MAE 229 - Introdução à Probabilidade e Estatística II
Resolução Lista 3

Professor: Pedro Morettin

Exercício 1

(a) A hipótese nula H_0 é de que a média de vendas μ permanece inalterada, enquanto que a hipótese alternativa H_1 é de que houve melhora nas vendas, i.e.:

$$H_0 : \mu = 320$$

$$H_1 : \mu > 320$$

(b) Neste teste a estatística considerada é a média amostral $\bar{X}$. A região crítica do teste é: $RC = \{\bar{X} \geqslant k\}$, em que k é tal que $P(\bar{X} \geqslant k|\mu = 320) = 0,05$. Ou seja, estamos fixando a probabilidade de erro tipo I em 5%.

Sob H_0, temos que $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$ para n suficientemente grande (aproximação normal). Neste caso, $n = 25$, $\mu = 320$ e $\sigma = 40$. Assim, sob H_0 temos que:

$$P\left(\frac{\bar{X} - 320}{40/\sqrt{25}} \geqslant \frac{k - 320}{40/\sqrt{25}}\right) = P\left(Z \geqslant \frac{k - 320}{8}\right) = 0,05.$$

Assim, $k = 333,16$ e rejeitamos H_0, se $\bar{X} \geqslant 333,16$.

(c) Para calcular o nível descritivo do teste temos calcular a probabilidade de se observar valores mais extremos do que o encontrado na amostra, supondo que a hipótese nula seja verdadeira:

$$P(\bar{X} \geqslant 335) = P(Z \geqslant \frac{335 - 320}{8}) = 0,0304$$

Assim, temos que o valor-p é de 3,04%, sendo menor que o nível de significância do teste $\alpha = 5\%$. Com isso, rejeitamos H_0. Logo, há evidências de que as vendas melhoraram. Chegaríamos a mesma conclusão usando o valor de k encontrado na letra (b), uma vez que $335 > 333,16$.

Exercício 2

(a) A hipótese nula H_0 é de que a proporção de pares defeituosos p continua em 10%, enquanto que a hipótese alternativa é de que houve piora, i.e:

$$H_0 : p = 0,10$$

$$H_1 : p > 0,10$$

(b) Neste teste a estatística considerada é a proporção amostral $\hat{p}$. A região crítica do teste é: $RC = \{\hat{p} \geqslant k\}$, em que k é tal que $P(\hat{p} \geqslant k|p = 0,10) = 0,09$. Ou seja, estamos fixando a probabilidade de erro tipo I em 9%.

Sob H_0, temos que $\hat{p} \sim N(p, \frac{p(1-p)}{n})$ para n suficientemente grande (aproximação normal). Neste caso, $n = 100$. Assim, sob H_0 temos que:

$$P\left(\frac{\hat{p} - 0,10}{0,03} \geqslant \frac{k - 0,10}{0,03}\right) = P\left(Z \geqslant \frac{k - 0,10}{0,03}\right) = 0,09.$$

Assim, $k = 0,1402$ e rejeitamos H_0, se $\hat{p} \geqslant 0,1402$.

(c) As remessas cujas proporções são 25%, 16%, 24% e 21% devem ser rejeitadas, uma vez que há evidências de que o processo piorou.

Exercício 3

A região crítica do teste é: $RC = \{\hat{p} \geqslant 0,20\}$, em que $P(\hat{p} \geqslant 0,20|p = 0,10) = \alpha$. Ou seja, estamos querendo encontrar a probabilidade de erro tipo I α dada esta região crítica, sendo que agora $n = 10$, ou seja, $\frac{p(1-p)}{n} = 0,009$.

$$P\left(\frac{\hat{p} - 0,10}{\sqrt{0,009}} \geqslant \frac{0,20 - 0,10}{\sqrt{0,009}}\right) = P\left(Z \geqslant 1,054\right) = 0,146.$$

Assim, para esta região crítica temos um nível de significância para o teste de 14,6%.

Exercício 4

(a) A hipótese nula H_0 é de que a altura média μ continua em até 8,5, enquanto que a hipótese alternativa é de aumentou, i.e:

$$H_0 : \mu \leqslant 8,5$$

$$H_1 : \mu > 8,5$$

Neste teste a estatística considerada é a média amostral $\bar{X}$. A região crítica do teste é: $RC = \{\bar{X} \geqslant k\}$, em que k é tal que $P(\bar{X} \geqslant k|\mu \leqslant 8,5) = 0,05$. Ou seja, estamos fixando a probabilidade de erro tipo I em 5%.

Sob H_0, temos que $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$ para n suficientemente grande (aproximação normal). Neste caso, $n = 100$, $\mu = 8,5$ e $\sigma = 1$. Assim, sob H_0 temos que:

$$P\left(\frac{\bar{X} - 8,5}{1/\sqrt{100}} \geqslant \frac{k - 8,5}{1/\sqrt{100}}\right) = P\left(Z \geqslant \frac{k - 8,5}{0,10}\right) = 0,05.$$

Assim, $k = 8,665$ e rejeitamos H_0, se $\bar{X} > 8,665$.

(b) Sim, rejeitamos a hipótese nula, uma vez que $8,8 > 8,665$, concluindo que há evidências do aumento na altura média, dado este valor de $\bar{X}$.

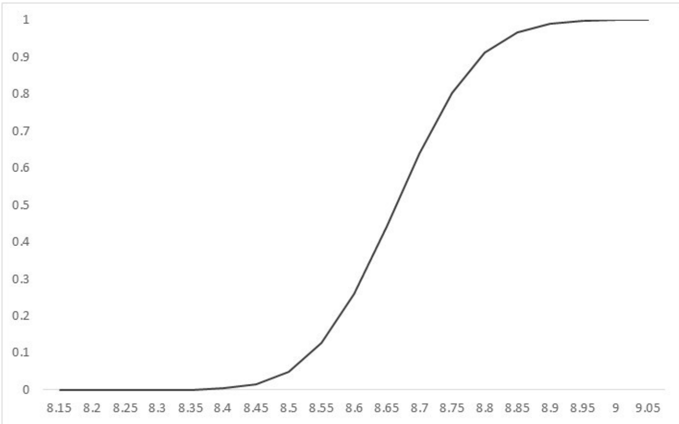
(c) A função poder do teste é definida como sendo:

$$\pi(\theta) = P(\hat{\theta} \in RC|\theta)$$

sendo θ o parâmetro a ser testado e $\hat{\theta}$ a estatística do teste. Neste teste específico, temos que desenhar o gráfico de:

$$\pi(\mu) = P(\bar{X} \in RC|\mu)$$

para diferentes valores de μ, sendo $RC = \{\bar{X} : \bar{X} \geqslant 8,665\}$.



(d) Queremos saber a probabilidade de não rejeitar H_0 ($\bar{X} \notin RC$) dado um certo valor para μ, ou seja, queremos saber $1 - \pi(\mu)$. Usando os valores para a função poder do teste obtidos e colocados em um gráfico na letra (c), temos que essas probabilidades são aproximadamente e respectivamente: (i) 0,558; (ii) 0,088 e (iii) 0. Note que quanto maior a média populacional, menor é a chance de não detectar uma melhora na altura média.

Exercício 5

A hipótese nula H_0 é de que a proporção p de pessoas que achou a pílula de açúcar (placebo) mais eficiente do que o medicamento é menor ou igual a 50 %, enquanto que a hipótese alternativa é de que esta proporção é maior que 50%, i.e:

$$H_0 : p \leqslant 0,5$$

$$H_1 : p > 0,5$$

Neste teste a estatística considerada é a proporção amostral $\hat{p}$. A região crítica do teste é: $RC = \{\hat{p} \geqslant k\}$, em que k é tal que $P(\hat{p} \geqslant k|p = 0,50) = 0,05$. Ou seja, estamos fixando a probabilidade de erro tipo I em 5%.

Sob H_0, temos que $\hat{p} \sim N(p, \frac{p(1-p)}{n})$ para n suficientemente grande (aproximação normal). Neste caso, $n = 20$. Assim, sob H_0 temos que:

$$P\left(\frac{\hat{p} - 0,50}{0,1118} \geqslant \frac{k - 0,50}{0,1118}\right) = P\left(Z \geqslant \frac{k - 0,50}{0,1118}\right) = 0,05.$$

Assim, $k = 0,6839$ e rejeitamos H_0, se $\hat{p} \geqslant 0,6830$. No caso, $\hat{p} = 15/20 = 0,75 > 0,6830$, logo rejeitamos H_0. Temos evidência, portanto, para a favor da afirmação do psiquiatra.

Exercício 6

Neste caso, queremos testar:

$$H_0 : \mu = 225$$

$$H_1 : \mu < 225$$

Sabemos que $\bar{X} = 200$, $n = 36$ e $s = 26$, sendo s^2 a variância amostral. Neste caso, a variância populacional é desconhecida, portanto, usamos a seguinte estatística de teste:

$$\frac{\bar{X} - \mu}{s/\sqrt{n}} \sim t(n-1).$$

Usando a tabela da distribuição t de Student com 35 graus de liberdade, temos:

$$P\left(\frac{\bar{X} - 225}{4,333} \leqslant \frac{k - 225}{4,333}\right) = P\left(t \leqslant \frac{k - 225}{4,333}\right) = 0,05.$$

Assim, $k = 217,6785$, e a região crítica do teste é $RC = \{\bar{X} : \bar{X} \leqslant 217,6785\}$. Como $\bar{X} = 200$, rejeitamos H_0, concluindo que há evidências para dizer que o conteúdo médio líquido é menor que 225 ml.

Exercício 7

(a) Denotando por X o consumo de gasolina em km/l, a hipótese a ser testada é:

$$H_0 : \mu = 15$$
$$H_1 : \mu = 14 < 15$$

Em particular, se H_0 é verdadeira, $X \sim N(15, 9)$ e $\bar{X} \sim N(15, 9/25)$. A região crítica deste teste é do tipo $RC = \{\bar{X} : \bar{X} \leqslant k\}$. Assim, a probabilidade do erro tipo I é:

$$P(\bar{X} \leqslant k | \bar{X} \sim N(15, 9/25)) = 0,06 = P\left(Z \leqslant \frac{k - 15}{\sqrt{9/25}}\right).$$

Portanto, temos $\frac{k-15}{\sqrt{9/25}} = -1,5547 \Rightarrow k = 14,0672$. Como $\bar{X} = 14,3 > 14,0672$, não rejeitamos H_0 a este nível de significância.

(b) A probabilidade de erro tipo II é a probabilidade de não rejeitar H_0, enquanto H_0 é falsa, i.e.:

$$P(\bar{X} > 14,0672 | \bar{X} \sim N(14, 9/25)) = P\left(Z > \frac{14,0672 - 14}{\sqrt{9/25}}\right) = P(Z > 0,1118) = 0,4554.$$

Exercício 8

Neste caso, queremos testar:

$$H_0 : \mu = 10$$

$$H_1 : \mu > 10$$

Sabemos que $\bar{X} = 11$, $n = 8$ e $s = 0,5$, sendo s^2 a variância amostral. Neste caso, a variância populacional é desconhecida, portanto, usamos a seguinte estatística de teste:

$$\frac{\bar{X} - \mu}{s/\sqrt{n}} \sim t(n-1).$$

Usando a tabela da distribuição t de Student com 7 graus de liberdade, temos:

$$P\left(\frac{\bar{X} - 10}{0,1767} \geqslant \frac{k - 10}{0,1767}\right) = P\left(t \geqslant \frac{k - 10}{0,1767}\right) = 0,05.$$

Assim, $k = 10,335$, e a região crítica do teste é $RC = \{\bar{X} : \bar{X} \geqslant 10,335\}$. Como $\bar{X} = 11$, rejeitamos H_0, concluindo que há evidências para dizer que houve aumento da produtividade com o novo fertilizante.

Exercício 9

Neste caso, queremos testar:

$$H_0 : \mu = 13$$
$$H_1 : \mu \neq 13$$

Sob H_0, a resistência dos cabos denotada por $X \sim N(13, 36)$ e $\bar{X} \sim N(13, 36/25)$. A região crítica deste teste é do tipo $RC = \{\bar{X} : \bar{X} > k_1 \cup \bar{X} < k_2\}$. Dado o nível de significância de 2%, temos que:

$$\begin{aligned}
P(\bar{X} > k_1 \cup \bar{X} < k_2 | X \sim N(13, 36/25)) &= P\left(Z > \frac{k_1 - 13}{\sqrt{36/25}} \cup Z < \frac{k_2 - 13}{\sqrt{36/25}}\right) = 0,02 \\
&= P(Z < -2,326 \cup Z > 2,326) \\
&= P(Z < -2,326) + P(Z > 2,326)
\end{aligned}$$

Assim, $k_2 = 10,208$ e $k_1 = 15,792$ e $RC = \{\bar{X} : \bar{X} > 15,792 \cup \bar{X} < 10,208\}$. Como $\bar{X} = 9,8 < 10,208 \in RC$, rejeitamos H_0. Temos evidência, portanto, que a resistência média é diferente de 13 Kgf.

Exercício 10

Neste caso, queremos testar:

$$H_0 : \mu = 100$$
$$H_1 : \mu < 100$$

Sabemos que $\bar{X} = 85$, $n = 16$ e $s = 12$, sendo s^2 a variância amostral. Neste caso, a variância populacional é desconhecida, portanto, usamos a seguinte estatística de teste:

$$\frac{\bar{X} - \mu}{s/\sqrt{n}} \sim t(n-1).$$

Usando a tabela da distribuição t de Student com 15 graus de liberdade, temos:

$$P\left(\frac{\bar{X} - 100}{3} \leqslant \frac{k - 100}{3}\right) = P\left(t \leqslant \frac{k - 100}{3}\right) = 0,05.$$

Assim, $k = 94,74$, e a região crítica do teste é $RC = \{\bar{X} : \bar{X} \leqslant 94,74\}$. Como $\bar{X} = 85$, rejeitamos H_0, concluindo que há evidências para dizer que houve melhora no processo a este nível de significância.

O IC para μ, considerando $\gamma = 0,9$ é:

$$\left[\bar{X} - \frac{t_\gamma s}{\sqrt{n}}; \bar{X} + \frac{t_\gamma s}{\sqrt{n}}\right]$$

em que $P(-t_\gamma \leqslant t \leqslant t_\gamma) = 0,9$.

Substituindo pelos valores de $\bar{X} = 85$, $n = 16$, $s = 12$ e $t_\gamma = 1,753$, temos que o $IC = [79,741; 90,259]$.

Exercício 11

(a) Usando a região crítica do enunciado, temos que:

$$P(\bar{X} \geqslant 1170|\mu = 1150; \sigma = 150) = \alpha.$$

Usando a aproximação normal, sob H_0, $X \sim N(1150, \frac{150^2}{100})$ e temos que:

$$P(\bar{X} \geqslant 1170|\bar{X} \sim N(1150, 225)) = P\left(Z \geqslant \frac{1170 - 1150}{\sqrt{225}}\right) = P(Z > 1,333) = 0,0912.$$

Assim, o nível de significância ou probabilidade de erro tipo I $\alpha = 9,12\%$.

(b) Temos que calcular a probabilidade de $\bar{X} \notin RC$, dado que H_1 é verdadeira:

$$P(\bar{X} < 1170|\mu = 1200; \sigma = 200) = \beta.$$

Usando a aproximação normal, sob $H - 1$, $X \sim N(1200, \frac{200^2}{100})$ e temos que:

$$P(\bar{X} < 1170|\bar{X} \sim N(200, 400)) = P\left(Z < \frac{1170 - 1200}{\sqrt{400}}\right) = P(Z < -1,5) = 0,0668.$$

Assim, a probabilidade de erro tipo II $\beta = 6,68\%$.

(c) Para que $\beta = \alpha$ teríamos que ter:

$$P(\bar{X} < k|\bar{X} \sim N(200, 400)) = P(\bar{X} \geqslant k|\bar{X} \sim N(1150, 225))$$

ou ainda

$$P\left(Z < \frac{k - 1200}{\sqrt{400}}\right) = P\left(Z \geqslant \frac{k - 1150}{\sqrt{225}}\right)$$

O valor aproximado de k que satisfaz a equação acima é $k = 1171,43$, ou seja, $RC = \{\bar{X} : \bar{X} \geqslant 1171,43\}$ Neste caso, $\beta = \alpha = 7,65\%$.

Exercício 12

Queremos testar:

$$H_0 : \mu = 100$$

$$H_1 : \mu > 100$$

A estatística considerada é a média amostral $\bar{X}$. A região crítica do teste é: $RC = \{\bar{X} \geqslant k\}$, em que k é tal que $P(\bar{X} \geqslant k|\mu = 100) = \alpha$. Ou seja, estamos fixando a probabilidade de erro tipo I em α.

Conforme enunciado,$P(\bar{X} < k|\mu = 100) = 0,95$ e $P(\bar{X} < k|\mu = 102) = 0,01$. Logo, $\alpha = 0,05$, uma vez que $\alpha = 1 - P(\bar{X} < k|\mu = 100)$.

Também foi enunciado que $P(\bar{X} < k|\mu = 102) = 0,01$. Note que esta equivale à probabilidade β de erro tipo II. Temos, assim, que o poder do teste é $1 - \beta = 0,99$.

Conforme já visto, sabemos que $\bar{X} \sim N(100, 9)$, sob H_0, e $\bar{X} \sim N(102, 9)$, sob H_1. Assim, temos duas equações com 2 incógnitas: n e k:

$$P\left(\frac{\bar{X} - 100}{3/\sqrt{n}} \geqslant \frac{k - 100}{3/\sqrt{n}}\right) = P\left(Z \geqslant \frac{\sqrt{n}(k - 100)}{3}\right) = 0,05$$

$$P\left(\frac{\bar{X} - 102}{3/\sqrt{n}} \geqslant \frac{k - 102}{3/\sqrt{n}}\right) = P\left(Z \geqslant \frac{\sqrt{n}(k - 102)}{3}\right) = 0,99$$

Segue que:

$$\begin{cases} \dfrac{\sqrt{n}(k - 100)}{3} = 1,645 \\ \dfrac{\sqrt{n}(k - 102)}{3} = -2,326 \end{cases}$$

Resolvendo o sistema, temos que $k = 100,82$ e $n \approx 36$.