Introdução à Probabilidade e à Estatística II
Solução da lista extra

1. $\bar{x}_A :=$ peso médio amostral do tecido do tipo A.
   $s_A :=$ desvio padrão amostral do tecido do tipo A.
   $\bar{x}_B :=$ peso médio amostral do tecido do tipo B.
   $s_B :=$ desvio padrão amostral do tecido do tipo B.

   Então $\bar{x}_A = 30,85$; $\bar{x}_B = 33,57$; $\sigma^2 = 49$; $n_A = 7$; $n_B = 7$;
   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:

   $$T = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{\sigma^2 \left[\left(\frac{1}{n_A} + \frac{1}{n_B}\right)\right]}} \text{ sob } H_0,\ T \sim t_{n_A+n_B-2} = t_{12}$$

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

   $RC = \{x \in \Re | x < -t_{12,2.5\%} \text{ ou } x > t_{12,97.5\%}\} = \{x \in \Re | x < -2,1788 \text{ ou } x > 2,1788\}$

   Valor observado:
   $$T_0 = \frac{30,85 - 33,57}{\sqrt{7^2 \left[\left(\frac{1}{7} + \frac{1}{7}\right)\right]}} \cong -0,726$$

   como T_0 não pertence a RC, então não rejeitamos a hipótese que os dois tipos de tecido tenham o mesmo peso, a 5% de significância.

2. Considere
   $\bar{x} :=$ média de resistência do concreto tipo X.
   $s_x^2 :=$ variância da resistência do concreto tipo X.
   $\bar{y} :=$ média de resistência do concreto tipo Y.
   $s_y^2 :=$ variância da resistência do concreto tipo Y.

   Então $\bar{x} = 55,6$; $\bar{y} = 53$; $s_x^2 = 17,3$; $s_y^2 = 2,5$; $n_x = 5$; $n_y = 5$;
   Primeiramente vamos fazer um teste para a variância.
   Teste de igualdade de variâncias:

   $$H_0 : \sigma_x^2 = \sigma_y^2 \quad \text{versus} \quad H_1 : \sigma_x^2 \neq \sigma_y^2$$

Estatística do teste: $W = S_x^2/S_y^2$. Sob H_0, $W \sim F(n_x - 1, n_y - 1) = F(4,4)$
Região Crítica: Observe que o teste é bilateral, Fixando $\alpha = 5\%$ (Consultando a Tabela F) então:

$$RC = \left\{x \in \Re | x \in \left(0, \frac{1}{F_{4,4,2.5\%}}\right) \text{ ou } x > F_{4,4,2.5\%}\right\} = \left\{x \in \Re | x \in \left(0, \frac{1}{9,6}\right) \text{ ou } x > 9,6\right\}$$

Valor observado: $W_0 = \frac{S_x^2}{S_y^2} = \frac{17{,}3}{2{,}5} = 6,92$, 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:

$$H_0 : \mu_x = \mu_y \quad \text{versus} \quad H_1 : \mu_x > \mu_y$$

Estatística do teste:

$$T = \frac{\bar{x}_x - \bar{x}_y}{\sqrt{s_p^2\left[\left(\frac{1}{n_x} + \frac{1}{n_y}\right)\right]}} \text{ com } s_p^2 = \frac{(n_x - 1)s_x^2 + (n_y - 1)s_y^2}{n_x + n_y - 2} \text{ sob } H_0,\ T \sim t_{n_x + n_y - 2} = t_8$$

. Região Crítica: Observe que o teste é unilateral, Fixando $\alpha = 10\%$ (Consultando a Tabela t) então:

$$RC = \{x \in \Re | x > t_{8,90.0\%}\} = \{x \in \Re | x > 1,397\}$$

Valor observado:

$$T_0 = \frac{55,6 - 53}{\sqrt{9,9\left[\left(\frac{1}{5} + \frac{1}{5}\right)\right]}} \cong 1,31 \text{ com } s_p^2 = \frac{(5-1)17,3 + (5-1)2,5}{5+5-2} = 9,9$$

. como T_0 não pertence a RC, então não rejeitamos H_0, ou seja a um nível de significância de 10%, não é possível afirmar que o concreto do tipo X seja mais resistente do que o concreto do tipo Y.

3. Considere
$\hat{p}_n$:= Proporção de curados com o tratamento novo.
$\hat{p}_a$:= Proporção de curados com o tratamento antigo.
Então $\hat{p}_n = \frac{36}{50} = 0,72$; $\hat{p}_a = \frac{29}{45} = 0,644$; $n_n = 50$; $n_a = 45$
Teste de igualdade de proporções:

$$H_0 : p_n = p_a \quad \text{versus} \quad H_1 : p_n > p_a$$

Estatística do teste:

$$Z = \frac{\hat{p}_n - \hat{p}_a}{\sqrt{\frac{\hat{p}_n(1-\hat{p}_n)}{n_n} + \frac{\hat{p}_a(1-\hat{p}_a)}{n_a}}} \text{ sob } H_0,\ Z \sim N(0,1)$$

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

$$RC = \{x \in \Re | x > z_{99\%}\} = \{x \in \Re | x > 2,36\}$$

Valor observado:

$$Z_0 = \frac{0,72 - 0,644}{\sqrt{\frac{0,72(1-0,28)}{50} + \frac{0,644(1-0,644)}{45}}} \cong 0,793$$

.

como Z_0 não pertence a RC, então não rejeitamos H_0, ou seja a um nível de significância de 1%, não temos evidência para rejeitar a hipótese nula. Então, o novo tratamento não é significativamente diferente do anterior.

o intervalo de confiança para diferença da proporção é dado por

$$IC(p_n - p_a, \alpha) = \bar{p}_n - \bar{p}_a \pm z_{1-\alpha/2}\sqrt{\frac{\hat{p}_n(1-\hat{p}_n)}{n_n} + \frac{\hat{p}_a(1-\hat{p}_a)}{n_a}}$$

aplicando os valores informados no exercício, e tomando $\alpha = 0,99$, logo $z_{0,995} = 2,32$, temos

$$IC(p_n - p_a, 0,99) = 0,72 - 0,644 \pm 2,32\sqrt{\frac{0,72(1-0,28)}{50} + \frac{0,644(1-0,644)}{45}}$$

$$IC(p_n - p_a, 0,99) = (-0,146; 0,298)$$

E como $0 \in$ IC(p_n - p_a, 0.99), não podemos rejeitar a hipótese que $p_n - p_a$. Note que o intervalo de confiança para a diferença de proporções é um teste aproximado, que pode inclusive dar resultados diferentes que o teste de hipótese formal, o teste é preferível sempre que possível.

4. Considere
$\hat{p}_d :=$ Proporção de cães machos domiciliados
$\hat{p}_n :=$ Proporção de cães machos não-domiciliados.
Então $\hat{p}_d = \frac{301}{510} = 0,59$; $\hat{p}_a = \frac{97}{230} = 0,42$; $n_d = 510$; $n_n = 230$
Teste de igualdade de proporções:

$$H_0 : p_d = p_n \quad \text{versus} \quad H_1 : p_d \neq p_n$$

Estatística do teste:

$$Z = \frac{\hat{p}_d - \hat{p}_n}{\sqrt{\frac{\hat{p}_d(1-\hat{p}_d)}{n_d} + \frac{\hat{p}_n(1-\hat{p}_n)}{n_n}}} \text{ sob } H_0,\ Z \sim N(0,1)$$

. 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:

$$Z_0 = \frac{0,59 - 0,42}{\sqrt{\frac{0,59(1-0,59)}{510} + \frac{0,42(1-0,42)}{230}}} \cong 4,34$$

.

como Z_0 pertence a RC, então rejeitamos H_0, Ou seja, observamos uma diferença estatística significativa entre as proporções de machos nos dois grupos.

5. $\sigma :=$ desvio padrão populacional da tensão.
   $S :=$ desvio padrão amostral da tensão.
   $n :=$ tamanho da amostra.

   Então $\sigma = 12$; $s = 8$; $n = 30$;
   Primeiramente vamos fazer um teste para a variância (comparando amostra com a população).
   Teste de igualdade de variâncias:

   $$H_0 : \sigma^2 = 144 \quad \text{versus} \quad H_1 : \sigma^2 < 144$$

   Estatística do teste: $\chi^2 = (n-1)S^2/\sigma^2$. Sob H_0, $\chi^2 \sim \chi^2_{(n-1)} = \chi^2_{(29)}$
   Região Crítica: Observe que o teste é unilateral, Fixando $\alpha = 5\%$ (Consultando a Tabela Qui-quadrado) então:

   $$RC = \left\{x \in \Re | x \in \left(0, \chi^2_{29;5\%}\right)\right\} = \{x \in \Re | x \in (0; 17,7)\}$$

   Valor observado: $\chi^2_0 = \frac{(29) \cdot 64}{144} \cong 12,89$, como χ^2_0 pertence a RC, então rejeitamos H_0, ou seja a um nível de significância de 5%, a instalação de novos transformadores reduziu a variação da tensão na rede de transmissão de energia elétrica.

6. $\sigma :=$ variância populacional do tempo de transmissão.
   $S :=$ variância amostral do tempo de transmissão.
   $n :=$ tamanho da amostra.

   Então $\sigma^2 = 1,3$; $s^2 = 0,304$(calculou-se o variância amostral); $n = 10$;
   Primeiramente vamos fazer um teste para a variância (comparando amostra com a população).
   Teste de igualdade de variâncias:

   $$H_0 : \sigma^2 = 1,3 \quad \text{versus} \quad H_1 : \sigma^2 \neq 1,3$$

   Estatística do teste: $\chi^2 = (n-1)S^2/\sigma^2$. Sob H_0, $\chi^2 \sim \chi^2_{(n-1)} = \chi^2_{(9)}$
   Região Crítica: Observe que o teste é bilateral, Fixando $\alpha = 5\%$ (Consultando a Tabela Qui-quadrado) então:

   $$RC = \left\{x \in \Re | x \in \left(0, \chi^2_{9;2,5\%}\right) \text{ ou } x > \chi^2_{9;97,5\%}\right\} = \{x \in \Re | x \in (0; 2,7) \text{ ou } x > 19\}$$

Valor observado: $\chi_0^2 = \frac{(9)\cdot 0{,}304}{1{,}3} \cong 2,1$, como χ_0^2 pertence a RC, então rejeitamos H_0, ou seja a um nível de significância de 5%, as mudanças realizadas na rede de computadores reduzem a variabilidade.

7. Considere
$\bar{x}_L :=$ média salarial de profissionais liberais.
$s_L :=$ desvio padrão salarial de profissionais liberais.
$\bar{x}_A :=$ média salarial de administradores.
$s_A :=$ desvio padrão salarial de administradores.

Então $\bar{x}_L = 9,87$; $\bar{x}_A = 9,22$; $s_L = 2,43$; $s_A = 0,88$; $n_L = 7$; $n_A = 8$;
Primeiramente vamos fazer um teste para a variância.
Teste de igualdade de variâncias:

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

Estatística do teste: $W = S_L^2/S_A^2$. Sob H_0, $W \sim F(n_L - 1, n_A - 1) = F(6,7)$
Região Crítica: Observe que o teste é bilateral, Fixando $\alpha = 5\%$ (Consultando a Tabela F) então:

$$RC = \left\{ x \in \Re | x \in \left(0, \frac{1}{F_{6,7,2.5\%}}\right) \text{ ou } x > F_{6,7,2.5\%} \right\} = \left\{ x \in \Re | x \in \left(0, \frac{1}{5,12}\right) \text{ ou } x > 5,12 \right\}$$

Valor observado: $W_0 = \frac{S_L^2}{S_A^2} = \frac{2{,}43^2}{0{,}88^2} = 7,62$, como W_0 pertence a RC, então rejeitamos H_0, Devemos aplicar o teste T para variâncias diferentes, desconhecidas.

Teste de igualdade de médias com variâncias diferentes:

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

Estatística do teste:

$$T = \frac{\bar{x}_L - \bar{x}_A}{\sqrt{\left(\frac{S_L^2}{n_L} + \frac{S_A^2}{n_A}\right)}} \text{ com } v = \frac{(C+D)^2}{C^2/(n_L - 1) + D^2/(n_A - 1)} \text{ com } C = S_L^2/n_L \text{e } D = S_A^2/n_A.$$

.

$$\text{sob } H_0,\ T \sim t_v$$

Calculando v encontramos $v = 7,39 \cong 7$, então sob H_0, $T \sim t_7$
Região Crítica: Observe que o teste é bilateral, Fixando $\alpha = 5\%$ (Consultando a Tabela t) então:

$$RC = \{x \in \Re | x < -t_{7,2.5\%} \text{ ou } x > t_{7,97.5\%}\} = \{x \in \Re | x < -2,364 \text{ ou } x > 2,364\}$$

Valor observado:

$$T_0 = \frac{9,87 - 9,22}{\sqrt{\left(\frac{2,43^2}{7} + \frac{0,88^2}{8}\right)}} \cong 0,67$$

. como T_0 não pertence a RC, então não rejeitamos H_0, ou seja a um nível de significância de 5%, não rejeitamos a hipótese de igualdade de médias.

8. Considere
   $\bar{x}_a :=$ vida média do transistor do Tipo A.
   $s_a :=$ desvio padrão do tempo de vida do transistor A.
   $\bar{x}_b :=$ vida média do transistor do Tipo B.
   $s_b :=$ desvio padrão do tempo de vida do transistor B.

   Então $\bar{x}_a = 1000$; $\bar{x}_b = 910$; $s_a = 60$; $s_b = 40$; $n_a = 20$; $n_b = 20$;
   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, $W \sim F(n_a - 1, n_a - 1) = F(19, 19)$
   Região Crítica: Observe que o teste é bilateral, Fixando $\alpha = 10\%$ (Consultando a Tabela F) então:

$$RC = \left\{x \in \Re | x \in \left(0, \frac{1}{F_{19,19,2.5\%}}\right) \text{ ou } x > F_{19,19,2.5\%}\right\} = \left\{x \in \Re | x \in \left(0, \frac{1}{2,16}\right) \text{ ou } x > 2,16\right\}$$

   Valor observado: $W_0 = \frac{S_a^2}{S_b^2} = \frac{60^2}{40^2} = 2,25$, como W_0 pertence a RC, então rejeitamos H_0, Devemos aplicar o teste T para variâncias diferentes, desconhecidas.

   Teste de igualdade de médias com variâncias diferentes:

$$H_0 : \mu_a - \mu_b = 60 \quad \text{versus} \quad H_1 : \mu_a - \mu_b > 60$$

   Estatística do teste:

$$T = \frac{\bar{x}_a - \bar{x}_b}{\sqrt{\left(\frac{S_a^2}{n_a} + \frac{S_b^2}{n_b}\right)}} \text{ com } v = \frac{(C + D)^2}{C^2/(n_a - 1) + D^2/(n_b - 1)} \text{ com } C = S_a^2/n_a \text{e } D = S_b^2/n_b.$$

   .

$$\text{sob } H_0,\ T \sim t_v$$

   Calculando v encontramos $v \cong 33$, então sob H_0, $T \sim t_{33}$
   Região Crítica: Observe que o teste é unilateral, Fixando $\alpha = 5\%$ (Consultando a Tabela t) então:

$$RC = \{x \in \Re | x > t_{33,95\%}\} = \{x \in \Re | x > 1,69\}$$

Valor observado:

$$T_0 = \frac{1000 - 910 - 60}{\sqrt{\left(\frac{60^2}{20} + \frac{40^2}{20}\right)}} \cong 1,861$$

. como T_0 pertence a RC, então rejeitamos H_0, ou seja a um nível de significância de 5%, posso afirmar que a vida média da marca A é pelo menos 60 horas maior que a marca B.

9. Como o estudo é feito com os mesmos individuos, estamos diante do caso em que se pode usar variáveis emparelhadas. Considere
$\bar{d}$:= Média da diferença de temperatura entre depois (Y) e antes (X).
S_d := Desvio padrão da diferença de temperatura entre depois (Y) e antes (X).
d = difença entre o depois e antes.
Então

$$\bar{d} = \frac{(37,8 - 37,5) + (36,4 - 36) + \ldots + (37,3 - 38,1)}{10} = -0,59$$

$$S_d = \sqrt{\frac{\sum_{i=1}^{10}(d_i - (-0,59))^2}{9}} \cong 0,61 \quad \text{com} \quad d = (0.3, 0.4, \ldots, -0.8)$$

Teste t-pareado

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

Estatística do teste:

$$T = \frac{\sqrt{n} \cdot \bar{d}}{S_d} \;\; \text{sob} \;\; H_0,\, T \sim t_{(n-1)} = t_9$$

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

$$RC = \left\{x \in \Re | x < t_{(9;5\%)}\right\} = \{x \in \Re | x < -1,83\}$$

Valor observado:

$$T_0 = \frac{\sqrt{10} \cdot -0,59}{0,61} \cong -3,05$$

como T_0 pertence a RC, então rejeitamos H_0, ou seja a um nível de significância de 5%, o medicamento é eficiênte na diminuição da temperatura corporal (em graus Celsius).

10. Exemplos de amostras pareadas:
Estudo feito do antes e depois no mesmo indivíduo
Estudo feito com gêmeos.