Estatística - ME 414: Gabarito da Lista 4 [1]

Questão 1.
X= número de pessoas acima de 40 anos que têm artrite Para uma amostra de $X = 240$ de $n = 4000$ tem-se:

a. $\hat{p} = \frac{X}{n} = \frac{240}{4000} = 0.06$

b. $IC(p; 0.95) = \left[\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$, $z_{\alpha/2}$ é o percentil $\alpha/2$ da distribuição Normal então $z_{\alpha/2} = z_{0.025} = 1.96$, assim:

$$\begin{aligned} IC(p; 0.95) &= \left[0.06 - 1.966\sqrt{\frac{0.06(1-0.06)}{4000}}, 0.06 + 1.96\sqrt{\frac{0.06(1-0.06)}{4000}}\right] \\ &= [0.052, 0.067] \end{aligned}$$

A um grau de confiança de 95% estimamos que a proporção de pessoas acima de 40 anos pertence ao intervalo [0.052,0.067]

Questão 2.
p=Proporção de pessoas que preferem a marca X de sabonete, $\hat{p} = 0.70$, n=625, $z_{\alpha/2} = z_{0.05} = 1.64$, então:

$$IC(p; 0.90) = \left[0.70 - 1.64\sqrt{\frac{0.70(1-0.70)}{625}}, 0.06 + 1.64\sqrt{\frac{0.70(1-0.70)}{625}}\right] = [0.67, 0.73] \tag{1}$$

a um grau de confiança de 90% é estimado que a proporção de pessoas preferem a marca X de sabonete pertence ao intervalo [0.67,0.73].

Questão 3.
p=Proporção de pessoas que são favoráveis a um determinado partido $\hat{p} = 0.6$ e $n = 100$.

- a. Seja ME a margem de erro $\alpha = 0.2$

$$ME = z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = z_{0.1}\sqrt{\frac{0.6(1-0.6)}{n}} = 0.01 \Rightarrow n = \frac{0.24(1.285)^2}{0.01^2} = 3962.94 \tag{2}$$

assim, o tamanho da amostra seria 3963 para que o erro cometido na estimativa seja não máximo 0.01 com compatibilidade 0.8

b. Se $n = 3963$ $\hat{p} = 0.55$ e $z_{\alpha/2} = z_{0.025} = 1.96$

$$\begin{aligned} IC(p; 0.95) &= \left[0.55 - 1.96\sqrt{\frac{0.55(1-0.55)}{3963}}, 0.55 + 1.96\sqrt{\frac{0.0.55(1-0.55)}{3963}}\right] \\ &= [0.534, 0.565] \end{aligned}$$

[1]PED: Lisbeth Corbacho Carazas, ra162526@ime.unicamp.br

com grau de confiança ao 95% é estimado que a proporção de pessoas são favoráveis a um determinado partido está no intervalo [0.534,0.565].

Questão 4.
p=Proporção de consumidores de um certo produto.
$n = 300$ $\hat{p} = \frac{1}{3}$.

a. $z_{\alpha/2} = z_{0.025} = 1.96$

$$IC(p; 0.95) = \left[\frac{1}{3} - 1.96\sqrt{\frac{\frac{1}{3}(1-\frac{1}{3})}{300}}, \frac{1}{3} + 1.96\sqrt{\frac{\frac{1}{3}(1-\frac{1}{3})}{300}}\right] = [0.280; 0.387] \tag{3}$$

b. ME margem de erro

$$\begin{aligned} ME &= z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = z_{0.025}\sqrt{\frac{\frac{1}{3}(1-\frac{1}{3})}{n}} = 1.96\sqrt{\frac{2/9}{n}} = 0.02 \\ \Rightarrow n &= \frac{2(1.96)^2}{9(0.02)^2} = 2134.22 \end{aligned}$$

Portanto, o tamanho da amostra seria 2135 para que o erro cometido na estimativa seja não máximo 0.02.

Questão 5.
Sejam $X_1, ..., X_n$ amostras aleatórias com distribuição $N(\mu, \sigma^2)$

$$E(T) = E\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} E(a_i X_i) = \sum_{i=1}^{n} a_i E(X_i) = \sum_{i=1}^{n} a_i \mu_i = \mu \sum_{i=1}^{n} a_i = \mu(1) = \mu.$$

Portanto, T é um estimador não viciado a média μ

Questão 6.

Seja $n = 50$, $\bar{x} = 20$ e $\sigma = 2$

a.

$$\begin{aligned} IC(\mu, 1-\alpha) &= \left[\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right] \\ IC(\mu, 0.99) &= \left[20 - 2.575\frac{2}{\sqrt{50}}, 20 + 2.575\frac{2}{\sqrt{50}}\right] = [19.772, 21.228] \end{aligned}$$

A um grau de confiança igual a 99% é estimado que a média populacional do tempo gasto para montar o brinquedo está entre 19.772 e 21.228.

b. $\alpha = 0.01 \Rightarrow z_{0.005} = 2.575$ e $ME = \epsilon = 0.10$

$$\epsilon = z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \Rightarrow n = \sigma^2\left(\frac{z_{\alpha/2}}{\epsilon}\right) = 4\left(\frac{2.575}{0.10}\right)^2 = 2652.25 \tag{4}$$

Portanto o tamanho amostral deverá ser $m = 2653$.

Questão 7.

Sejam $n = 25$, $\alpha = 0.05$ então $t_{24;0.025} = 2.064$ $\bar{x} = 97.64$ e $s = 17.821$.

$$\begin{aligned} IC(\mu, 1-\alpha) &= \left[\bar{x} - t_{n-1,\alpha/2}\frac{s}{\sqrt{n}}, \bar{x} + t_{n-1,\alpha/2}\frac{s}{\sqrt{n}}\right] \\ IC(\mu, 0.95) &= \left[97.64 - 2.064\frac{17.821}{\sqrt{25}}, 97.64 + 2.064\frac{17.821}{\sqrt{25}}\right] = [90.283, 104.997]. \end{aligned}$$

A um grau de confiança do 95% estima-se a média populacional do nível de glicêmico pertence ao intervalo [90.283,104.997].

Questão 8.

Sejam $n = 100$, $\bar{x} = 500$, $\sigma = 5$ e $\alpha = 0.05$ então $z_{0.025} = 1.96$

$$\begin{aligned} IC(\mu, 1-\alpha) &= \left[\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right] \\ IC(\mu, 0.95) &= \left[500 - 1.96\frac{5}{\sqrt{100}}, 500 + 1.96\frac{5}{\sqrt{100}}\right] = [499.02, 500.98]. \end{aligned}$$

A um grau de confiança do 95% estima-se a média populacional do tempo de vida de uma peça de equipamento pertence ao intervalo [499.02,500.98].

Questão 9.

Seja $\alpha = 0.10$ então $z_{0.05} = 1.645$ e C o cumprimento do intervalo de confiança

$$\begin{aligned} IC(\mu, 1-\alpha) &= \left[\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right] \\ \Rightarrow \quad & C = 2z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < 0.2\sigma \\ \Rightarrow \quad & n > \left(\frac{2z_{\alpha/2}\sigma}{0.20\sigma}\right)^2 = \left(\frac{2(1.645)}{0.20}\right) = 16.45^2 = 270.60 \\ \Rightarrow \quad & n \geq 271 \end{aligned}$$

Questão 10.

Os dois grupos são amostras aleatórias de duas populacionais independentes e normalmente distribuídas. Os tamanhos das amostras são $n_1 = n_2 = 18$, as medidas

amostrais $\bar{x}_1 = 6.622$ e $\bar{x}_2 = 5.750$ e as variâncias amostrais $s_1^2 = 1.325$ e $s_2^2 = 0.739$. Vamos supor que as variâncias populacionais são iguais e desconhecidas, portanto o estimador da variança populacional é:

$$s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} = \frac{17(1.325)+17(0.739)}{34} = 1.032. \tag{5}$$

Assim, o intervalo de confiança ao 95% para diferença das médias populacionais para $\alpha = 0.05$, $t_{34;0.025} = 2.032$ é dado por:

$$\begin{aligned}
IC(\mu_1-\mu_2, 1-\alpha) &= \left[(\bar{x_1}-\bar{x_2}) - t_{n_1+n_2-2,\alpha/2}\sqrt{s_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}, (\bar{x_1}-\bar{x_2}) + t_{n_1+n_2-2,\alpha/2}\sqrt{s_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}\right] \\
IC(\mu_1-\mu_2, 0.95) &= \left[(6.622-5.750) - 2.032\sqrt{1.032\left(\frac{2}{18}\right)}, (6.622-5.750) + 2.032\sqrt{1.032\left(\frac{2}{18}\right)}\right] \\
IC(\mu_1-\mu_2, 0.95) &= [0.184, 1.560].
\end{aligned}$$

A um grau de confiança de 95% estima-se que a diferença entre os resultados médios obtidos entre o grupo 1 e grupo 2 está no intervalo [0.184,1.560].

Questão 11.
Sejam $n_A = 100$ $\hat{p}_A = 0.75$, $n_B = 100$, $\hat{p}_B = 0.65$ e $\hat{p} = \frac{75+65}{200} = 0.70$

Será testado:

$$H_0 : p_A - p_B = 0 \quad vs \quad H_a : p_A - p_B > 0$$

$$z_{obs} = \frac{\hat{p}_A - \hat{p}_B}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_A}+\frac{1}{n_B}\right)}} = \frac{0.75-0.65}{\sqrt{0.70(1-0.70)\left(\frac{1}{100}+\frac{1}{100}\right)}} = 1.543 \tag{6}$$

valor de p: $P(Z \geq z_{obs}) = P(Z \geq 1.543) = 1 - P(Z < 1.543) = 1 - 0.939 = 0.061$. Se consideramos $\alpha \leq 0.05$ então com valor de p maior que α não rejeitamos a hipótese nula de que a verdadeira proporção de curados usando soro é igual a verdadeira proporção de curados não usando soro.

Questão 12.
Sejam $\hat{p}_1 = \frac{54}{150} = 0.36$, $\hat{p}_2 = \frac{33}{100} = 0.348$ e $\hat{p} = \frac{75+65}{200} = 0.70$

Será testado:

$$H_0 : p_1 - p_2 = 0 \quad vs \quad H_a : p_1 - p_2 > 0$$

$$z_{obs} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{0.36-0.33}{\sqrt{0.348(1-0.348)\left(\frac{1}{150}+\frac{1}{100}\right)}} = 0.488 \tag{7}$$

valor-de-p: $P(Z \geq z_{obs}) = P(Z \geq 0.488) = 1 - P(Z < 0.488) = 1 - 0.6872 = 0.3128$. Como o valor-de-p maior que 0.05 não rejeitamos a hipótese nula, isto é, que dois métodos tem a mesma proporção de sucesso para provocar chuva.

Questão 13.

Seja p: a proporção de associados que apoiam a política de privatização do governo.

Sejam $\alpha = 0.05$ e $n = 80$.

Será testado:

$$H_0 : p = 0 \quad vs \quad H_a : p > 0.60$$

Estatística do Teste: $Z = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$.

Valor crítico $z_{crit} = z_{0.05} = 1.64$.

Região de rejeição: Rejeitamos H_0 se $Z \geq 1.64$ então:

$$C = \left\{ (X_1, X_2, ..., X_n) : \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \geq 1.64 \right\}.$$

Questão 14.

a. Sejam $n = 15$, $\bar{X} = 83.9$, $s = 18.2$ e $\alpha = 0.10$.
Será testado:

$$H_0 : \mu = 85 \quad vs \quad H_a : \mu < 85$$

Estatística do Teste: $t_{obs} = \frac{\bar{X}-\mu_0}{\frac{s}{\sqrt{n}}} = \frac{83.9-85}{\frac{18.2}{\sqrt{15}}} = -0.234$.

Valor crítico $t_{crit} = t_{14:0.10} = -1.345$ menor que o valor observado da estatística do teste ($t_{obs} = -0.234$) não rejeitamos a hipótese que a média é 85.

b. Sejam $n = 15$, $\bar{X} = 79.1$, $s = 11.8$ e $\alpha = 0.10$.
Será testado:

$$H_0 : \mu = 76 \quad vs \quad H_a : \mu \neq 76$$

Estatística do Teste: $t_{obs} = \frac{\bar{X}-\mu_0}{\frac{s}{\sqrt{n}}} = \frac{70.10-76}{\frac{11.8}{\sqrt{15}}} = 1.017$
Valor crítico $t_{crit} = t_{14:0.05} = 1.761$ menor que o valor observado da estatística do teste ($t_{obs} = 1.017$) não rejeitamos a hipótese que a média é 76.

Questão 15.

Sejam $n = 16$, $\bar{X} = 94.32$, $\sigma = 1.20$ e $\alpha = 0.10$.

a. Será testado:

$$H_0 : \mu = 95 \quad vs \quad H_a : \mu \neq 95$$

Estatística do Teste: $z_{obs} = \frac{\bar{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{94.32-95}{\frac{1.2}{\sqrt{16}}} = -2.267$.

Valor-de-p $P(|Z| \geq |z_{obs}|) = P(|Z| \geq 2.267) = 2P(Z \geq 2.267) = 2[1 - P(Z < 2.267)] = 0.0234$.

Com valor-de-p maior que 0.01 é concluído que existe evidência para não rejeitar a hipótese de que o ponto de desvanecimento médio de uma certe marca de vegetais hidrogenados é 95.

b. Seja $\alpha = 0.01$ então $z_{crit} = 2.575$ e $\mu' = 94$ assim $\bar{X} \sim N(\mu', \sigma^2/n))$.

$$\begin{aligned}
\beta &= P(\text{não rejeitar} H_0 | H_0 \text{é falsa}) = P(|Z| < z_{crip} | \mu' = 94) \\
&= P\left(\left|\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right| < 2.575 | \mu' = 94\right) \\
&= P\left(-2.575 < \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} < 2.575 | \mu' = 94\right) \\
&= P\left(-2.575 < \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} < 2.575 | \mu' = 94\right) \\
&= P\left(-2.575 + \frac{\mu_0-\mu'}{\sigma/\sqrt{n}} < \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} < 2.575 + \frac{\mu_0-\mu'}{\sigma/\sqrt{n}} | \mu' = 94\right) \\
&= P\left(-2.575 + \frac{95-94}{1.20/\sqrt{16}} < \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} < 2.575 + \frac{95-94}{1.20/\sqrt{16}}\right) \\
&= P(0.758 < Z < 5.908) = 0.2242
\end{aligned}$$

Questão 16.

Sejam $n = 16$, $\bar{X} = 5.25$, $\sigma = 0.30$.
Será testado:

$$H_0 : \mu = 5.5 \quad vs \quad H_a : \mu \neq 5.5$$

Estatística do Teste: $z_{obs} = \frac{\bar{X}-\mu_0}{\frac{s}{\sqrt{n}}} = \frac{5.25-5.5}{\frac{0.30}{\sqrt{16}}} = -3.33$.

Valor-de-p: $P(|Z| \geq |z_{obs}|) = P(|Z| \geq 3.333) = 2(1 - P(Z < 3.333)) = 0.0008$.

Com um valor-de-p muito menor que 0.05 rejeitamos a hipótese de que o ponto médio de SiO_2 em certo tipo de cimento aluminoso é de 5.5.

Questão 17.

Sejam $n_H = 97$, $\bar{X}_H = 10.40$, $\sigma_H = 4.83$, $n_M = 148$, $\bar{X}_M = 9.26$, $\sigma_M = 4.86$ e $\alpha = 0.05$.

Será testado:

$$H_0 : \mu_H - \mu_M = 0 \;\; vs \;\; H_a : \mu_M - \mu_H > 0.60$$

Assumindo que a classificação fornecida pela escala possui distribuição normal e que as variancas são conhecidas e diferentes.

Estatística do Teste: $z_{obs} = \frac{\bar{X}_H - \bar{X}_M - \triangle_0}{\sqrt{\frac{\sigma_H^2}{\sqrt{n_H}} + \frac{\sigma_M}{\sqrt{n_M}}}} = \frac{10.40-9.26}{\sqrt{\frac{4.83^2}{97} + \frac{4.86^2}{148}}} = 1.802.$

Valor-de-p: $P(Z > z_{obs}) = P(Z > 1.802) = 1 - P(Z < 1.802) = 0.036.$

Com valor-de-p menor que 0.05 existe evidência para rejeitar a hipótese de que não existe diferença entre a facilidade que os estudantes homens es as estudantes mulheres entediam-se.

Questão 18.

Supondo que a duração das superfícies de rodagem possuem distribuição Normal. Sejam $\alpha = 0.05$, $m = 40$, $\bar{X} = 36500$, $\sigma_1 = 2200$, $n = 40$, $\bar{Y} = 33400$ e $\sigma_2 = 1900$.

Será testado:

$$H_0 : \mu_1 - \mu_2 = 0 \;\; vs \;\; H_a : \mu_1 - \mu_2 \neq 0$$

Estatística do Teste: $z_{obs} = \frac{\bar{X} - \bar{Y} - \triangle_0}{\sqrt{\frac{\sigma_1}{\sqrt{m}} + \frac{\sigma_2}{\sqrt{n}}}} = \frac{36500-33400}{\sqrt{\frac{2200^2}{40} + \frac{1900^2}{40}}} = 6.745.$

Valor-de-p: $P(|Z| > z_{obs}) = P(|Z| > 6.745) = 2(1 - P(Z < 6.745)) \simeq 0.$

Com valor-de-p menor que 0.05 existe evidência para rejeitar a hipótese de que os pontos médios de duração das superfícies de rodagem são iguais.

Questão 19.

Supondo que os dado possuem uma distribuição Normal. Sejam $\alpha = 0.01$, $m = 40$, $\bar{X} = 18.12$, $\sigma_1 = 1.60$, $n = 32$, $\bar{Y} = 16.87$ e $\sigma_2 = 1.40$.

Será testado:

$$H_0 : \mu_1 - \mu_2 = 0 \;\; vs \;\; H_a : \mu_1 - \mu_2 > 0$$

Estatística do Teste: $z_{obs} = \frac{\bar{X} - \bar{Y} - \triangle_0}{\sqrt{\frac{\sigma_1}{\sqrt{m}} + \frac{\sigma_2}{\sqrt{n}}}} = \frac{18.12-16.87}{\sqrt{\frac{1.60^2}{40} + \frac{1.40^2}{32}}} = 3.532.$

Valor-de-p: $P(Z > z_{obs}) = P(Z > 3.532) = 1 - P(Z < 3.532) = 1 - 0.9998 = 0.0002.$

Com valor-de-p menor que 0.01 existe evidência para rejeitar a hipótese nula de que a resistência média de coesão à tensão do morteiro modificado de látex de polímeros é igual a do morteiro não modificado.

Questão 20.

Supondo que o= conteúdo de nicotina das duas marcas tem distribuição Normal que as duas variançãs populacionais são iguais. Sejam $\alpha = 0.05$, $n_A = 40$, $\bar{X}_A = 20$, $s_A^2 = 6$, $n_B = 5$, $\bar{X}_B = 21$ e $s_B^2 = 5$. Com as variançãs populacionais desconhecidas e iguais, será determinado o estimados da variança:

$$S_p^2 = \frac{S_A^2(n_A - 1) + S_B^2(n_B - 1)}{n_A + n_B - 2} = \frac{6(4-1) + 5(4-1)}{4 + 5 - 2} = 5.429.$$

Será testado:

$$H_0 : \mu_1 - \mu_2 = 0 \quad vs \quad H_a : \mu_1 - \mu_2 \neq 0$$

Estatística do Teste: $t_{obs} = \frac{\bar{X}_A - \bar{X}_B - \triangle_0}{\sqrt{S_p^2\left(\frac{1}{\sqrt{n_A}} + \frac{1}{\sqrt{n_B}}\right)}} = \frac{20-21}{\sqrt{5.429(\frac{1}{4}+\frac{1}{5})}} = -0.639.$

Valor-de-p: $P(|T| > t_{obs}) = P(|T| > 0.639) = 2(1 - P(Z < 0.639)) = 2(1 - 0.728) = 0.544$

Com valor-de-p menor que 0.05 existe evidência para não rejeitar a hipótese de que o conteúdo médio de nicotina das duas marcas de cigarros é o mesmo.