MAE126 - Noções de Estatística II

Turma Prof. Popov

4 de junho de 2006

Correção da Lista 5

1. Sabe-se que, durante muitos anos, a altura média de habitantes adultos de uma região remota foi de 1,67m, com desvio padrão de 0,092m. Um pesquisador deseja verificar se, depois de contatos intensos com a civilização, a altura média mudou (presume-se que o desvio padrão contnua o mesmo). Para isso, ele escolheu uma amostra de 83 pessoas, e verificou que a altura média para essa amostra ficou 1,683m.

   (a) Formule este problema como teste de hipóteses (especificando quais são as hipóteses nula e alternativa).
   $H_0 : \mu = 1,67$ versus $H_a : \mu \neq 1,67$.

   (b) Quais são os significados dos erros dos tipos I e II?
   Erro Tipo I: concluir que a média não é 1,67 (rejeitar H_0) quando a média é 1,67 (H_0 verdadeira). Erro Tipo II: concluir que a média é 1,67 (aceitar H_0) quando a média não é 1,67 (H_0 falsa).

   (c) Construa a região crítica para o nível de significância $\alpha = 0,04$. Com base nesta região crítica, qual deve ser a decisão?
   Queremos uma região do tipo:

   $$RC = \{\overline{x} \leq x_{rc1} \text{ ou } \overline{x} \geq x_{rc2}\}$$

   Pela simetria da normal, vamos calcular somente x_{rc2}:

   $$\begin{aligned}
   P\left(\overline{X} \geq x_{rc2}\right) &= P\left(\frac{\overline{X} - 1,67}{0,092/\sqrt{83}} \geq \frac{x_{rc2} - 1,67}{0.092/\sqrt{83}}\right) \\
   &= P\left(Z \geq \frac{x_{rc2} - 1,67}{0,01}\right) = 0,02 \Rightarrow \\
   &= \frac{x_{rc2} - 1,67}{0,01} = 2,054 \Rightarrow x_{rc2} = 1,67 + 0,021 \Rightarrow \\
   &= x_{rc2} = 1,691, x_{rc1} = 1,649
   \end{aligned}$$

   Assim:

   $$RC = \{\overline{x} \leq 1,649 \text{ ou } \overline{x} \geq 1,691\}$$

   Como $\overline{x}_{obs} = 1,683 \notin RC$, não rejeitamos H_0, concluindo que a média de altura *não* mudou.

(d) Suponha que a altura média, de fato, mudou, e é agora 1,70m. Qual é a probabilidade do erro tipo II?

Pela definição de Erro Tipo II:

$$\begin{aligned}
P(\text{Erro Tipo II}) &= P(\overline{x} \in RC^C | H_0 \text{ falsa })\\
&= P(\overline{x} \in RC^C | \mu = 1,7) = P(1,649 \leq \overline{x} \leq 1,691 | \mu = 1,7)\\
&= P\left(\frac{1,649 - 1,7}{0,092/\sqrt{83}} \leq \frac{\overline{x} - 1,7}{0,092/\sqrt{83}} \leq \frac{1,691 - 1,7}{0,092/\sqrt{83}}\right)\\
&= P\left(-5,05 \leq Z \leq -0,891\right) = 0,186
\end{aligned}$$

2. Um candidato A afirma que a intenção de voto nele é de 53%. Um concorrente dele deseja contestar essa afirmação. para isso, o concorrente contrata uma empresa de pesquisas, que entrevista 300 pessoas, verificando que 151 tem a intenção de vota em A.

(a) Formule este problema como teste de hipóteses.

$H_0 : p = 0,53$, versus $H_a : p \leq 0,53$.

(b) Quais são os significados dos erros dos tipos I e II?

Erro Tipo I: concluir que a intenção de voto não é 0,53 quando na verdade ela é 0,53. Erro Tipo II: concluir que a intenção de voto é 0,53, quando na verdade ela é menor que 0,53.

(c) Construa a região crítica para o nível de significância $\alpha = 0,06$. Com base nesta região crítica, qual deve ser a decisão?

Queremos uma região do tipo:

$$RC = \{p_{obs} \leq p_c\}$$

$$\begin{aligned}
P(\text{Erro Tipo I}) &= P(\hat{p} \leq p_c | H_0 \text{ verdadeira }) = 0,06 \Rightarrow\\
&= P\left(\frac{\hat{p} - p}{\sqrt{p(1-p)/n}} \leq \frac{p_c - p}{\sqrt{p(1-p)/n}} | H_0 \text{ verdadeira}\right)\\
&= P\left(Z \leq \frac{p_c - 0,53}{\sqrt{0,53(0,47)/300}}\right)\\
&= P\left(Z \leq \frac{p_c - 0,53}{0,029}\right) = 0,06\\
&\Rightarrow \frac{p_c - 0,53}{0,029} = -1,555 \Rightarrow p_c = 0,53 - 0,045 = 0,485
\end{aligned}$$

Logo a região de rejeição é dada por:

$$RC = \{p_{obs} \leq 0,485\}$$

Como $p_{obs} = 151/300 = 0,503 \notin RC$, não rejeitamos a hipótese nula.

(d) Qual é a probabilidade do erro tipo II, se na verdade a intenção de voto em A é 49%?

$$\begin{aligned} P(\text{Erro Tipo II}) &= P(\hat{p} \in RC^C | H_0 \text{ falsa }) \\ &= P(\hat{p} > 0,485 | p = 0,49) \\ &= P\left(\frac{\hat{p} - 0,49}{\sqrt{0,49(0,51)/300}} > \frac{0,485 - 0,49}{\sqrt{0,49(0,51)/300}}\right) \\ &= P(Z > -0,173) = 0,568. \end{aligned}$$

3. Para os problemas 1 e 2, calcule o nível descritivo (p-valor, probabilidade de significância), e tome a decisão ao nível de significância de 1%.

Pr 1.

$$\begin{aligned} \alpha^* &= 2P(\overline{X} \geq 1,683 | H_0) \\ &= 2P\left(\frac{\overline{X} - 1,67}{0,092/\sqrt{83}} \geq \frac{1,683 - 1,67}{0,092/\sqrt{83}}\right) \\ &= 2P(Z \geq 1,287) = 0,198 \end{aligned}$$

Pr 2.

$$\begin{aligned} \alpha^* &= P(\hat{p} < 0,503 | H_0) \\ &= P\left(\frac{\hat{p} - 0,53}{\sqrt{0,53(0,47)/300}} < \frac{0,503 - 0,53}{\sqrt{0,53(0,47)/300}}\right) \\ &= P(Z < -0,937) = 0,174 \end{aligned}$$

4. Numa linha de produção, é importante que o tempo gasto numa determinada operação não varie muito de empregado para empregado. Especificamente, é considerado satisfatório que a variância deste tempo não ultrapasse 9 segundos^2. Desejando verificar se há sintonia entre os empregados, o empresário recolhe uma amostra destes tempos para 15 empregados: 20, 27, 16, 24, 25, 19, 29, 27, 21, 29, 23, 18, 29, 25, 29.

(a) Formule este problema como teste de hipóteses.
$H_0 : \sigma^2 \leq 9$, versus $H_a : \sigma^2 > 9$.

(b) Quais são os significados dos erros tipos I e II?
Erro Tipo I: concluir que a variância é maior que 9, quando na verdade ela é menor. Erro Tipo II: concluir que a variância é menor ou igual a 9, quando na verdade ela é maior.

(c) Construa a região crítica para o nível de significância $\alpha = 0,07$. Com base nesta região crítica, qual deve ser a decisão?
Queremos uma região crítica do tipo:

$$RC = \{S^2_{obs} > S^2_c\}$$

$$
\begin{aligned}
P(\hat{S}^2 > S_c^2) &= 0,07 \\
&= P\left(\frac{(n-1)\hat{S}^2}{\sigma_0^2} > \frac{14S_c^2}{9}\right) \\
&= P\left(\chi_{14} > \frac{14S_c^2}{9}\right) = 0,07 \\
&\Rightarrow \frac{14S_c^2}{9} = 22,44 \Rightarrow S_c^2 = 14,42
\end{aligned}
$$

Assim:

$$RC = \{S_{obs}^2 > 14,42\}$$

Como $S_{obs}^2 = 19,29 \in RC$, rejeitamos a hipótese nula, concluindo que a variância aumentou.

5. Sabemos que, há dois anos, a renda média mensal dos habitantes de uma cidade era de 590 R$. Queremos saber se agora a situação mudou (para melhor ou para pior). Para isso, selecionou-se uma amostra de tamanho 23, apresentando os seguintes resultados: 510, 730, 455, 650, 500, 470, 490, 690, 340, 420, 710, 560, 500, 550, 600, 590, 410, 470, 510, 380, 450, 490, 580.

   (a) Formule este problema como teste de hipóteses.
   $H_0 : \mu = 590$ versus $H_a : \mu \neq 590$.

   (b) Quais são os significados dos erros tipos I e II?
   Erro Tipo I: concluir que a renda média mensal mudou, quando ela não mudou. Erro Tipo II: concluir que a renda mensal média não mudou, quando ela mudou.

   (c) Construa a região crítica e tome a decisão.
   Queremos uma região crítica do tipo:

   $$RC = \{\overline{x} \leq x_{rc1} \text{ ou } \overline{x} \geq x_{rc2}\}$$

   Pela simetria da t de Student, e notando que $S_{obs} = 103,14$ vamos calcular somente x_{rc2}

   $$
   \begin{aligned}
   P\left(\overline{X} \geq x_{rc2}\right) &= P\left(\frac{\overline{X} - 590}{103,14/\sqrt{23}} \geq \frac{x_{rc2} - 590}{103,14/\sqrt{23}}\right) \\
   &= P\left(t_{(22)} \geq \frac{x_{rc2} - 590}{21,51}\right) = 0,025 \Rightarrow \\
   &= \frac{x_{rc2} - 590}{21,51} = 2,074 \Rightarrow x_{rc2} = 590 + 44,61 \Rightarrow \\
   &= x_{rc2} = 634,61, x_{rc1} = 545,39
   \end{aligned}
   $$

   Assim:

$$RC = \{\overline{x} \leq 545,39 \text{ ou } \overline{x} \geq 634,61\}$$

Como $\overline{x}_{obs} = 524,13 \in RC$, rejeitamos a hipótese nula, concluindo que a média de renda mensal mudou (piorou).