Errata

Aula de 27/10

Pág 2 3ª linha onde está $n=36$, correto é $n=30$

Pag 12 Linha 10
onde está $\bar{X} \sim N(0,1)$ o correto é $\frac{\bar{X}-10}{1} \sim N(0,1)$.

Pag.7 Linha 4
onde está $\bar{X} \leq c_1$ o correto é $\bar{X} \leq x_{c_1}$.

Aula de 23/10
Pag 11 3ª linha onde está Dist. de $\bar{X}$ $N\left(\mu, \frac{36}{\sqrt{30}}\right)$
o correto é $\bar{X}$ $N\left(\mu, \frac{36}{30}\right)$

Alterar a parte b) do ex. 24, pág 255, livro para
Repita o problema utilizando a proporção amostral obtida em a) mas supondo que a amostra foi de 360 dias. Compare as amplitudes dos dois intervalos (item a) e b))

Exercício

Seja X o custo de manutenção de um tear. Sabe-se que $X \sim N(\mu, 400)$. Para testar $H_0: \mu = 200$ contra $H_a: \mu > 200$, utilizou-se uma amostra de 25 teares.

a) Fixando α em 5%, encontre a região crítica.

b) Adotando a região crítica do item a), se $\mu = 210$, qual é a probabilidade desse fato não ser detectado?

c) Se o custo médio para a amostra de 25 teares foi de 205, qual será a conclusão?

$H_0: \mu = 200$

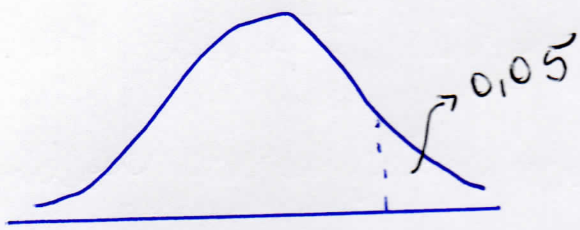
$H_a: \mu > 200$

$$\bar{X} \sim N\Big(\underbrace{\mu}_{\text{desconhecido}}, \underbrace{\frac{400}{25}}_{16}\Big) \qquad RC: \bar{X} \geq x_c$$

a) $P(\bar{X} \geq x_c \mid \mu = 200) = 0{,}05 = P\left(\frac{\bar{X} - 200}{4} \geq \frac{x_c - 200}{4}\right)$

$\mu = 200 \Rightarrow \bar{X} \sim N(200, 16)$

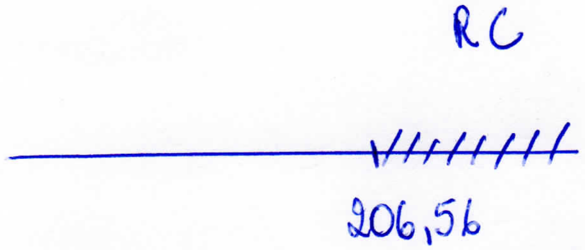
$= P\left(Z \geq \frac{x_c - 200}{4}\right)$



$\frac{x_c - 200}{4} = 1{,}64$

$$\frac{x_c - 200}{4} = 1{,}64$$

$x_c = 206{,}56$ $\qquad \Rightarrow RC: \bar{X} \geq 206{,}56$



b) $P(\bar{X} < 206{,}56 \mid \mu = 210) = P\left(\frac{\bar{X} - 210}{4} < \frac{206{,}56 - 210}{4}\right)$

$$\frac{\bar{X} - 210}{4} \sim N(0,1)$$

$= P(Z < -0{,}86) = P(Z > 0{,}86) = 0{,}5 - 0{,}3051 = 0{,}1949$

$= \beta(210) = P(\text{erro do tipo II} \mid \mu = 210)$

c) $\bar{x}_{obs} = 205 \notin RC$

Não rejeitamos H_0. Ao nível de significância de 0,05, não rejeitamos a hipótese de que o custo médio é 200.

Teste de Hipóteses para a proporção

Seja p - proporção de elementos em uma população portadores de uma certa característica

$H_0: p = p_0 \qquad 0 < p_0 < 1$

$H_a: p \neq p_0$

Estatística de teste:

$\hat{p}$ = proporção amostral

Distribuição de probabilidades de $\hat{p}$:

$$\hat{p} \approx N\left(p, \frac{p(1-p)}{n}\right)$$

Região crítica:

$\hat{p} \geq p_{c_2}$ ou $\hat{p} \leq p_{c_1}$ $\qquad p_{c_1}$ e p_{c_2} tais que

$$P(\hat{p} \geq p_{c_2} \mid p = p_0) = \frac{\alpha}{2} \qquad P(\hat{p} \leq p_{c_1} \mid p = p_0) = \frac{\alpha}{2}$$

Exemplo 8.4 Magalhães e Lima

Um relatório afirma que 40% da água obtida em poços artesianos do nordeste é salobra. Há controvérsias para mais e para menos. Para avaliar o fato, uma amostra de 400 poços foi sorteada e observou-se que em 120 deles a água era salobra. Qual é a conclusão ao nível de significância de 0,01?

p - proporção desconhecida de água salobra na região

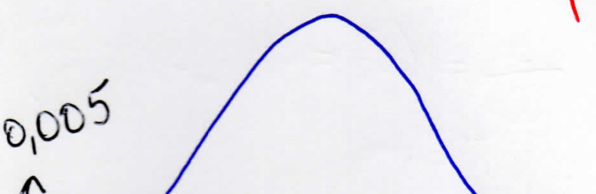
$H_0: p = 0{,}4$ $\qquad n = 400 \qquad \hat{p} \approx N\left(p, \dfrac{p(1-p)}{400}\right)$

$H_a: p \neq 0{,}4$

$RC: \hat{p} \geq p_{c_2}$ ou $\hat{p} \leq p_{c_1}$

$$P(\hat{p} \leq p_{c_1} \mid p = 0{,}4) = 0{,}005 = P\left(\overbrace{\frac{\hat{p} - 0{,}4}{\sqrt{\frac{0{,}24}{400}}}}^{Z \sim N(0,1)} \leq \frac{p_{c_1} - 0{,}4}{\sqrt{\frac{0{,}24}{400}}}\right)$$

$$\hat{p} \approx N\left(0{,}4, \frac{0{,}4 \cdot 0{,}6}{400}\right)$$



$$\frac{p_{c_1} - 0{,}4}{\sqrt{\frac{0{,}24}{400}}} = -2{,}58$$

$$p_{c_1} = 0{,}4 - 2{,}58 \cdot \frac{0{,}49}{20}$$

$$p_{c_1} = 0{,}336$$

Analogamente,

$$P\left(\hat{p} \geq pc_2 \mid p=0{,}4\right) = P\left(Z \geq \frac{pc_2 - 0{,}4}{\sqrt{\frac{0{,}24}{400}}}\right) = 0{,}005$$

$$\frac{pc_2 - 0{,}4}{\frac{0{,}49}{20}} = 2{,}58 \qquad pc_2 = 0{,}4 + 2{,}58 \cdot \frac{0{,}49}{20} = 0{,}463$$

RC: $\hat{p} \leq 0{,}336$ ou $\hat{p} \geq 0{,}463$

$\hat{p}_{obs} = \frac{120}{400} = 0{,}3$ $\quad \hat{p}_{obs} \in$ R.C. rejeitamos H_0

Ao nível de significância de 0,01, os dados sugerem que a proporção de água salobra não é 0,4.

Este foi o teste bicaudal

Teste Unicaudal à esquerda para a proporção

$H_0: p = p_0 \qquad 0 < p_0 < 1 \qquad RC: \hat{p} \leq p_c$

$H_a: p < p_0$

Um fabricante garante que pelo menos 90% dos equipamentos que fornece estão dentro das especificações. A análise de uma amostra de 200 peças revelou 175 peças dentro das especificações. Testar a afirmação do fabricante ao nível de significância de 0,01.

p : proporção desconhecida de peças dentro das especificações

$H_0: p = 0{,}9 \quad (\Leftrightarrow p \geq 0{,}9) \rightarrow$ afirmação do fabricante é verdadeira

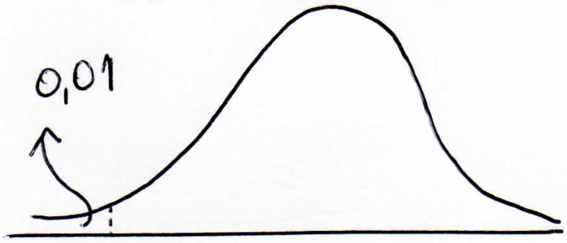
$H_a: p < 0{,}9 \rightarrow$ afirmação do fabricante é falsa

$RC: \hat{p} \leq p_c \qquad \hat{p} \overset{\approx}{\sim} N\left(p, \frac{p(1-p)}{n}\right)$

$$\alpha = 0{,}01 = P(\hat{p} \leq p_c \mid p = 0{,}9) = P\left(\frac{\hat{p} - 0{,}9}{\sqrt{\frac{0{,}9 \cdot 0{,}1}{200}}} \leq \frac{p_c - 0{,}9}{\sqrt{\frac{0{,}9 \cdot 0{,}1}{200}}}\right)$$

$$\hat{p} \overset{\approx}{\sim} N\left(0{,}9, \frac{0{,}1 \cdot 0{,}9}{200}\right)$$

$= P\left(Z \leq \frac{p_c - 0{,}9}{0{,}021}\right)$

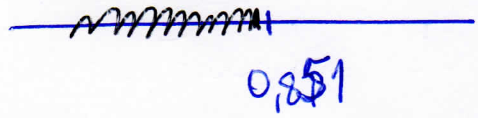


$Z \sim N(0,1)$

$\frac{p_c - 0{,}9}{0{,}021} = -2{,}33$

$\Rightarrow p_c = 0{,}851$

RC: $\hat{p} \leq 0{,}851$



$\hat{p}_{obs} = \frac{175}{200} = 0{,}875$

$\hat{p}_{obs} \notin RC$

Não rejeitamos Ho. Os dados não nos levam a duvidar da informação do fabricante.

$0{,}01 = P($Concluir que o fabricante não fala a verdade | quando fala a verdade$)$

Com essa regra de decisão, qual é a probabilidade de considerarmos o fabricante honesto quando na verdade ele não é e a proporção de peças dentro da especificação é 0,8?

$\beta(0{,}8) = P(\text{aceitar } H_0 \mid H_0 \text{ é falsa, } p = 0{,}8) =$

$P(\hat{p} > 0{,}851 \mid p = 0{,}8) = P\left(\frac{\hat{p} - 0{,}8}{\sqrt{0{,}0008}} > \frac{0{,}851 - 0{,}8}{\sqrt{0{,}0008}}\right) =$

$\hat{p} \approx N\left(0{,}2 \,;\, \frac{0{,}2 \cdot 0{,}8}{200}\right)$

$$= P\left(Z > 1{,}82\right) = 0{,}0344$$

Esta é a probabilidade do erro do tipo II para $p = 0{,}8$ (valor de p fixado em H_a) $\begin{cases} H_0: p = 0{,}9 \\ H_a: p < 0{,}9 \end{cases}$

O problema da parada desnecessária é um problema de teste de hipóteses

Procedimento de controle de qualidade planejado para máximo 10% itens defeituosos.
Sorteia-se amostra de 100 peças, havendo mais de 15% defeituosas pára-se a produção.
Probabilidade da parada desnecessária.

p - proporção desconhecida de peças defeituosas

$H_0: p = 0{,}10$ $(p \leq 0{,}10)$ produção sob controle

$H_a: p > 0{,}10$ produção fora de controle

$$\hat{p} = \frac{X}{n} = \frac{n^{\underline{o}}\ \text{defeituosas}}{100}$$

Parar a produção $\Longleftrightarrow$ rejeitar H_0 $\Longleftrightarrow$ $\hat{p} > 0{,}15$

RC: $\hat{p} > 0{,}15$

Parada desnecessária $\Longleftrightarrow \hat{p} > 0{,}15$ quando $p = 0{,}10$

$\Longleftrightarrow$ Rejeitar H_0 quando H_0 é verdadeira $\Longleftrightarrow$

erro do tipo I

$P(\text{parada desnecessária}) = \alpha = P(\text{erro do tipo I})$

Na aula de exercícios foi calculada para

$n = 40 \qquad \alpha = 0{,}1469$

$n = 60 \qquad \alpha = 0{,}0985$

Exercício: Calcular para $n = 100$.

Teste de Igualdade de Proporções associadas a duas populações (exemplo 9.9 - Magalhães e Lima)

Deseja-se avaliar a eficiência de um novo método de ensino com relação ao método tradicional. Para isso tomou-se duas amostras similares de 100 alunos cada. Em uma das amostras, os alunos foram submetidos ao método novo e na outra ao método tradicional. Posteriormente, os alunos foram submetidos a uma prova e foram estimadas as probabilidades de "bom desempenho" em ambos os casos. Formule o problema como um teste de hipóteses.

p_1 - probabilidade de "bom desempenho" na população "hipotética" de todos os alunos sujeitos ao método novo

p_2 - " " " " tradicional

$H_0: p_1 = p_2$ método novo é equivalente ao tradicional

$H_a: p_1 > p_2$ método novo é superior ao tradicional

n_1 - tamanho da amostra da pop. 1 (Novo)

n_2 - tamanho da amostra da pop. 2 (Tradicional)

$$\hat{p}_1 = \frac{X_1}{n_1} \qquad \hat{p}_2 = \frac{X_2}{n_2}$$

X_1 - nº de alunos com bom desempenho na amostra da pop 1 (método novo)

X_2 - nº de alunos com bom desempenho na amostra da pop 2 (método tradicional)

Construção Teórica do Teste

Parâmetro de interesse: $p_1 - p_2$

Estimador: $\hat{p}_1 - \hat{p}_2$

$$\hat{p}_1 \overset{\sim}{\approx} N\left(p_1, \frac{p_1(1-p_1)}{n_1}\right) \qquad \hat{p}_2 \overset{\sim}{\approx} N\left(p_2, \frac{p_2(1-p_2)}{n_2}\right)$$

$$E(\hat{p}_1 - \hat{p}_2) = E(\hat{p}_1) - E(\hat{p}_2) = p_1 - p_2$$

Propriedades já vistas de $\hat{p}$

$$Var(\hat{p}_1 - \hat{p}_2) = Var(\hat{p}_1) + Var(\hat{p}_2) =$$

independência das amostras

$$= \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2} = V$$

$$\hat{p}_1 - \hat{p}_2 \overset{\sim}{\approx} N(p_1 - p_2, V)$$

Sob H_0, $p_1 = p_2 = p$, utiliza-se o estimador da proporção comum p: $\hat{p} = \dfrac{X_1 + X_2}{n_1 + n_2} = \dfrac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2}$

Também sob $H_0: p_1 = p_2 = p$

$$V = \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2} = \frac{p(1-p)}{n_1} + \frac{p(1-p)}{n_2}$$

$$= p(1-p)\left[\frac{1}{n_1} + \frac{1}{n_2}\right] \quad e$$

$$\hat{V} = \hat{p}(1-\hat{p})\left[\frac{1}{n_1} + \frac{1}{n_2}\right]$$

Verifica-se ainda que

$$\frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\hat{V}}} \overset{\sim}{\approx} N(0,1)$$

Sob H_0, $p_1 - p_2 = 0$ e $Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{V}}} \overset{\sim}{\approx} N(0,1)$ $\square$

Grandes valores de $Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{V}}}$ sugerem que H_a é verdadeira.

$\therefore$ Rejeita-se H_0 se $Z \geq z_c$, z_c |

$$P(Z \geq z_c) = \alpha \qquad Z \sim N(0,1)$$

No exemplo, se 75 alunos sujeitos ao método novo e 65 sujeitos ao método tradicional obtiveram bom desempenho, qual é a conclusão ao nível de significância de 0,05?

$$\hat{p}_1 = 0{,}75 \qquad \hat{p}_2 = 0{,}65$$

$$\hat{p} = \frac{n_1 \hat{p}_1 + n_2 \hat{p}_2}{n_1 + n_2} = \frac{100.0{,}75 + 100.0{,}65}{200} = 0{,}7$$

$$\hat{V} = 0{,}7 . 0{,}3 \left(\frac{1}{100} + \frac{1}{100} \right) = 0{,}0042$$

$$\alpha = 0{,}05 \implies z_c = 1{,}64 \qquad RC: Z \geq 1{,}64$$

$$Z_{obs} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{V}}} = \frac{0{,}75 - 0{,}65}{\sqrt{0{,}0042}} = 1{,}543 \qquad Z_{obs} \notin RC$$

Não rejeitamos H_0. Ao nível de significância de 0,05, não rejeitamos a hipótese de equivalência dos métodos.