151 + 3,250\frac{0,106}{\sqrt{10}}\right] \\ &= [0,151 - 0,109\ ;\ 0,151 + 0,109] \\ &= [0,042\ ;\ 0,260] \end{aligned}$$

Exercício 4:

Suponha que a voltagem que um cabo elétrico com um certo isolamento pode suportar varia de acordo com uma distribuição Normal. Para uma amostra de 12 cabos as falhas ocorreram nos seguintes níveis de voltagem:

52 64 38 68 66 52 60 44 48 46 70 62

(a) Construa um intervalo de confiança a 95% para o valor esperado da voltagem que um cabo elétrico pode suportar.

(b) Construa um intervalo de confiança a 95% para a variância da voltagem que um cabo elétrico pode suportar.

Solução:

Seja X : "Voltagem que um cabo elétrico com um certo isolamento pode suportar", $X \sim N(\mu, \sigma^2)$. Do exercício tem-se que $n = 12$, e calcula–se

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{12} x_i = \frac{52 + \cdots + 62}{12} = \frac{670}{12} = 55,833$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{12}(x_i - \overline{x})^2 = \frac{(52 - 55,833)^2 + \cdots + (62 - 55,833)^2}{11} = \frac{1219,667}{11} = 110,8788$$

$$s = \sqrt{s^2} = 10,5299$$

(a) como o desvio padrão não é conhecido e é estimado da amostra e n é pequeno

$$IC(\mu, \gamma\%) = \left[\overline{x} - t\frac{s}{\sqrt{n}}\ ;\ \overline{x} + t\frac{s}{\sqrt{n}}\right],$$

em que $\gamma = P(-t \leq T \leq t)$ com T tendo distribuição t–Student com $n-1$ graus de liberdade, como $\gamma = 0,95$, então t é tal que $P(T \leq t) = 0,975$, ver tabela da t–student do livro do Morettin: $p = 0,05$ e portanto, para 11 graus de liberdade obtemos $t = 2,201$. Logo,

$$\begin{aligned} IC(\mu, 95\%) &= \left[55,833 - 2,201\frac{10,5299}{\sqrt{12}}\ ;\ 55,833 + 2,201\frac{10,5299}{\sqrt{12}}\right] \\ &= [55,833 - 6,690\ ;\ 55,833 + 6,690] \\ &= [49,143\ ;\ 62,523] \end{aligned}$$

(b)

$$
\begin{aligned}
IC(\sigma^2, 95\%) &= \left[\frac{(n-1)s^2}{\chi^2_2}\ ;\ \frac{(n-1)s^2}{\chi^2_1}\right] \\
&= \left[\frac{11(110,8788)}{21,920}\ ;\ \frac{11(110,8788)}{3,816}\right] \\
&= [55,642\ ;\ 319,640]
\end{aligned}
$$

Os valores de χ^2_2 e χ^2_1 podem ser encontrados no livro do Morettin.

Exercício 5:

Um profissional de boliche jogou 8 partidas num torneio, tendo obtido as seguintes pontuações:

117.0 220.2 199.5 237.2 249.5 179.8 259.2 248.5

Admitindo a normalidade das pontuações, construa um intervalo de confiança a 95% para a variância e para o desvio padrão (este último fornece uma medida da consistência da prestação do jogador).

Solução:

Seja X : "Pontuação obtida num jogo de boliche", $X \sim N(\mu, \sigma^2)$. Do exercício se calcula primeiro

$$\overline{x} = \frac{117,0 + \cdots + 248,5}{8} = \frac{1710,9}{8} = 213,8625,$$

$$s^2 = \frac{(117,0 - 213,8625)^2 + \cdots + (248,5 - 213,8625)^2}{7} = \frac{15858,96}{7} = 2265,566,$$

assim

$$\begin{aligned} IC(\sigma^2, 95\%) &= \left[\frac{(n-1)s^2}{\chi_2^2} \ ; \ \frac{(n-1)s^2}{\chi_1^2}\right] \\ &= \left[\frac{7(2265,566)}{16,013} \ ; \ \frac{7(2265,566)}{1,690}\right] \\ &= [990,3953 \ ; \ 9384,7290] \end{aligned}$$

e

$$\begin{aligned} IC(\sigma, 95\%) &= \left[\sqrt{\frac{(n-1)s^2}{\chi_2^2}} \ ; \ \sqrt{\frac{(n-1)s^2}{\chi_1^2}}\right] \\ &= \left[\sqrt{990,3953} \ ; \ \sqrt{9384,7290}\right] \\ &= [31,471 \ ; \ 96,875] \end{aligned}$$

Os valores de χ_2^2 e χ_1^2 podem ser encontrados no livro do Morettin.

Exercício 6:

Um determinado banco deseja ter informação sobre o tempo de utilização de seus caixas eletrônicos pelos clientes, em determinada região, nos fins de semana. Mais especificamente, deseja estimar a proporção p de usuários dessa região, que demoram 2 minutos ou mais para realizarem suas operações. Uma amostra aleatória de clientes que utilizam caixas eletrônicos em fins de semana nessa região será coletada, e o tempo de utilização de cada um será registrado.

(a) Qual deve ser o tamanho da amostra, para que o erro de sua estimativa seja no máximo 0,08 com um grau de confiança de 0,85?

(b) A direção do banco sabe que, nas condições descritas, essa proporção p não ultrapassa 15%. Com essa informação seria possível considerar em (a) uma amostra de tamanho menor? Se sim, de quanto? Se não, por quê?

(c) Uma amostra de 80 clientes forneceu as seguintes medidas desse tempo (em minutos): 1,2 1,2 1,1 1,3 1,5 0,9 2,0 1,3 1,4 1,6 1,6 1,3 2,2 1,6 1,0 0,8 1,5 2,3 1,7 1,6 2,4 1,2 1,2 1,0 0,9 2,2 1,7 1,5 1,3 1,2 1,9 0,9 1,3 1,3 1,8 1,3 2,7 1,4 0,9 1,2 1,3 2,4 2,1 1,0 1,0 1,1 1,6 1,3 1,1 1,9 1,1 2,2 2,1 1,7 1,5 0,9 2,0 1,1 1,4 1,6 1,4 1,7 2,3 1,6 1,0 0,8 1,5 1,3 2,7 1,2 0,9 1,2 1,3 1,4 2,8 1,0 1,0 1,1 1,6 1,3.

(d) Dê uma estimativa pontual para p e, com base nela, construa intervalos de 90% de confiança, otimista e conservador, para p. Qual é o erro amostral de sua estimativa?

Solução:

Seja p : "proporção de usuários de uma determinada regiaõ que demoram 2 minutos ou mais para relaizarem suas operações".

(a) Lembre que

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

do exercício tem–se que, $\epsilon \leq 0,08$, $\gamma = 0,85$, logo, $0,85 = P(-z \leq Z \leq z) \Rightarrow P(Z \leq z) = 0,925 \Rightarrow z = 1,44$. Como não se tem informação sobre a variância se considera a maior possível, isto é quando $p = 1/2$, logo, $p(1-p) = 1/4$. Assim

$$n = \left(\frac{1,44}{0,08}\right)^2 \frac{1}{4} = 81,$$

se deve tomar uma amostra de tamanho 81 para que o erro de sua estimativa seja no máximo 0,08 com um grau de confiança de 0,85.

(b) Sabe–se que $p \leq 0,15$, portanto

$$n = \left(\frac{z}{\epsilon}\right)^2 \sigma^2 = \left(\frac{1,44}{0,08}\right)^2 0,15(1-0,15) = \left(\frac{1,44}{0,08}\right)^2 0,1275 = 41,31 \approx 42,$$

é possível considerar uma amostra de tamanho menor, neste caso de tamanho 42, pois com a informação adicional sobre p também se tem informação sobre a variância, sendo ela menor do que a considerada no item (a).

(c) Primeiro se devem converter os dados para saber se a pessoa demora 2 minutos ou mais realizando as operações.

0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 1 0 0 1 0 0 0 0 1 0
0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 1 0 0 0 0 0 0 0 0 1 1 0
0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0

Assim

$$\widehat{p} = \frac{14}{80} = 0,175,$$

considerando um graude confiança γ o intervalo de confiança otimista é

$$IC(p, \gamma\%) = \left[\widehat{p} - z\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}} \ ; \ \widehat{p} + z\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}}\right],$$

como $\gamma = 0,9 = P(-z \leq Z \leq z)$, então z é tal que $P(Z \leq z) = 0,95$, logo, $z = 1,645$.

$$\begin{aligned} IC(p, 90\%) &= \left[0,175 - 1,645\sqrt{\frac{0,175(1-0,175)}{80}} \ ; \ 0,175 + 1,645\sqrt{\frac{0,175(1-0,175)}{80}}\right] \\ &= [0,175 - 0,070 \ ; \ 0,175 + 0,070] \\ &= [0,105 \ ; \ 0,245] \end{aligned}$$

Considerando $\widehat{p} = 1/2$ o intervalo de confiança conservador

$$IC(p, \gamma\%) = \left[\widehat{p} - z\sqrt{\frac{1}{4n}} \ ; \ \widehat{p} + z\sqrt{\frac{1}{4n}}\right],$$

logo

$$
\begin{aligned}
IC(p, 90\%) &= \left[0,175 - 1,645\sqrt{\frac{1}{4 \times 80}} \ ; \ 0,75 + 1,645\sqrt{\frac{1}{4 \times 80}}\right] \\
&= [0,175 - 0,092 \ ; \ 0,175 + 0,092] \\
&= [0,083 \ ; \ 0,267]
\end{aligned}
$$

O erro amostral ϵ é

$$\epsilon = z\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}}$$

no caso do intervalo de confiança otimista é

$$\epsilon = 1,645\sqrt{\frac{0,175(1-0,175)}{80}} = 0,070$$

no caso do intervalo de confiança conservador é

$$\epsilon = 1,645\sqrt{\frac{1}{4 \times 80}} = 0,092.$$

Exercício 7:

Realizou-se uma pesquisa de opinião via telefone para estimar a proporção da população de um país favorável a uma reforma fiscal.

(a) Determine o tamanho da amostra que garanta, com um grau de confiança de 95%, que o erro máximo cometido seja inferior a 0,01.

(b) Supondo que, em uma amostra de 50 pessoas, 8 pessoas se manifestaram desfavoráveis à reforma fiscal, determine um intervalo de confiança a 95% para a proporção da população favorável à reforma fiscal.

(c) Mantendo as condições da alínea anterior, diga quantas pessoas teriam de ser ouvidas para reduzir para metade a amplitude do intervalo obtido anteriormente.

Solução:

Seja p : "proporção da população de um país favorável a uma reforma fiscal".

(a) Tem–se que $\epsilon \leq 0,01$ e $\gamma = 0,95 = P(-z \leq Z \leq z)$, então z é tal que $P(Z \leq z) = 0,975$, $z = 1,96$. Como não se tem informação sobre a variância se considera a maior possível, isto é quando $p = 1/2$.

$$n = \left(\frac{z}{\epsilon}\right)^2 p(1-p) = \left(\frac{1,96}{0,01}\right)^2 \frac{1}{4} = 9604$$

(b) considerando $n = 50$ e sendo que 8 pessoas se manifestaram desfavoráveis, tem–se que

$$\widehat{p} = \frac{42}{50} = 0,84,$$

como $\gamma = 0,95$, $z = 1,96$ pelo item (a). O intervalo de confiança otimista é

$$\begin{aligned} IC(p, 95\%) &= \left[\widehat{p} - z\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}}\ ;\ \widehat{p} + z\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}}\right] \\ &= \left[0,84 - 1,96\sqrt{\frac{0,84(1-0,84)}{50}}\ ;\ 0,84 + 1,96\sqrt{\frac{0,84(1-0,84)}{50}}\right] \\ &= [0,84 - 0,102\ ;\ 0,84 + 0,102] \\ &= [0,738\ ;\ 0,942] \end{aligned}$$

O intervalo de confiança conservador é

$$
\begin{aligned}
IC(p, 95\%) &= \left[\widehat{p} - z\sqrt{\frac{1}{4 \times n}}\ ;\ \widehat{p} + z\sqrt{\frac{1}{4 \times n}}\right] \\
&= \left[0,84 - 1,96\sqrt{\frac{1}{200}}\ ;\ 0,84 + 1,96\sqrt{\frac{1}{200}}\right] \\
&= [0,84 - 0,139\ ;\ 0,84 + 0,139] \\
&= [0,701\ ;\ 0,979]
\end{aligned}
$$

(c) A amplitude do intervalo no item (b) no caso do intervalo de confiança otimista é

$$2z\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}} = 0,204,$$

como se quer reduzir para metade, se quer

$$2z\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}} = 0,102,$$

logo

$$2 \times 1,96\sqrt{\frac{0,84(1-0,84)}{n}} = 0,102 \Rightarrow n = \frac{4(1,96)^2(0,84(1-0,84))}{(0,102)^2} = 198,505 \approx 199$$

se devem ouvir 199 pessoas para reduzir para metade a amplitude do intervalo de confiança otimista obtido anteriormente.

A amplitude do intervalo no item (b) no caso do intervalo de confiança conservador é

$$2z\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}} = 0,278,$$

como se quer reduzir para metade, se quer

$$2z\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}} = 0,139,$$

logo

$$2 \times 1,96\sqrt{\frac{1}{4n}} = 0,139 \Rightarrow n = \frac{4(1,96)^2}{4(0,139)^2} = 198,830 \approx 199$$

se devem ouvir 199 pessoas para reduzir para metade a amplitude do intervalo de confiança conservador obtido anteriormente. Neste caso coincide com o obtido para o intervalo de confiança otimista.