Capítulo 6 - Estimação por intervalo de confiança

Resolução de alguns exercícios

Estimação por intervalo de confiança

6.1 a) Uma vez que a população segue uma distribuição Normal de valor médio μ desconhecido e variância σ^2 conhecida, vamos considerar variável pivot:

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1).$$

Para a amostra observada e com base nas seguintes informações:

- informação populacional: $\sigma = 20$ g
- informação amostral: $n = 20$, $\overline{x} = 320$ g
- nível de confiança $1 - \alpha = 0.90 \Leftrightarrow \alpha = 0.10$, pelo que $z_{\alpha/2} = z_{0.05} = \Phi^{-1}(0.95) \simeq 1.64$

obtemos

$$\begin{aligned} IC_{90\%}(\mu) &\simeq \left]320 - 1.64\frac{20}{\sqrt{20}}, 320 + 1.64\frac{20}{\sqrt{20}}\right[\\ &\simeq \left]312.67, 327.33\right[. \end{aligned}$$

Caso tivéssemos usado a aproximação $z_{0.05} \simeq 1.645$ viria

$$IC_{90\%}(\mu) \simeq \left]312.64, 327.36\right[.$$

6.1 b) A amplitude do intervalo de confiança, Δ, é dada por

$$\Delta = 2 \times z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}}.$$

Pretendemos garantir que a amplitude do intervalo é no máximo 1, assim

$$2 \times 1.64 \frac{20}{\sqrt{n}} \leq 1 \Leftrightarrow n \geq (2 \times 1.64 \times 20)^2 \quad (\simeq 4303.36)$$

pelo que, sendo $n \in \mathbb{N}$, se tem finalmente que $n \geq 4304$. Para uma amplitude no máximo de 5 teremos que ter $n \geq 173$.

Caso tivéssemos usado a aproximação $z_{0.05} \simeq 1.645$ viria $n \geq 4330$ no primeiro caso e $n \geq 174$ no segundo caso.

6.2 a) Uma vez que a população segue uma distribuição Normal de valor médio μ desconhecido e variância σ^2 desconhecida vamos considerar a variável pivot:

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}.$$

Assim, com base nas seguintes informações:

- informação amostral: $n = 8$, $s = 10$, $\overline{x} = 36$
- nível de confiança: $1 - \alpha = 0.90 \Leftrightarrow \alpha = 0.1 \Leftrightarrow \alpha/2 = 0.05$, pelo que $t_{n-1,\alpha/2} = t_{7,0.05} \simeq 1.895$

obtemos

$$IC_{90\%}(\mu) \simeq \left]36 - 1.895\frac{10}{\sqrt{8}}, 36 + 1.895\frac{10}{\sqrt{8}}\right[\simeq]29.300, 42.700[\,.$$

(usando a aproximação $t_{7,0.05} \simeq 1.89$ viria $IC_{90\%}(\mu) \simeq]29.318, 42.682[$)

De forma análoga, para um nível de confiança $1 - \alpha = 0.95$ temos $\alpha = 0.05$ e $t_{n-1,\alpha/2} = t_{7,0.025} \approx 2.365$, pelo que

$$IC_{95\%}(\mu) \simeq \left]36 - 2.365\frac{10}{\sqrt{8}}, 36 + 2.365\frac{10}{\sqrt{8}}\right[\simeq]27.638, 44.362[\,.$$

(usando a aproximação $t_{7,0.05} \simeq 1.89$ viria $IC_{95\%}(\mu) \simeq]27.656, 44.344[$)

Observamos que $IC_{90\%}(\mu) \subseteq IC_{95\%}(\mu)$, o que resulta directamente da própria construção dos ICs.

6.2 b) Nesta alínea a única alteração é a dimensão da amostra que passa a ser 50 mas continuamos a ter uma população segue uma distribuição Normal de valor médio μ desconhecido e variância σ^2 desconhecida, pelo que a variável pivot e a respetiva distribuição são

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}.$$

Assim, a resolução pode ser feita de forma semelhante à da alínea a)

- para um nível de confiança $1 - \alpha = 0.90$ temos $\alpha = 0.1$ e $t_{n-1,\alpha/2} = t_{49,0.05} \simeq 1.677$, pelo que

$$\begin{aligned} IC_{90\%}(\mu) &\simeq \left]36 - 1.677\frac{10}{\sqrt{50}}, 36 + 1.677\frac{10}{\sqrt{50}}\right[\\ &\simeq \left]33.628, 38.372\right[. \end{aligned}$$

- para um nível de confiança $1 - \alpha = 0.95$ temos $\alpha = 0.05$ e $t_{n-1,\alpha/2} = t_{49,0.025} \simeq 2.010$, pelo

$$\begin{aligned} IC_{95\%}(\mu) &\simeq \left]36 - 2.010\frac{10}{\sqrt{50}}, 36 + 2.010\frac{10}{\sqrt{50}}\right[\\ &\simeq \left]33.157, 38.843\right[. \end{aligned}$$

6.2 b) (continuação)

Observamos que

- aumentando o tamanho amostral é então possível aumentar também a precisão dos intervalos de confiança;

- a dimensão da amostra é maior que 30 pelo que se não soubéssemos o valor dos quantis da distribuição t poderíamos ter aproximado a distribuição exata pela distribuição Normal considerando assim, na resolução do exercício, a variável pivot

$$Z = \frac{\overline{X} - \mu}{S/\sqrt{n}} \overset{a}{\sim} N(0,1).$$

6.5 População com distribuição desconhecida com valor médio μ e variância σ^2 desconhecidos. Dado que $n = 35 \geq 30$ podemos considerar a variável pivot:

$$Z = \frac{\overline{X} - \mu}{S/\sqrt{n}} \overset{a}{\sim} N(0,1).$$

Assim, para a amostra observada e com base nas seguintes informações:

- informação amostral: $n = 35$, $\overline{x} = 22.1°C$ e $s = 3.2°C$
- nível de confiança $1 - \alpha = 0.95 \Leftrightarrow \alpha = 0.05$, pelo que $z_{\alpha/2} = z_{0.025} = \Phi^{-1}(0.975) \simeq 1.96$

obtemos

$$\begin{aligned} IC_{95\%}(\mu) &\simeq \left] 22.1 - 1.96\frac{3.2}{\sqrt{35}}, 22.1 + 1.96\frac{3.2}{\sqrt{35}} \right[\\ &\simeq \left] 21.04, 23.16 \right[. \end{aligned}$$

6.7 a) A estimativa pontual é $\overline{x} = \dfrac{1}{40}\sum_{i=1}^{40} x_i = 1.73\ m$.

b) Uma vez que a população segue uma distribuição Normal de valor médio μ desconhecido e variância σ^2 desconhecida vamos considerar a variável pivot:

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}.$$

Assim, para a amostra observada e com base nas seguintes informações:

- informação amostral: $n = 40$,

$$\overline{x} = \frac{1}{40}\sum_{i=1}^{40} x_i = 1.73 \quad \text{e} \quad s = \sqrt{\frac{1}{39}\sum_{i=1}^{40}(x_i - \overline{x})^2} = 0.08$$

- nível de confiança $1 - \alpha = 0.92 \Leftrightarrow \alpha = 0.08$, pelo que $t_{n-1,\alpha/2} = t_{39,0.04} \simeq 1.797$ (Nota: este valor não está na tabela da distribuição t mas pode ser obtido facilmente com o software R)

obtemos

$$\begin{aligned} IC_{92\%}(\mu) &\simeq \left]1.73 - 1.797\frac{0.08}{\sqrt{40}}, 1.73 + 1.797\frac{0.08}{\sqrt{40}}\right[\\ &\simeq \left]1.707, 1.753\right[. \end{aligned}$$

6.9 Pretende-se construir um IC a 98% para verdadeira proporção de pacotes inadequados na produção total, p.

Usando a variável pivot

$$Z = \frac{\hat{P} - p}{\sqrt{\hat{P}(1-\hat{P})/n}} \overset{a}{\sim} N(0,1)$$

deduzimos então

$$IAC_{(1-\alpha)\times 100\%}(p) \underset{a}{\simeq} \left]\hat{P} - z_{\alpha/2}\sqrt{\hat{P}(1-\hat{P})/n}\ ;\ \hat{P} + z_{\alpha/2}\sqrt{\hat{P}(1-\hat{P})/n}\right[.$$

Dado que se tem

- informação amostral: $n = 100$, $\hat{p} = 18/100$
- nível de confiança: $1-\alpha = 0.98 \Leftrightarrow \alpha = 0.02$, vem $z_{\alpha/2} = z_{0.01} \simeq 2.33$

vem finalmente

$$IC_{98\%}(p) \underset{a}{\simeq} \left]\frac{18}{100} - 2.33\sqrt{\frac{18}{100}\left(1-\frac{18}{100}\right)/100}\,,\ \frac{18}{100} + 2.33\sqrt{\frac{18}{100}\left(1-\frac{18}{100}\right)/100}\right[$$

$$\simeq\ ]0.09, 0.27[$$

Estimação por intervalo de confiança

6.10 a) A estimativa pontual para p é dada por $\hat{p} = 12/200 = 0.06$

b) Pretende-se determinar que dimensão de amostra deve ser considerado para que a amplitude seja inferior a 0.01. Sabe-se que, para uma dada amostra, a estimativa por intervalo de confiança para p é dada por

$$IC_{90\%}(p) \underset{a}{\simeq} \left]\hat{p} - z_{0.05}\ \sqrt{\hat{p}(1-\hat{p})/n}\ ,\ \hat{p} + z_{0.05}\ \sqrt{\hat{p}(1-\hat{p})/n}\right[$$

sendo a amplitude deste IC dada por

$$\Delta = 2 \times z_{0.05}\sqrt{\hat{p}(1-\hat{p})/n}$$

e onde notamos que a estimativa pontual $\hat{p}$ também depende de n. Neste caso deve considerar-se que a estimativa determinada na alínea anterior $\hat{p} = 0.06$ é uma estimativa fiável do verdadeiro valor de p resolvendo portanto a inequação

$$\begin{aligned} & 2 \times 1.64 \times \sqrt{0.06(1-0.06)/n} < 0.01 \quad \Rightarrow \\ & \Rightarrow \quad n > (2 \times 1.64 \times \sqrt{0.06(1-0.06)}/0.01)^2 \Leftrightarrow \quad n > 6067.7376 \\ & \Rightarrow \quad n \underset{n \in \mathbb{N}}{\geq} 6068 \end{aligned}$$

($n \geq 6105$ no caso de ter usado $z_{0.05} \simeq 1.645$)

6.11 Pretende-se construir um IC a 99% para a variância do tempo de espera. Sendo a população em causa uma **população Normal** com variância σ^2 **desconhecida** a variável pivot que nos permite construir esse IC é

$$X^2 = \frac{(n-1)\,S^2}{\sigma^2} \sim \chi^2_{n-1}$$

vindo o intervalo aleatório de confiança a $(1-\alpha)\times 100\%$ para σ^2 dado por

$$IAC_{(1-\alpha)100\%}(\sigma^2) = \left]\frac{(n-1)\,S^2}{\chi^2_{n-1,\alpha/2}}, \frac{(n-1)\,S^2}{\chi^2_{n-1,1-\alpha/2}}\right[.$$

Dado que se tem:

- informação amostral: $n = 15,\ s^2 = 0.167$
- nível de confiança: $1-\alpha = 0.99 \Leftrightarrow \alpha = 0.01 \Leftrightarrow \frac{\alpha}{2} = 0.005$ vindo
  - $\chi^2_{n-1,1-\alpha/2} = \chi^2_{14,0.995} \underset{tabela}{\simeq} 4.075$
  - $\chi^2_{n-1,\alpha/2} = \chi^2_{14,0.005} \underset{tabela}{\simeq} 31.319$

(cont.) obtemos finalmente

$$IC_{99\%}(\sigma^2) = \left] \frac{(n-1)\, s^2}{\chi^2_{n-1,\alpha/2}}, \frac{(n-1)\, s^2}{\chi^2_{n-1,1-\alpha/2}} \right[\simeq \left]0.075, 0.574\right[\, .$$

6.12 Pretende-se construir um intervalo de confiança a 95% para a variância e para o desvio padrão populacionais. Sendo a população em causa uma **população Normal** com variância σ^2 **desconhecida** a variável pivot que nos permite construir esse IC é

$$X^2 = \frac{(n-1)\, S^2}{\sigma^2} \sim \chi^2_{n-1}$$

vindo o intervalo aleatório de confiança a $(1-\alpha) \times 100\%$ para σ^2 dado por

$$IAC_{(1-\alpha)100\%} = \left] \frac{(n-1)\, S^2}{\chi^2_{n-1,\alpha/2}}, \frac{(n-1)\, S^2}{\chi^2_{n-1,1-\alpha/2}} \right[.$$

Dado que se tem:

- informação amostral: $n = 8$, $s^2 = 2265.566$
- nível de confiança: $1-\alpha = 0.95 \Leftrightarrow \alpha = 0.05 \Leftrightarrow \frac{\alpha}{2} = 0.025$ vindo
  - $\chi^2_{n-1,1-\alpha/2} = \chi^2_{7,0.975} \underset{tabela}{\simeq} 1.690$
  - $\chi^2_{n-1,\alpha/2} = \chi^2_{7,0.025} \underset{tabela}{\simeq} 16.013$

(cont.) obtemos finalmente

$$IC_{95\%}(\sigma^2) = \left] \frac{(n-1)\, s^2}{\chi^2_{n-1,\alpha/2}}, \frac{(n-1)\, s^2}{\chi^2_{n-1,1-\alpha/2}} \right[\simeq \,]990.380, 9384.001[$$

e

$$IC_{95\%}(\sigma) = \left] \sqrt{\frac{(n-1)\, s^2}{\chi^2_{n-1,\alpha/2}}}, \sqrt{\frac{(n-1)\, s^2}{\chi^2_{n-1,1-\alpha/2}}} \right[\simeq \,]31.470, 96.871[\, .$$