6ª Lista de PE – Solução

1. a) Sabemos que

$$\sigma^2 = \text{Var}(X_i) = \mathcal{E}\left(X_i^2\right) - [\mathcal{E}(X_i)]^2 = \mathcal{E}\left(X_i^2\right) - \mu^2$$

portanto $\mathcal{E}\left(X_i^2\right) = \mu^2 + \sigma^2$.

Analogamente,

$$\frac{\sigma^2}{n} = \text{Var}\left(\overline{X}\right) = \mathcal{E}\left(\overline{X}^2\right) - \left[\mathcal{E}\left(\overline{X}\right)\right]^2 = \mathcal{E}\left(\overline{X}^2\right) - \mu^2$$

portanto $\mathcal{E}\left(\overline{X}^2\right) = \mu^2 + \dfrac{\sigma^2}{n}$.

b)

$$\begin{aligned}
\mathcal{E}(X_i\overline{X}) &= \frac{1}{n}\mathcal{E}(X_iX_1 + X_2X_i + \cdots + X_i^2 + \cdots + X_iX_n) \\
&= \frac{1}{n}\left\{\mathcal{E}(X_iX_1) + \cdots + \mathcal{E}(X_i^2) + \cdots + \mathcal{E}(X_iX_n)\right\} \\
&= \frac{1}{n}\left\{\mathcal{E}(X_i)\mathcal{E}(X_1) + \cdots + \mu^2 + \sigma^2 + \cdots + \mathcal{E}(X_i)\mathcal{E}(X_n)\right\} \\
&= \mu^2 + \frac{\sigma^2}{n}.
\end{aligned}$$

c) Para saber se um estimador é viciado calculamos sua esperança:

$$\begin{aligned}
\mathcal{E}(T) &= \frac{1}{n}\sum_{i=1}^{n}\mathcal{E}\left(X_i - \overline{X}\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\mathcal{E}\left(X_i^2 - 2X_i\overline{X} + \overline{X}^2\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}\left\{\mathcal{E}\left(X_i^2\right) - 2\mathcal{E}\left(X_i\overline{X}\right) + \mathcal{E}\left(\overline{X}^2\right)\right\} \\
&= \frac{1}{n}\sum_{i=1}^{n}\left\{\mu^2 + \sigma^2 - 2\left(\mu^2 + \frac{\sigma^2}{n}\right) + \mu^2 + \frac{\sigma^2}{n}\right\} \\
&= \sigma^2 - \frac{\sigma^2}{n} \\
&= \sigma^2\left(\frac{n-1}{n}\right).
\end{aligned}$$

Como $\mathcal{E}(T) \neq \sigma^2$, o estimador T é viciado para a variância.

d) Para que um estimador seja não viciado para a variância, sua esperança deve ser σ^2. Observe que se passarmos o termo $(n-1)/n$ para o outro lado da igualdade obtida no **item c)** multiplicando, teremos

$$\frac{n}{n-1} \times \mathcal{E}\left(T\right) = \mathcal{E}\left(\frac{n}{n-1} \times T\right) = \sigma^2.$$

Então basta definirmos o estimador

$$S^2 = \frac{n}{n-1} \times T = \sum_{i=1}^{n} \frac{\left(X_i - \overline{X}\right)^2}{n-1},$$

e S^2 é não viciado para a variância.

É importante notar que essa é razão para usarmos S^2 como variância amostral ao invés de usarmos T.

2. Primeiramente, lembre-se que

$$\rho = \frac{\mathcal{E}\left[\left(X-\mu_x\right)\left(Y-\mu_y\right)\right]}{\sigma(X)\sigma(Y)}.$$

Como a correlação envolve duas AAS, $(X_1, \ldots, X_n)$ e $(Y_1, \ldots, Y_n)$, um estimador natural do parâmetro ρ seria

$$\hat{\rho} = \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\sqrt{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 \sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2}}.$$

3. a) Como sabemos, no modelo de Poisson, $\mathcal{E}(X) = \mathrm{Var}(X) = \lambda$. Portanto, pelo método dos momentos, podemos estimar λ por meio dos dois estimadores, já encontrados em sala, abaixo:

$$\hat{\lambda}_1 = \frac{\sum_{i=1}^{n} X_i}{n} \qquad \text{e} \qquad \hat{\lambda}_2 = \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{n}.$$

b) Usando a amostra coletada, os valores obtidos para os estimadores são

$$\hat{\lambda}_1 = 3,17 \qquad \text{e} \qquad \hat{\lambda}_2 = 9,07.$$

4. Como $X_i \sim \mathrm{Exp}(\lambda)$, para $0 < x_i < \infty$, $i = 1, \ldots, n$, a verossimilhança do parâmetro λ será dada por

$$\begin{aligned} L(\lambda; x) &= f_{X_1}(x_1) \cdots f_{X_n}(x_n) \\ &= \lambda e^{-\lambda x_1} \cdots \lambda e^{-\lambda x_n} \\ &= \lambda^n \exp\left\{-\lambda \sum_{i=1}^{n} x_i\right\}. \end{aligned}$$

Portanto a logverossimilhança é dada por

$$\begin{aligned} l(\lambda;x) &= \log\lambda^n + \log\left[\exp\left\{-\lambda\sum_{i=1}^n x_i\right\}\right] \\ &= n\log\lambda - \lambda\sum_{i=1}^n x_i. \end{aligned}$$

Derivando a logverossimilhança temos que

$$l'(\lambda;x) = \frac{n}{\lambda} - \sum_{i=1}^n x_i = 0 \iff \lambda = \frac{n}{\sum_{i=1}^n x_i} = \frac{1}{\overline{x}}.$$

Como

$$l''(\lambda;x) = -\frac{n}{\lambda^2} < 0$$

concluímos que

$$\hat{\lambda} = \frac{1}{\overline{X}}.$$

5. Usando os dados e o estimador encontrado na **Questão 4** temos o valor $\hat{\lambda} = 1/3$. Desse modo, segue que

$$\mathcal{P}(X > 6) = 1 - F(6) = e^{\left(-\frac{1}{3}\right)(6)} = e^{-2} \simeq 0,1362.$$

6. Como $X_i \sim \text{Pois}(\lambda)$, para $0 < x_i < \infty$, $i = 1, \ldots, n$, a verossimilhança do parâmetro λ será dada por

$$\begin{aligned} L(\lambda;x) &= f_{X_1}(x_1)\cdots f_{X_n}(x_n) \\ &= \frac{e^{-\lambda}\lambda^{x_1}}{x_1!}\cdots\frac{e^{-\lambda}\lambda^{x_n}}{x_n!} \\ &= \frac{e^{-n\lambda}\lambda^{\sum x_i}}{x_1!\cdots x_n!}. \end{aligned}$$

Portanto a logverossimilhança é dada por

$$\begin{aligned} l(\lambda;x) &= \log e^{-n\lambda} + \log\lambda^{\sum x_i} - \log(x_1!\cdots x_n!) \\ &= -n\lambda + \log\lambda\sum_{i=1}^n x_i - \log(x_1!\cdots x_n!). \end{aligned}$$

Derivando a logverossimilhança temos que

$$l'(\lambda;x) = -n + \frac{\sum_{i=1}^n x_i}{\lambda} = 0 \iff \lambda = \frac{\sum_{i=1}^n x_i}{n} = \overline{x}.$$

Como

$$l''(\lambda;x) = -\frac{\sum_{i=1}^n x_i}{\lambda^2} < 0$$

concluímos que

$$\hat{\lambda} = \overline{X}.$$

7. Usando os dados e o estimador encontrado na **Questão 6** temos o valor $\hat{\lambda} = 2,70$. Desse modo, segue que

$$\begin{aligned} \mathcal{P}(X \leqslant 2) &= \mathcal{P}(X=0) + \mathcal{P}(X=1) + \mathcal{P}(X=2) \\ &= \frac{e^{-2,70}(2,70)^0}{0!} + \frac{e^{-2,70}(2,70)^1}{1!} + \frac{e^{-2,70}(2,70)^2}{2!} \\ &\simeq 0,4936. \end{aligned}$$

8. Como $X_i \sim \mathcal{N}(\mu, \sigma^2)$, a verosimilhança dos parâmetros μ e σ^2 é

$$\begin{aligned} L\left(\mu, \sigma^2; x\right) &= f_{X_1}(x_1) \cdots f_{X_n}(x_n) \\ &= \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[\frac{-(x_i-\mu)^2}{2\sigma^2}\right] \\ &= \left(\frac{1}{2\pi}\right)^{n/2} \frac{1}{\sigma^n} \exp\left[-\sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\sigma^2}\right]. \end{aligned}$$

Portanto a logverossimilhança é dada por

$$l\left(\mu, \sigma^2; x\right) = -\frac{n}{2}\log(2\pi) - n\log\sigma - \frac{\sum(x_i-\mu)^2}{2\sigma^2}.$$

Derivando a logverossimilhança com relação aos parâmetros μ e σ temos

$$\frac{\partial l}{\partial \mu} = \frac{\sum(x_i-\mu)}{\sigma^2} \tag{1}$$

$$\frac{\partial l}{\partial \sigma} = -\frac{n}{\sigma} + \frac{\sum(x_i-\mu)^2}{\sigma^3}. \tag{2}$$

Igualando as equações (1) e (2) a zero e resolvendo o sistema linear para μ e σ, encontramos facilmente

$$\mu = \frac{\sum_{i=1}^{n} x_i}{n} \quad \text{e} \quad \sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n}}.$$

Calculado as derivadas parciais segundas obtemos

$$A = \frac{\partial^2 l}{\partial \mu^2} = -\frac{n}{\sigma^2}, \quad B = \frac{\partial^2 l}{\partial \mu \partial \sigma} = -\frac{2\sum(x_i-\mu)}{\sigma^2} \quad \text{e} \quad C = \frac{\partial^2 l}{\partial \sigma^2} = \frac{n}{\sigma^2} - \frac{3\sum(x_i-\mu)^2}{\sigma^4}.$$

Como $A < 0$ e $AC - B^2 > 0$, concluímos que

$$\hat{\mu} = \frac{\sum_{i=1}^{n} X_i}{n} \quad \text{e} \quad \hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{n}}.$$

9. Tomando o logaritmo natural dos valores fornecidos temos,

0,7885	1,2238	0,4700	0,2231	0,9933	1,1939	0,4700	1,0296	0,9163	0,6419

Usando os estimadores encontrados na **Questão 8**, temos

$$\overline{x} = 0,7504 \quad \text{e} \quad s = 0,4351.$$

Concluímos assim que $\log(X) \sim \mathcal{N}(0,7504; 0,4351^2)$.

Desse modo, podemos normalizar a variável para encontrar a probabilidade

$$\begin{aligned}
\mathcal{P}(2 < X < 3) &= \mathcal{P}(\log(2) < \log(X) < \log(3)) \\
&= \mathcal{P}\left(\frac{\log(2) - 0,7504}{0,4351} < \frac{\log(X) - 0,7504}{0,4351} < \frac{\log(3) - 0,7504}{0,4351}\right) \\
&\simeq \mathcal{P}(-0,1316 < Z < 0,8003) \\
&= \mathcal{P}(Z < 0,8003) - \mathcal{P}(Z < -0,1316) \\
&= \mathcal{P}(Z < 0,8003) + \mathcal{P}(Z < 0,1336) - 1 \\
&\simeq 0,78814 + 0,55172 - 1 \\
&= 0,3399.
\end{aligned}$$

10. Para $x_i \geqslant \theta$, $i = 1, \ldots, n$ a verossimilhança do parâmetro θ é

$$\begin{aligned}
L\left(\mu, \sigma^2; x\right) &= f_{X_1}(x_1) \cdots f_{X_n}(x_n) \\
&= e^{-(x_1-\theta)} \cdots e^{-(x_n-\theta)} \\
&= e^{-\sum x_i + n\theta}.
\end{aligned}$$

Note que a função acima é crescente para o parâmetro θ, entretanto a restrição $\theta \leqslant x_i$ para todo $i = 1, \ldots, n$ nos leva a

$$\theta = \text{Min}(x_1, \cdots, x_n)$$

como solução e, portanto

$$\hat{\theta} = \text{Min}(X_1, \cdots, X_n).$$

11. Nós temos $n = 20$, $\sigma = 2$ e, a partir dos dados, $\overline{x} = 4,745$. Queremos um intervalo de confiança de 95%, ou seja, $\alpha = 0,05$.

$$\alpha = 0,05 \Longrightarrow z_{\alpha/2} = z_{0,025} = 1,96.$$

Portanto,

$$\text{IC}(\mu; 0,95) = \left(4,745 - 1,96\frac{2}{\sqrt{20}}\ ;\ 4,745 + 1,96\frac{2}{\sqrt{20}}\right) = (3,868\ ;\ 5,622).$$

12. Nós temos $n = 625$ e $p = 0,7$. Queremos um intervalo de confiança de 90%, ou seja, $\alpha = 0,1$.

$$\alpha = 0,1 \Longrightarrow z_{\alpha/2} = z_{0,05} = 1,64.$$

Portanto,

$$\text{IC}(p; 0,90) = \left(0,7 - 1,64\sqrt{\frac{(0,7)(0,3)}{625}}\ ;\ 0,7 + 1,64\sqrt{\frac{(0,7)(0,3)}{625}}\right) = (0,67\ ;\ 0,73).$$

13. Nós temos $n = 25$, $\sigma = 4$ e $\overline{x} = 8$. Para as confianças 80%, 85%, 90% e 95% temos, respectivamente, $\alpha = 0,2$, $\alpha = 0,15$, $\alpha = 0,10$ e $\alpha = 0,05$. Para esses valores, nós temos $z_{\alpha/2}$ igual a 1,28, 1,44, 1,64 e 1,96 respectivamente. Como sabemos

$$\mathrm{IC}(\mu; 1-\alpha) = \left(\overline{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \ ; \ \overline{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right).$$

Substituindo os valores chegamos aos intervalos respectivos

$$(6,976 \ ; \ 9,024), \quad (6,848 \ ; \ 9,152), \quad (6,688 \ ; \ 9,312), \quad \text{e} \quad (6,432 \ ; \ 9,568).$$

Podemos observar que quanto maior a confiança que exigimos, maior o tamanho do intervalo de confiança para a média. Isso ocorre pois, para termos uma confiança maior, precisamos diminuir a precisão das estimativas.

14. Nós temos $\sigma = 9$. Para uma confiança de 90% temos $\alpha = 0,1$ e, portanto, $z_{\alpha/2} = 1,64$. Como sabemos,

$$\mathrm{IC}(\mu; 0,90) = \left(\overline{x} - 1,64\frac{\sigma}{\sqrt{n}} \ ; \ \overline{x} + 1,64\frac{\sigma}{\sqrt{n}}\right)$$

e então, a amplitude será

$$\left(\overline{x} + 1,64\frac{\sigma}{\sqrt{n}}\right) - \left(\overline{x} - 1,64\frac{\sigma}{\sqrt{n}}\right) = 3,28\left(\frac{\sigma}{\sqrt{n}}\right).$$

Para $n = 30$, $n = 50$ e $n = 100$ os valores calculados, respectivamente, para a amplitude dos dados são

$$5,39; \quad 4,17, \quad \text{e} \quad 2,95.$$

15. Inicialmente temos que $X \sim \mathcal{N}(\mu, 30)$ e, portanto, $\overline{X} \sim \mathcal{N}(\mu, 30/n)$. Segue que

$$\begin{aligned}
&\mathcal{P}(|\overline{X} - \mu| < 3) \geqslant 0,92 \iff \\
&\mathcal{P}\left(\frac{|\overline{X} - \mu|}{\sqrt{30/n}} < \frac{3}{\sqrt{30/n}}\right) \geqslant 0,92 \iff \\
&\mathcal{P}\left(|Z| < \frac{3\sqrt{n}}{\sqrt{30}}\right) \geqslant 0,92 \iff \\
&2\mathcal{P}\left(Z < \frac{3\sqrt{n}}{\sqrt{30}}\right) - 1 \geqslant 0,92 \iff \\
&\mathcal{P}\left(Z < \frac{3\sqrt{n}}{\sqrt{30}}\right) \geqslant 0,96 \iff \\
&\frac{3\sqrt{n}}{\sqrt{30}} \geqslant 1,76 \iff \\
&n \geqslant \left(\frac{1,76\sqrt{30}}{3}\right)^2 \iff \\
&n \geqslant 10,325.
\end{aligned}$$

Concluímos que o tamanho da amostra deve ser 11.

16. Nós temos $n = 40$ e $\sigma = 2$.

   a) Neste caso temos $\overline{x} = 9,3$. Para uma confiança de 94% temos $\alpha = 0,06$ e, portanto, $z_{\alpha/2} = 1,88$. Segue que,

$$\text{IC}(\mu; 0,94) = \left(\overline{x} - 1,88\frac{\sigma}{\sqrt{n}}\ ;\ \overline{x} + 1,88\frac{\sigma}{\sqrt{n}}\right) = (8,7055\ ;\ 9,8945).$$

   b) Considerando a informação sobre a amplitude nós temos que

$$\frac{2\sigma}{\sqrt{n}} z_{\alpha/2} = 1,5 \iff z_{\alpha/2} = \frac{1,5\sqrt{n}}{2\sigma} = 2,37.$$

   Por outro lado $\mathcal{P}(Z \leqslant 2,37) = 0,99$ de modo que

$$\alpha/2 = 1\% \Rightarrow \alpha = 2\% \Rightarrow \gamma = 98\%.$$

17. Inicialmente sabemos que $n = 100$, $\sigma = 2$ e $\alpha = 5\%$.

   a) $\overline{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 35,21 \iff \overline{x} = 35,21 + 1,96 \times \frac{2}{10} = 35,602.$

   b) Neste caso $z_{\alpha/2} = 1,64$ e temos

$$\text{IC}(\mu; 0,90) = \left(\overline{x} - 1,64\frac{\sigma}{\sqrt{n}}\ ;\ \overline{x} + 1,64\frac{\sigma}{\sqrt{n}}\right) = (35,274\ ;\ 35,930).$$

18. Do enunciado temos que $X \sim \mathcal{N}(\mu, 2)$ e, portanto, $\overline{X} \sim \mathcal{N}(\mu, 2/n)$. Com confiança de 90%, segue que

   a)

$$\begin{aligned}
&\mathcal{P}(|\overline{X} - \mu| < 0,5) \geqslant 0,9 \iff \\
&\mathcal{P}\left(\frac{|\overline{X} - \mu|}{\sqrt{2/n}} < \frac{0,5}{\sqrt{2/n}}\right) \geqslant 0,9 \iff \\
&\mathcal{P}\left(|Z| < \frac{0,5\sqrt{n}}{\sqrt{2}}\right) \geqslant 0,9 \iff \\
&2\mathcal{P}\left(Z < \frac{0,5\sqrt{n}}{\sqrt{2}}\right) - 1 \geqslant 0,9 \iff \\
&\mathcal{P}\left(Z < \frac{0,5\sqrt{n}}{\sqrt{2}}\right) \geqslant 0,95 \iff \\
&\frac{0,5\sqrt{n}}{\sqrt{2}} \geqslant 1,64 \iff \\
&n \geqslant \left(\frac{1,64\sqrt{2}}{0,5}\right)^2 \iff \\
&n \geqslant 21,5168.
\end{aligned}$$

   Concluímos que o tamanho da amostra deve ser 22.

b) Com as informações adicionais do item anterior temos $n = 22$, $\sigma = 2$ e $\alpha = 10\%$. Segue que $\overline{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 57,6 \iff \overline{x} = 57,6 + 1,64 \times \frac{\sqrt{2}}{\sqrt{22}} = 58,09$.

c) Do item anterior temos

$$\begin{aligned}
\mathcal{P}(58 < X < 62) &= \mathcal{P}\left(\frac{58 - 58,01}{\sqrt{2}} < Z < \frac{62 - 58,01}{\sqrt{2}}\right) \\
&= \mathcal{P}\left(\frac{58 - 58,1}{\sqrt{2}} < Z < \frac{62 - 58,1}{\sqrt{2}}\right) \\
&= \mathcal{P}(-0,07 < Z < 2,76) \\
&= \mathcal{P}(Z < 2,76) - \{1 - \mathcal{P}(Z < 0,07)\} \\
&= 0,99711 - \{1 - 0,52790\} \\
&= 0,52501.
\end{aligned}$$

19. Nós temos $n = 100$, $\sigma = 1,1$ e $\overline{x} = 2,5$. Para uma confiança de 95% temos $\alpha = 0,05$ e, portanto, $z_{\alpha/2} = 1,96$. Segue que,

$$\text{IC}(\mu; 0,95) = \left(\overline{x} - 1,96\frac{\sigma}{\sqrt{n}}\ ;\ \overline{x} + 1,96\frac{\sigma}{\sqrt{n}}\right) = (2,2844\ ;\ 2,7156).$$

20. Nós temos $n = 50$ e $\hat{p} = 0,34$. Queremos um intervalo de confiança de 94%, ou seja, $\alpha = 0,06$.

$$\alpha = 0,06 \Longrightarrow z_{\alpha/2} = z_{0,03} = 1,88.$$

Como sabemos, se o valor de p for desconhecido, o intervalo de confiança otimista para a proporção é

$$\left(\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\ ;\ \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\ \right)$$

e o intervalo conservador é

$$\left(\hat{p} - z_{\alpha/2}\sqrt{\frac{1}{4n}}\ ;\ \hat{p} + z_{\alpha/2}\sqrt{\frac{1}{4n}}\ \right).$$

Desse modo, chegamos aos valores respectivos $(0,214\ ;\ 0,466)$ e $(0,207\ ;\ 0,473)$.

21. As informações iniciais são que $n = 100$ e $p = 0,6$.

a) Com confiança de 80%, segue que

$$
\begin{aligned}
&\mathcal{P}(|\hat{p}-p| \leqslant 0,01) \geqslant 0,8 \iff \\
&\mathcal{P}\left(\frac{|\hat{p}-p|}{\sqrt{p(1-p)/n}} \leqslant \frac{0,01}{\sqrt{p(1-p)/n}}\right) \geqslant 0,8 \iff \\
&\mathcal{P}\left(|Z| \leqslant 0,02\sqrt{n}\right) \geqslant 0,8 \iff \\
&2\mathcal{P}\left(Z \leqslant 0,02\sqrt{n}\right) - 1 \geqslant 0,8 \iff \\
&\mathcal{P}\left(Z \leqslant 0,02\sqrt{n}\right) \geqslant 0,9 \iff \\
&0,02\sqrt{n} \geqslant 1,28 \iff \\
&n \geqslant \left(\frac{1,28}{0,02}\right)^2 \iff \\
&n \geqslant 4.096.
\end{aligned}
$$

Concluímos que o tamanho da amostra deve ser 4.096.

b) Nesse caso temos que $\hat{p} = 0,55$ e $z_{\alpha/2} = 1,96$. Segue que

$$
\left(\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\ ;\ \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\ \right) = (0,5347\ ;\ 0,5652).
$$

22. a) Como p é desconhecido, utilizamos o fato de que o valor máximo do produto $p(1-p)$ é $1/4n$ e, então, com confiança de 90% temos que

$$
\begin{aligned}
&\mathcal{P}(|\hat{p}-p| \leqslant 0,05) \geqslant 0,9 \iff \\
&\mathcal{P}\left(\frac{|\hat{p}-p|}{\sqrt{1/4n}} \leqslant \frac{0,05}{\sqrt{1/4n}}\right) \geqslant 0,9 \iff \\
&\mathcal{P}\left(|Z| \leqslant 0,1\sqrt{n}\right) \geqslant 0,9 \iff \\
&2\mathcal{P}\left(Z \leqslant 0,1\sqrt{n}\right) - 1 \geqslant 0,9 \iff \\
&\mathcal{P}\left(Z \leqslant 0,1\sqrt{n}\right) \geqslant 0,95 \iff \\
&0,1\sqrt{n} \geqslant 1,64 \iff \\
&n \geqslant \left(\frac{1,64}{0,1}\right)^2 \iff \\
&n \geqslant 268,96.
\end{aligned}
$$

Concluímos que o tamanho da amostra deve ser 269.

b) Basta repetirmos os passos do item anterior substituindo $1/4n$ por $p(1-p)/n = (0,8 \times 0,2)/n$. Chegamos assim à desigualdade

$$
n \geqslant \left(\frac{1,64\sqrt{0,8 \times 0,2}}{0,05}\right)^2 = 172,13.
$$

Concluímos que o tamanho da amostra deve ser 173.

c) Para uma amostra de tamanho 81 temos que

$$\mathcal{P}(|\hat{p}-p| \leqslant \epsilon) = 0,9 \iff$$
$$\mathcal{P}\left(\frac{|\hat{p}-p|}{\sqrt{1/(4\times 81)}} \leqslant \frac{\epsilon}{\sqrt{1/(4\times 81)}}\right) = 0,9 \iff$$
$$\epsilon\sqrt{4\times 81} = 1,28 \iff$$
$$\epsilon = 0,07.$$

d) Tomando $n = 81$ e o erro no valor de 0,08 temos

$$\begin{aligned}
\mathcal{P}(|\hat{p}-p| \leqslant 0,08) &= \mathcal{P}\left(\frac{|\hat{p}-p|}{\sqrt{1/(4\times 81)}} \leqslant \frac{0,08}{\sqrt{1/(4\times 81)}}\right) \\
&= \mathcal{P}(|Z| \leqslant 1,44) \\
&= 2\mathcal{P}(Z \leqslant 1,44) - 1 \\
&= 2(0,92507) - 1 \\
&= 0,8504.
\end{aligned}$$

23. Do enunciado nós temos que $n = 300$ e $\hat{p} = 1/3$.

a) Das informações e do fato que $z_{\alpha/2} = 1,96$ temos

$$\left(\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\ ;\ \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right) = (0,2799\ ;\ 0,3867).$$

b) Como p é desconhecido, utilizamos o fato de que o valor máximo do produto $p(1-p)$ é $1/4n$ e, então, com confiança de 95% temos que

$$\mathcal{P}(|\hat{p}-p| \leqslant 0,02) \geqslant 0,95 \iff$$
$$\mathcal{P}\left(\frac{|\hat{p}-p|}{\sqrt{1/4n}} \leqslant \frac{0,02}{\sqrt{1/4n}}\right) \geqslant 0,95 \iff$$
$$\mathcal{P}\left(|Z| \leqslant 0,04\sqrt{n}\right) \geqslant 0,95 \iff$$
$$2\mathcal{P}\left(Z \leqslant 0,04\sqrt{n}\right) - 1 \geqslant 0,95 \iff$$
$$\mathcal{P}\left(Z \leqslant 0,04\sqrt{n}\right) \geqslant 0,975 \iff$$
$$0,04\sqrt{n} \geqslant 1,96 \iff$$
$$n \geqslant \left(\frac{1,96}{0,04}\right)^2 \iff$$
$$n \geqslant 2.401.$$

Concluímos que o tamanho da amostra deve ser 2.401.

24. Com base na tabela da distribuição t-Student seguem os resultados abaixo:

a)

$$\begin{aligned}
\mathcal{P}(-3,365 < t_5 < 3,365) &= \mathcal{P}(t_5 < 3,365) - \mathcal{P}(t_5 < -3,365) \\
&= \mathcal{P}(t_5 < 3,365) - \mathcal{P}(t_5 > 3,365) \\
&= \mathcal{P}(t_5 < 3,365) - \{1 - \mathcal{P}(t_5 < 3,365)\} \\
&= 2\mathcal{P}(t_5 < 3,365) - 1 \\
&\simeq 2(0,99) - 1 \\
&= 0,98
\end{aligned}$$

b)

$$\begin{aligned}
\mathcal{P}(|t_8| < 1,4) &= \mathcal{P}(-1,4 < t_8 < 1,4) \\
&= \mathcal{P}(t_8 < 1,4) - \mathcal{P}(t_8 < -1,4) \\
&= \mathcal{P}(t_8 < 1,4) - \mathcal{P}(t_8 > 1,4) \\
&= \mathcal{P}(t_8 < 1,4) - \{1 - \mathcal{P}(t_8 < 1,4)\} \\
&= 2\mathcal{P}(t_8 < 1,4) - 1 \\
&\simeq 2(0,9) - 1 \\
&= 0,8
\end{aligned}$$

c)

$$\begin{aligned}
\mathcal{P}(-1,1 < t_{14} < 2,15) &= \mathcal{P}(t_{14} < 2,15) - \mathcal{P}(t_{14} < -1,1) \\
&= \mathcal{P}(t_{14} < 2,15) - \mathcal{P}(t_{14} > 1,1) \\
&= \mathcal{P}(t_{14} < 2,15) - \{1 - \mathcal{P}(t_{14} < 1,1)\} \\
&= \mathcal{P}(t_{14} < 2,15) + \mathcal{P}(t_{14} < 1,1) - 1 \\
&\simeq 0,975 + 0,85 - 1 \\
&= 0,825
\end{aligned}$$

d)

$$\begin{aligned}
\mathcal{P}(t_9 > a) &= 0,02 \qquad \Longleftrightarrow \\
\mathcal{P}(t_9 \leqslant a) &= 0,98 \qquad \Longleftrightarrow \\
a &= 2,3984
\end{aligned}$$

e)

$$\begin{aligned}
\mathcal{P}(t_{16} \leqslant b) &= 0,05 \qquad \Longleftrightarrow \\
\mathcal{P}(t_{16} \geqslant -b) &= 0,05 \qquad \Longleftrightarrow \\
\mathcal{P}(t_{16} \leqslant -b) &= 0,95 \qquad \Longleftrightarrow \\
-b &= 1,7459 \\
b &= -1,7459
\end{aligned}$$

25. Como a variância é desconhecida, devemos estimar tanto a média (por meio da média amostral) como a variância (por meio da variância amostral) para obter $\overline{x} = 10,07$ e $s = 2,74$. Sabemos que o intervalo de confiança para a média com variância desconhecida é

$$\mathrm{IC}(\mu; 1-\alpha) = \left(\overline{x} - t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}\ ;\ \overline{x} + t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}\right).$$

Uma vez que o tamanho da amostra é 30, temos para as confianças de 90% e 95%

$$\begin{aligned}\alpha &= 0,05 \Longrightarrow t_{\alpha/2,n-1} = t_{0,025;29} = 2,0452\\ \alpha &= 0,10 \Longrightarrow t_{\alpha/2,n-1} = t_{0,050;29} = 1,6991\end{aligned}$$

e os intervalos de confiança são, respectivamente,

$$(9,0469\ ;\ 11,0931) \quad \text{e} \quad (9,2200\ ;\ 10,9199).$$

26. Definindo da mesma maneira $\mathcal{P}(Z \leqslant z_\alpha) = 1-\alpha$, basta observar que

$$\begin{aligned}1-\alpha &= \mathcal{P}(Z \leqslant z_\alpha)\\ &= \mathcal{P}\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \leqslant z_\alpha\right)\\ &= \mathcal{P}\left(\overline{X} < z_\alpha\frac{\sigma}{\sqrt{n}} + \mu\right)\\ &= \mathcal{P}\left(\mu > \overline{X} - z_\alpha\frac{\sigma}{\sqrt{n}}\right).\end{aligned}$$

e o intervalo de confiança superior, após observamos os dados, é

$$\mathrm{IC}(\mu; 1-\alpha) = \left(\overline{x} - z_\alpha\frac{\sigma}{\sqrt{n}}\ ;\ +\infty\right).$$

Por outro lado, se $\mathcal{P}(Z \leqslant z_\alpha) = 1-\alpha$ então $\mathcal{P}(Z > -z_\alpha) = 1-\alpha$ e então

$$\begin{aligned}1-\alpha &= \mathcal{P}(Z > -z_\alpha)\\ &= \mathcal{P}\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} > -z_\alpha\right)\\ &= \mathcal{P}\left(\overline{X} > -z_\alpha\frac{\sigma}{\sqrt{n}} + \mu\right)\\ &= \mathcal{P}\left(\mu < \overline{X} + z_\alpha\frac{\sigma}{\sqrt{n}}\right).\end{aligned}$$

e o intervalo de confiança inferior, após observamos os dados, é

$$\mathrm{IC}(\mu; 1-\alpha) = \left(-\infty\ ;\ \overline{x} + z_\alpha\frac{\sigma}{\sqrt{n}}\right).$$

Se a variância for desconhecida, basta lembrar que no lugar de σ^2 usamos o estimador S^2 e no lugar da distribuição normal usamos a distribuição t de Student, obtendo, assim, os intervalos de confiança superior

$$\mathrm{IC}(\mu; 1-\alpha) = \left(\overline{x} - t_{\alpha,n-1}\frac{s}{\sqrt{n}}\ ;\ +\infty\right)$$

e inferior

$$\mathrm{IC}(\mu; 1-\alpha) = \left(-\infty\ ;\ \overline{x} + t_{\alpha,n-1}\frac{s}{\sqrt{n}}\right).$$

27. Da **Questão 25** já sabemos que $\overline{x} = 10,07$ e $s = 2,74$.

$$\alpha = 0,05 \Longrightarrow t_{\alpha,n-1} = t_{0,05;29} = 1,6991.$$

Desse modo, os intervalos de confiança superior e inferior são respectivamente

$$(9,2200\ ;\ +\infty) \quad \text{e} \quad (-\infty\ ;\ 10,9199).$$

28. Para obtermos um intervalo de confiança para $\mu_1 - \mu_2$ precisamos conhecer a distribuição de $\overline{X} - \overline{Y}$. Como

$$\begin{aligned} \overline{X} &\sim \mathcal{N}\left(\mu_1, \sigma_1^2/n\right) \\ \overline{Y} &\sim \mathcal{N}\left(\mu_2, \sigma_2^2/m\right) \end{aligned}$$

e cada $X_1, \ldots, X_n, Y_1, \ldots, Y_m$ são todos independentes entre si, podemos afirmar que

$$\overline{X} - \overline{Y} \sim \mathcal{N}\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}\right).$$

Concluímos assim que

$$\frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} \sim \mathcal{N}(0,1)$$

e, portanto

$$\mathcal{P}\left(-z_{\alpha/2} < \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} < z_{\alpha/2}\right) = 1 - \alpha$$

ou equivalentemente

$$\mathcal{P}\left(\overline{X} - \overline{Y} - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} < \mu_1 - \mu_2 < \overline{X} - \overline{Y} + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}\right) = 1 - \alpha$$

e, após os dados serem observados, o intervalo de confiança é

$$\mathrm{IC}\left(\mu_1 - \mu_2, 1 - \alpha\right) = \left(\overline{x} - \overline{y} - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}\ ;\ \overline{x} - \overline{y} + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}\right).$$

29. Da **Questão 28**, nós temos que

$$\mathrm{IC}\left(\mu_A - \mu_B, 1 - \alpha\right) = \left(\overline{x}_A - \overline{x}_B - z_{\alpha/2}\sqrt{\frac{\sigma_A^2}{n} + \frac{\sigma_B^2}{m}}\ ;\ \overline{x}_A - \overline{x}_B + z_{\alpha/2}\sqrt{\frac{\sigma_A^2}{n} + \frac{\sigma_B^2}{m}}\right).$$

A partir dos valores $n_A = 14$, $n_B = 12$, $\overline{x}_A = 42,79$, $\overline{x}_B = 55,83$, $\sigma_A^2 = 40$ e $\sigma_B^2 = 100$ encontramos

$$\mathrm{IC}(\mu_A - \mu_B; 0,95) = (-19,59\ ;\ -6,48).$$