

Introdução à Probabilidade e à Estatística II

Solução Lista 3

1. Considere

$\bar{x}_D$:= média da produtividade dos operários do período diurno.

s_D^2 := variância amostral da produtividade dos operários do período diurno.

$\bar{x}_N$:= média da produtividade dos operários do período noturno.

s_N^2 := variância amostral da produtividade dos operários do período noturno.

$$\text{Então } \bar{x}_D = \frac{\sum_{i=1}^{15} x_i}{15} = 12; \bar{x}_N = \frac{\sum_{i=1}^{15} y_i}{15} = 10; n_D = 15; n_N = 15;$$

$$s_D^2 = \frac{\sum_{i=1}^{15} x_i^2 - \frac{\left(\sum_{i=1}^{15} x_i\right)^2}{15}}{15-1} = 35,7; s_N^2 = \frac{\sum_{i=1}^{15} y_i^2 - \frac{\left(\sum_{i=1}^{15} y_i\right)^2}{15}}{15-1} = 105,7. \text{ Primeiramente vamos fazer um teste para a variância.}$$

Teste de igualdade de variâncias:

$$H_0 : \sigma_D^2 = \sigma_N^2 \quad \text{versus} \quad H_1 : \sigma_D^2 \neq \sigma_N^2$$

Estatística do teste: $W = S_N^2/S_D^2$. Sob H_0 , $W \sim F(n-1, n-1) = F(14, 14)$

Região Crítica: Observe que o teste é bilateral, Fixando $\alpha = 5\%$ (Consultando a Tabela F) então:

$$RC = \left\{ x \in \mathbb{R} \mid x \in \left(0, \frac{1}{F_{14,14,2.5\%}} \right) \text{ ou } x > F_{14,14,2.5\%} \right\} = \left\{ x \in \mathbb{R} \mid x \in \left(0, \frac{1}{2,98} \right) \text{ ou } x > 2,98 \right\}$$

Valor observado: $W_0 = \frac{S_N^2}{S_D^2} = \frac{105,7}{35,7} = 2,96$, como W_0 não pertence a RC, então não rejeitamos H_0 , ou seja a um nível de significância de 5% as variâncias dos dois grupos são iguais.

Teste de igualdade de médias com variâncias iguais:

$$H_0 : \mu_D = \mu_N \quad \text{versus} \quad H_1 : \mu_D \neq \mu_N$$

Estatística do teste:

$$T = \frac{\bar{x}_D - \bar{x}_N}{\sqrt{s_p^2 \left[\left(\frac{1}{n_D} + \frac{1}{n_N} \right) \right]}} \text{ com } s_p^2 = \frac{(n_D - 1)s_D^2 + (n_N - 1)s_N^2}{n_D + n_N - 2} \text{ sob } H_0, T \sim t_{n_D+n_N-2} = t_{28}$$

. Região Crítica: Observe que o teste é bilateral, Fixando $\alpha = 5\%$ (Consultando a Tabela t) então:

$$RC = \{ x \in \mathbb{R} \mid x < -t_{28,2.5\%} \text{ ou } x > t_{28,2.5\%} \} = \{ x \in \mathbb{R} \mid x < -2,0484 \text{ ou } x > 2,0484 \}$$

Valor observado:

$$T_0 = \frac{12 - 10}{\sqrt{70,7 \left[\left(\frac{1}{15} + \frac{1}{15} \right) \right]}} \cong 0,65 \text{ com } s_p^2 = \frac{(15-1)35,7 + (15-1)105,7}{15+15-2} = 70,7$$

. como T_0 não pertence a RC, então não rejeitamos H_0 , ou seja a um nível de significância de 5%, não há evidências de que as produtividades médias dos dois períodos sejam diferentes.

2. Considere

$\hat{p}_M$:= Porcentagem de votos masculinos ao seu favor.

$\hat{p}_F$:= Porcentagem de votos femininos ao seu favor.

Então $\hat{p}_M = \frac{170}{400} = 0,425$; $\hat{p}_F = \frac{194}{625} = 0,3104$; $n_M = 400$; $n_F = 625$

Teste de igualdade de proporções:

$$H_0 : p_M - p_F = 0,1 \quad \text{versus} \quad H_1 : p_M - p_F \neq 0,1$$

Estatística do teste:

$$Z = \frac{\hat{p}_M - \hat{p}_F - 0,1}{\sqrt{\frac{\hat{p}_M(1-\hat{p}_M)}{n_M} + \frac{\hat{p}_F(1-\hat{p}_F)}{n_F}}} \text{ sob } H_0, Z \sim N(0,1)$$

. Região Crítica: Observe que o teste é bilateral, Fixando $\alpha = 5\%$ (Consultando a Tabela N(0,1)) então:

$$RC = \{x \in \Re | x < -z_{2,5\%} \text{ ou } x > z_{97,5\%}\} = \{x \in \Re | x < -1,96 \text{ ou } x > 1,96\}$$

Valor observado:

$$Z_0 = \frac{0,425 - 0,3104 - 0,1}{\sqrt{\frac{0,425(1-0,425)}{400} + \frac{0,3104(1-0,3104)}{625}}} \cong 0,43$$

. como Z_0 não pertence a RC, então não rejeitamos H_0 , ou seja a um nível de significância de 5%, não há evidências de que a afirmação do partido seja falsa.

3. Como o estudo é feito com os mesmos operários, estamos diante do caso em que se pode usar variáveis emparelhadas. Considere

$\bar{d}$:= Média da diferença de peças produzidas na semana (Y) e (X).

S_d := Desvio padrão da diferença de peças produzidas na semana (Y) e (X).

Então

$$\bar{d} = \frac{(28 - 23) + (38 - 35) + (29 - 29) + (37 - 33) + (43 - 42) + (32 - 30)}{6} = 1,5$$

$$S_d = \sqrt{\frac{\sum_{i=1}^6 (d_i - 1,5)^2}{5}} \cong 2,9 \quad \text{com} \quad d = (5; 3; 0; 4; -1; -2)$$

Teste t-pareado

$$H_0 : \mu_d = 0 \quad \text{versus} \quad H_1 : \mu_d > 0$$

Estatística do teste:

$$T = \frac{\sqrt{n} \cdot \bar{d}}{S_d} \quad \text{sob} \quad H_0, \quad T \sim t_{(n-1)} = t_5$$

. Região Crítica: Observe que o teste é unilateral a direita, Fixando $\alpha = 5\%$ (Consultando a Tabela t) então:

$$RC = \{x \in \Re | x > t_{(5;95\%)}\} = \{x \in \Re | x > 2,015\}$$

Valor observado:

$$T_0 = \frac{\sqrt{6} \cdot 1,5}{2,9} \cong 1,27$$

como T_0 não pertence a RC, então não rejeitamos H_0 , ou seja a um nível de significância de 5%, não há evidências de que a pausa aumente a produtividade média dos trabalhadores.

4.

a) considere:

$\bar{x}_A$:= vida média em horas da lâmpada produzida pela fábrica A.

s_A := desvio padrão em horas da lâmpada produzida pela fábrica A.

$\bar{x}_B$:= vida média em horas da lâmpada produzida pela fábrica B.

s_B := desvio padrão em horas da lâmpada produzida pela fábrica B.

Então $\bar{x}_A = 1190$; $s_A = 90$; $\bar{x}_B = 1230$; $s_B = 120$; $n_A = 100$; $n_B = 75$;

Primeiramente vamos fazer um teste para a variância.

Teste de igualdade de variâncias:

$$H_0 : \sigma_A^2 = \sigma_B^2 \quad \text{versus} \quad H_1 : \sigma_A^2 \neq \sigma_B^2$$

Estatística do teste: $W = S_A^2/S_B^2$. Sob H_0 , $W \sim F(n_A - 1, n_B - 1) = F(99, 74)$

Região Crítica: Observe que o teste é bilateral, Fixando $\alpha = 5\%$ (Consultando a Tabela F por meio do software R) então:

$$RC = \left\{x \in \Re | x \in \left(0, \frac{1}{F_{99,74,2.5\%}}\right) \text{ ou } x > F_{99,74,2.5\%}\right\} = \left\{x \in \Re | x \in \left(0, \frac{1}{1,53}\right) \text{ ou } x > 1,53\right\}$$

Valor observado: $W_0 = \frac{S_A^2}{S_B^2} = \frac{90^2}{120^2} = 0,75$, como W_0 não pertence a RC, então não rejeitamos H_0 , ou seja a um nível de significância de 5% as variâncias dos dois grupos são iguais.

b) **Teste de igualdade de médias com variâncias iguais:**

$$H_0 : \mu_A = \mu_B \quad \text{versus} \quad H_1 : \mu_A \neq \mu_B$$

Estatística do teste:

$$T = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{s_p^2 \left[\left(\frac{1}{n_A} + \frac{1}{n_B} \right) \right]}} \text{ com } s_p^2 = \frac{(n_A - 1)s_A^2 + (n_B - 1)s_B^2}{n_A + n_B - 2}$$

$$\text{sob } H_0, T \sim t_{n_A+n_B-2} = t_{173} \cong N(0, 1)$$

. Região Crítica: Observe que o teste é bilateral, Fixando $\alpha = 5\%$ (Consultando a Tabela $N(0,1)$) então:

$$RC = \{x \in \mathbb{R} | x < -z_{2.5\%} \text{ ou } x > z_{97.5\%}\} = \{x \in \mathbb{R} | x < -1,96 \text{ ou } x > 1,96\}$$

Valor observado:

$$T_0 = \frac{1190 - 1230}{\sqrt{10795,8 \left[\left(\frac{1}{100} + \frac{1}{75} \right) \right]}} \cong -2,52 \text{ com } s_p^2 = \frac{(100 - 1)90^2 + (75 - 1)120^2}{100 + 75 - 2} = 10795,8$$

como T_0 pertence a RC, então rejeitamos H_0 , ou seja a um nível de significância de 5%, concluímos que as lâmpadas da fábrica B têm vida média populacional maior que a fábrica A.

O intervalo de confiança é dado por:

$$IC(\Delta; \gamma) = (\bar{x}_B - \bar{x}_A) \pm t_{gl, \gamma} \cdot s_p \sqrt{1/n_B + 1/n_A}$$

$$IC(\Delta; 0,95) = (1230 - 1190) \pm 1,96 \cdot 103,9 \sqrt{1/75 + 1/100} =]8,89; 71,1[$$

Logo a cada 100 intervalos construídos, espero que em 95 deles contenha a diferença entre as médias entre 8,89 e 71,1.

5. μ := média de salários de todos empregados.

σ := desvio padrão de salários de todos empregados.

$\bar{x}_T$:= média de salários dos torneiros.

s_T := desvio padrão de salários dos torneiros.

Então $\mu = 3,64$; $\sigma = 0,85$; $\bar{x}_T = 4,22$; $s_T = 1,25$; $n_T = 25$;

Primeiramente vamos fazer um teste para a variância (comparando amostra com a população).

Teste de igualdade de variâncias:

$$H_0 : \sigma_T^2 = 0,85^2 \quad \text{versus} \quad H_1 : \sigma_T^2 \neq 0,85^2$$

Estatística do teste: $\chi^2 = (n_T - 1)S_T^2/\sigma^2$. Sob H_0 , $\chi^2 \sim \chi_{(n_T-1)}^2 = \chi_{(24)}^2$

Região Crítica: Observe que o teste é bilateral, Fixando $\alpha = 5\%$ (Consultando a Tabela Qui-quadrado) então:

$$RC = \{x \in \Re | x \in (0, \chi_{24;2,5\%}^2) \text{ ou } x > \chi_{24;97,5\%}^2\} = \{x \in \Re | x \in (0; 12,4) \text{ ou } x > 39,36\}$$

Valor observado: $\chi_0^2 = \frac{(24)S_T^2}{\sigma^2} = \frac{(24) \cdot 1,25^2}{0,85^2} \cong 51,903$, como χ_0^2 pertence a RC, então rejeitamos H_0 , ou seja a um nível de significância de 5% concluímos que a variância dos salários dos torneiros mecânicos é maior que a variância dos salários dos empregados da indústria como um todo.

Teste de igualdade de médias

$$H_0 : \mu_T = 3,64 \quad \text{versus} \quad H_1 : \mu_T \neq 3,64$$

Estatística do teste:

$$T = \frac{\sqrt{n_T} \cdot (\bar{x}_T - \mu)}{S_T} \quad \text{sob } H_0, T \sim t_{(n_T-1)} = t_{24}$$

. Região Crítica: Observe que o teste é bilateral, Fixando $\alpha = 5\%$ (Consultando a Tabela t) então:

$$RC = \{x \in \Re | x < -t_{24,2,5\%} \text{ ou } x > t_{24,97,5\%}\} = \{x \in \Re | x < -2,064 \text{ ou } x > 2,064\}$$

Valor observado:

$$T_0 = \frac{5 \cdot (4,22 - 3,64)}{1,25} \cong 2,32$$

como T_0 pertence a RC, então rejeitamos H_0 , ou seja a um nível de significância de 5%, concluímos que a média dos salários dos torneiros mecânicos é maior que a média dos salários dos empregados da indústria como um todo.

6. Considere

$\bar{x}_A$:= média das notas dos alunos de economia da turma A.

s_A := desvio padrão das notas dos alunos de economia da turma A.

$\bar{x}_B$:= média das notas dos alunos de economia da turma B.

s_B := desvio padrão das notas dos alunos de economia da turma B.

Então $\bar{x}_A = 7,8$; $s_A = 0,6$; $\bar{x}_B = 7,4$; $s_B = 0,8$; $n_A = 30$; $n_B = 32$;

Primeiramente vamos fazer um teste para a variância.

Teste de igualdade de variâncias:

$$H_0 : \sigma_A^2 = \sigma_B^2 \quad \text{versus} \quad H_1 : \sigma_A^2 \neq \sigma_B^2$$

Estatística do teste: $W = s_A^2/s_B^2$. Sob H_0 , $W \sim F(n_A - 1, n_B - 1) = F(29, 31)$

Região Crítica: Observe que o teste é bilateral, Fixando $\alpha = 5\%$ (Consultando a Tabela F como uso do software R) então:

$$RC = \left\{ x \in \mathbb{R} \mid x \in \left(0, \frac{1}{F_{29,31,2.5\%}} \right) \text{ ou } x > F_{29,31,2.5\%} \right\} = \left\{ x \in \mathbb{R} \mid x \in \left(0, \frac{1}{2,07} \right) \text{ ou } x > 2,07 \right\}$$

Valor observado: $W_0 = \frac{s_A^2}{s_B^2} = \frac{0,6^2}{0,8^2} = 0,75$, como W_0 não pertence a RC, então não rejeitamos H_0 , ou seja a um nível de significância de 5% as variâncias dos dois grupos são iguais.

Teste de igualdade de médias com variâncias iguais:

$$H_0 : \mu_A = \mu_B \quad \text{versus} \quad H_1 : \mu_A > \mu_B$$

Estatística do teste:

$$T = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{s_p^2 \left[\left(\frac{1}{n_A} + \frac{1}{n_B} \right) \right]}} \text{ com } s_p^2 = \frac{(n_A - 1)s_A^2 + (n_B - 1)s_B^2}{n_A + n_B - 2} \text{ sob } H_0, T \sim t_{n_A+n_B-2} = t_{60}$$

. Região Crítica: Observe que o teste é unilateral a direita, Fixando $\alpha = 5\%$ (Consultando a Tabela t) então:

$$RC = \{x \in \mathbb{R} \mid x > t_{60,95\%}\} = \{x \in \mathbb{R} \mid x > 1,6707\}$$

Valor observado:

$$T_0 = \frac{7,8 - 7,4}{\sqrt{0,5 \left[\left(\frac{1}{30} + \frac{1}{32} \right) \right]}} \cong 2,22 \text{ com } s_p^2 = \frac{(30 - 1)0,6^2 + (32 - 1)0,8^2}{30 + 32 - 2} = 0,5$$

. como T_0 pertence a RC, então rejeitamos H_0 , ou seja a um nível de significância de 5%, concluímos que a média das notas da turma A é superior a média das notas da turma B.

7.

a) Como o estudo é feito com os mesmos pacientes, estamos diante do caso em que se pode usar variáveis emparelhadas. Considere

$\bar{d}$:= Média da diferença da pressão sanguínea (depois - antes).

S_d := Desvio padrão da diferença da pressão sanguínea (depois - antes).

Então

$$\bar{d} = \frac{(125 - 120) + (126 - 124) + \dots + (136 - 130)}{7} \cong 3,29$$

$$S_d = \sqrt{\frac{\sum_{i=1}^7 (d_i - 3,29)^2}{6}} \cong 3,25 \quad \text{com} \quad d = (5; 2; 8; -1; 3; 0; 6)$$

Teste t-pareado

$$H_0 : \mu_d = 0 \quad \text{versus} \quad H_1 : \mu_d > 0$$

Estatística do teste:

$$T = \frac{\sqrt{n} \cdot \bar{d}}{S_d} \quad \text{sob} \quad H_0, \quad T \sim t_{(n-1)} = t_6$$

. Região Crítica: Observe que o teste é unilateral a direita, Fixando $\alpha = 5\%$ (Consultando a Tabela t) então:

$$RC = \{x \in \Re | x > t_{(6;95\%)}\} = \{x \in \Re | x > 1,9432\}$$

Valor observado:

$$T_0 = \frac{\sqrt{7} \cdot 3,29}{3,25} \cong 2,67$$

como T_0 pertence a RC, então rejeitamos H_0 , ou seja a um nível de significância de 5%, concluimos o aumento da pressão diastólica após receberem a droga.

b) O intervalo de confiança é dado por:

$$IC(\mu_d; \gamma) = \bar{d} \pm t_{6,\gamma} \cdot s_d / \sqrt{n}$$

$$IC(\mu_d; 0,90) = 3,29 \pm 1,94 \cdot 3,25 / \sqrt{7} =]0,91; 5,67[$$

Logo a cada 100 intervalos construídos, espero que em 90 deles contenha a diferença entre as médias entre 0,91 e 5,67; ou seja a diferença é significativa pois não contém o 0 no intervalo.

8. Considere

$\bar{x}_A$:= média da amostra da população A.

s_A^2 := variância da amostra da população A.

$\bar{x}_B$:= média da amostra da população B.

s_B^2 := variância da amostra da população B.

Então $\bar{x}_A = 14$; $s_A^2 = 1,51^2$; $\bar{x}_B = 15,5$; $s_B^2 = 1,93^2$; $n_A = 8$; $n_B = 8$;

Primeiramente vamos fazer um teste para a variância.

Teste de igualdade de variâncias:

$$H_0 : \sigma_A^2 = \sigma_B^2 \quad \text{versus} \quad H_1 : \sigma_A^2 \neq \sigma_B^2$$

Estatística do teste: $W = S_A^2 / S_B^2$. Sob H_0 , $W \sim F(n_A - 1, n_B - 1) = F(7, 7)$

Região Crítica: Observe que o teste é bilateral, Fixando $\alpha = 5\%$ (Consultando a Tabela F) então:

$$RC = \left\{ x \in \Re | x \in \left(0, \frac{1}{F_{7,7,2.5\%}} \right) \text{ ou } x > F_{7,7,2.5\%} \right\} = \left\{ x \in \Re | x \in \left(0, \frac{1}{4,99} \right) \text{ ou } x > 4,99 \right\}$$

Valor observado: $W_0 = \frac{S_A^2}{S_B^2} = \frac{1,51^2}{1,93^2} = 0,78$, como W_0 não pertence a RC, então não rejeitamos H_0 , ou seja a um nível de significância de 5% as variâncias dos dois grupos são iguais.

Teste de igualdade de médias com variâncias iguais:

$$H_0 : \mu_A = \mu_B \quad \text{versus} \quad H_1 : \mu_A \neq \mu_B$$

Estatística do teste:

$$T = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{s_p^2 \left[\left(\frac{1}{n_A} + \frac{1}{n_B} \right) \right]}} \text{ com } s_p^2 = \frac{(n_A - 1)s_A^2 + (n_B - 1)s_B^2}{n_A + n_B - 2} \text{ sob } H_0, T \sim t_{n_A + n_B - 2} = t_{14}$$

. Região Crítica: Observe que o teste é bilateral, Fixando $\alpha = 5\%$ (Consultando a Tabela t) então:

$$RC = \{x \in \Re | x < -t_{14, 2.5\%} \text{ ou } x > t_{14, 97.5\%}\} = \{x \in \Re | x < -2,14 \text{ ou } x > 2,14\}$$

Valor observado:

$$T_0 = \frac{14 - 15,5}{\sqrt{3,03 \left[\left(\frac{1}{8} + \frac{1}{8} \right) \right]}} \cong 0,65 \text{ com } s_p^2 = \frac{(8 - 1)1,53^2 + (8 - 1)1,93^2}{8 + 8 - 2} = -1,72$$

. como T_0 não pertence a RC, então não rejeitamos H_0 , ou seja a um nível de significância de 5%, não há evidências de que a média das duas amostras diferem.