MAE0229 – Introdução à Probabilidade e à Estatística II

I semestre de 2018

Lista de exercícios 4 - Testes de hipóteses (1 população) - C A S A (gabarito)

Exercício 1:

Sabe–se que o consumo mensal per capita de um determinado produto tem distribuição normal, com desvio padrão de 2 kg. A diretoria de uma firma que fabrica esses produto resolveu que retiraria o produto da linha de produção se a média do consumo per capita fosse menor que 8 kg. Caso contrário, continuaria a fabricá–lo. Foi realizada uma pesquisa de mercado, tomando–se uma amostra de 25 indivíduos, e verificou-se que $\sum_{i=1}^{25} X_i = 180kg$, em que X_i representa o consumo mensal do i–ésimo indivíduo da amostra.

(a) Construa um teste de hipóteses adequado, utilizando $\alpha = 5\%$, e com base na amostra colhida determine a decisão a ser tomada pela diretoria.

(b) Qual a probabilidade β de se tomar uma decisão errada se, na realidade, a média populacional for $\mu = 7.8$ kg?

(c) Se fosse fixado $\alpha = 1\%$, a decisão da diretoria seria a mesma? E se fosse fixado $\alpha = 10\%$? (justifique as suas respostas.)

(d) Se o desvio populacional fosse 4kg, qual seria a decisão, com $\alpha = 5\%$? (Justifique a sua resposta.)

Solução:

Seja X : "Consumo mensal percapita de um determinado produto". $X \sim N(\cdot; 2^2 kg^2)$

(a) Considere

$$\begin{cases} H_0: & \mu \geq 8kg, \text{ a média do consumo per capita é maior ou igual a 8kg.} \\ H_1: & \mu < 8kg, \text{ a média do consumo per capita é menor que 8kg.} \end{cases}$$

A estatística de teste é dada por

$$Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \underset{\mu=\mu_0=8}{\sim} N(0,1).$$

A região crítica é

$$\text{RC} = \{\overline{X} : \overline{X} \leq c\}$$

como $\alpha = 0,05$, tem–se que $z = 1,645$, portanto

$$\begin{aligned} c &= \mu_0 - z_{(1-\alpha)}\sqrt{\frac{\sigma^2}{n}} \\ &= 8 - 1,645\frac{2}{5} \\ &= 7,342. \end{aligned}$$

Assim,

$$\text{RC} = \{\overline{X} : \overline{X} \leq 7,342\},$$

como $\overline{x} = 7,2 \in \text{RC}$, se rejeita a hipótese nula, logo, ao nível de significância do 5% pode–se concluir que a diretoria retiraria o produto da linha de produção.

(b) Se $\mu = 7,8kg$, se quer saber $P(\text{Não rejeitar } H_0 | H_0 \text{ é falsa}) = \beta(7,8)$, lembre que sob H_1 $\overline{X} \sim N(7,8; 2^2kg^2/25)$.

$$\begin{aligned} P(\overline{X} > 7,342 | \mu = 7,8) &= P\left(\frac{\overline{X} - 7,8}{2/5} > \frac{7,342 - 7,8}{2/5}\right) \\ &= P\left(Z > \frac{-0,458}{2/5}\right) \\ &= P(Z > -1,145) \\ &= P(Z \leq 1,145) \\ &= 0,874, \end{aligned}$$

logo $\beta(7,8) = 87,4\%$.

(c)
1. Se $\alpha = 1\%$ fixado, $z = 2,33$

$$\text{RC} = \{\overline{X} : \overline{X} \leq c\},$$

em que

$$\begin{aligned} c &= \mu_0 - z_{(1-\alpha)}\sqrt{\frac{\sigma^2}{n}} \\ &= 8 - 2,33\frac{2}{5} \\ &= 7,068. \end{aligned}$$

Portanto $\text{RC} = \{\overline{X} : \overline{X} \leq 7,068\}$, como $\overline{x} = 7,2$ é maior que 7,068, $\overline{x} \notin \text{RC}$, logo pode–se concluir ao nível de significância do 1% que a diretoria não deveria retirar o produto.

2. Se $\alpha = 10\%$ fixado, tem–se que, $z = 1,282$,

$$\begin{aligned} c &= 8 - 1,282\frac{2}{5} \\ &= 7,4872, \end{aligned}$$

logo, $\text{RC} = \{\overline{X} : \overline{X} \leq 7,4872\}$, como $\overline{x} = 7,2$ é menor que 7,4872, $\overline{x} \in \text{RC}$ logo ao nível de significância do 10% pode–se concluir que a diretoria deve retirar o produto.

(d) Se $\sigma = 4kg$ e $\alpha = 5\%$, $z = 1,645$, portanto RC $= \{\overline{X} : \overline{X} \leq c\}$, em que

$$\begin{aligned} c &= \mu_0 - z_{(1-\alpha)}\sqrt{\frac{\sigma^2}{n}} \\ &= 8 - 1,645\frac{4}{5} \\ &= 6,684, \end{aligned}$$

logo, RC $= \{\overline{X} : \overline{X} \leq 6,684\}$, como $\overline{x} = 7,2 > 6,684$, $\overline{x} \notin$ RC, logo ao nível de significância do 5% pode–se concluir que a diretoria não deve retirar o produto.

Exercício 2:

Uma empresa de comidas rápidas registra o ingresso médio bruto de R\$3000, 00 por dia. Para saber se uma crise econômica tem afetado os ingressos, o departamento financeiro encomendou um estudo que registrou os ingressos de 8 dias seguidos. Os valores registrados foram: R\$3050, 00; R\$3212, 00; R\$2880, 00; R\$3121, 00; R\$3205, 00; R\$3018, 00; R\$2980, 00; R\$3188, 00.

(a) Formule este problema como um problema de teste de hipóteses.

(b) Os dados são suficientemente significativos, ao nível de significância de 5%, para provar que houve uma mudança?

(c) E ao nível de significância de 1%?

(d) Calcule o nível descritivo e interprete.

Solução:

Seja X : "ingresso bruto diário de uma empresa de comidas rápidas". Admitindo que $X \sim N(\mu, \sigma^2)$, com $\sigma^2 < \infty$.

(a) Se deseja testar

$$\begin{cases} H_0: & \mu \geq 3000, \text{ a crise não afetou os ingressos.} \\ \\ H_1: & \mu < 3000, \text{ a crise afetou os ingressos.} \end{cases}$$

(b) Note que aqui a variância populacional não é conhecida e o tamanho amostral é pequeno, a variância deve ser estimada da amostra, assim como a média, logo

$$\overline{x} = 3081, 75 \text{ e } s = 120, 1865.$$

Será considerada a seguinte estatística de teste

$$T = \frac{\overline{X} - \mu_0}{S/\sqrt{n}} \underset{\mu=\mu_0=3000}{\sim} t_{n-1}$$

como $\alpha = 5\%$, na tabela da t–Student do livro do Morettin $5\% = p/2$, logo, $p = 10\%$, assim $t_7 = 1,895$, portanto RC $= \{\overline{X} : \overline{X} < c\}$, em que

$$\begin{aligned} c &= \mu_0 - t_c\sqrt{\frac{S^2}{n}} \\ &= 3000 - 1,895\frac{120,1865}{\sqrt{8}} \\ &= 2919,477, \end{aligned}$$

logo, RC $= \{\overline{X} : \overline{X} < 2919,477\}$, como $\overline{x} = 3081,75 > 2919,477$, $\overline{x} \notin$ RC, logo ao nível de significância de 5% pode–se concluir que houve mudança, e portanto existem evidências de que a crise econômica não tem afetado os ingressos.

(c) Se $\alpha = 1\%$, na tabela da t–Student do livro do Morettin $1\% = p/2$, logo, $p = 2\%$, assim o valor de $t_7 = 2,998$, e

$$\begin{aligned} c &= \mu_0 - t_c\sqrt{\frac{S^2}{n}} \\ &= 3000 - 2,998\frac{120,1865}{\sqrt{8}} \\ &= 2872,392 \end{aligned}$$

assim, RC $= \{\overline{X} : \overline{X} < 2872,392\}$, como $\overline{x} = 3081,75 > 2872,392$, $\overline{x} \notin$ RC, ao nível de significância do 1% pode–se concluir que houve mudanças, e portanto existem evidências de que a crise econômica não tem afetado os ingressos.

(d) O nível descritivo é

$$\begin{aligned} \text{p-value} = \widehat{\alpha} &= P(\overline{X} < 3081,75 | \mu_0 = 3000) \\ &= P\left(\frac{\overline{X} - 3000}{120,1865/\sqrt{8}} < \frac{3081,75 - 3000}{120,1865/\sqrt{8}}\right) \\ &= P(T < 1,924) \\ &= 0,952 \end{aligned}$$

Com base no nível descritivo, como $\widehat{\alpha} = 0,952 > 0,05 = \alpha$, ao nível de significância do 5% não se rejeita a hipótese nula, isto é, pode–se concluir que houve mudança, a crise econômica não tem afetado os ingressos, a mesma conclusão obtida no item (b). Considerando $\alpha = 1\%$, como $\widehat{\alpha} = 0,952 > 0,01 = \alpha$, ao nível de significância do 1% não se rejeita a hipótese nula, isto é, pode–se concluir que houve mudança, a crise econômica não tem afetado os ingressos, a mesma conclusão obtida no item (c).

Exercício 3:

O tempo de permanência de economistas recém-formados no primeiro emprego, em anos, foi estudado considerando um modelo Normal de média e variância desconhecidas. Para uma amostra de 15 economistas obteve-se uma média de 2,7 anos e o desvio–padrão de 1,4 anos.

(a) Por analogia a outras categorias profissionais, deseja-se testar se a média é de 2 anos contra a alternativa de ser diferente desse valor. Ao nível de 1% qual foi a conclusão do teste?

(b) Por analogia a outras categorias profissionais, deseja-se testar se a variância é de 1,5 anos contra a alternativa de ser diferente desse valor. Ao nível de 1% qual foi a conclusão do teste?

Solução:

Seja X :"tempo de permanência de economistas recém-formados no primeiro emprego, em anos", $X \sim N(\cdot;\cdot)$.

(a) Se deseja testar

$$\begin{cases} H_0: & \mu = 2, \text{ a média do tempo de permanência de economistas recém formados} \\ & \text{no primeiro emprego é de 2 anos.} \\ & \\ H_1: & \mu \neq 2, \text{ a média do tempo de permanência de economistas recém formados} \\ & \text{no primeiro emprego é diferente de 2 anos.} \end{cases}$$

Como o desvio padrão não é conhecido e o tamanho amostral é pequeno, se considera a estatística de teste baseada na distribuição t–Student

$$T = \frac{\overline{X} - \mu_0}{S/\sqrt{n}} \underset{\mu=\mu_0=2}{\sim} t_{n-1}$$

como $\alpha = 1\%$ e o teste é bilateral na tabela da t–Student do livro do Morettin $1\%/2 = p/2$, logo, $p = 1\%$, assim o valor de $t_{14} = 2,977$, a região crítica é dada por RC $= \{\overline{X} : \overline{X} \leq c \text{ ou } \overline{X} \geq d\}$, em que

$$\begin{aligned} c &= \mu_0 - t_c \frac{S}{\sqrt{n}} \\ &= 2 - 2,977 \frac{1,4}{\sqrt{15}} \\ &= 0,9239, \end{aligned}$$

e

$$\begin{aligned} d &= \mu_0 + t_c \frac{S}{\sqrt{n}} \\ &= 2 + 2,977 \frac{1,4}{\sqrt{15}} \\ &= 3,076, \end{aligned}$$

assim,

$$\begin{aligned} \text{RC} &= \{\overline{X} : \overline{X} \leq c \ \text{ ou } \ \overline{X} \geq d\} \\ &= \{\overline{X} : \overline{X} \leq 0,9239 \ \text{ ou } \ \overline{X} \geq 3,076\}, \end{aligned}$$

como $\overline{x} = 2,7$, tem–se que $0,9239 \leq \overline{x} \leq 3,076$, logo $\overline{x} \notin \text{RC}$, ao nível de significância de 1% pode–se concluir que a média do tempo de permanência de economistas recém-formados no primeiro emprego é de 2 anos.

(b) Se deseja testar

$$\begin{cases} H_0: & \sigma^2 = 1,5, \text{ a variância do tempo de permanência de economistas recém formados} \\ & \quad \text{no primeiro emprego é de 1,5 anos.} \\ \\ H_1: & \sigma^2 \neq 1,5, \text{ a variância do tempo de permanência de economistas recém formados} \\ & \text{no primeiro emprego é diferente de 1,5 anos.} \end{cases}$$

A estatística de teste é dada por

$$\frac{(n-1)S^2}{\sigma^2} \underset{\sigma^2 = {\sigma^2}_0 = 1,5}{\sim} \chi^2_{(n-1)},$$

a região crítica é dada por $\text{RC} = \{S^2 : S^2 \leq a \ \text{ ou } \ S^2 \geq b\}$, em que a é tal que

$$\begin{aligned} P(S^2 \leq a) &= \frac{\alpha}{2} \\ P\left(\chi^2_{(n-1)} \leq \frac{(n-1)a}{\sigma^2}\right) &= \frac{\alpha}{2} \end{aligned}$$

como $\alpha = 1\%$, então $\frac{(n-1)a}{\sigma^2} = \chi^2_1$ e $\chi^2_1 : P(\chi^2_{14} \leq \chi^2_1) = 0,005\%$, logo

$$
\begin{aligned}
a &= \frac{\sigma^2}{(n-1)}\chi_1^2 \\
&= \frac{1,5}{14}4,075 \\
&= 0,4366
\end{aligned}
$$

e b é tal que

$$
\begin{aligned}
P(S^2 \geq b) &= \frac{\alpha}{2} \\
P\left(\chi^2_{(n-1)} \geq \frac{(n-1)b}{\sigma^2}\right) &= \frac{\alpha}{2}
\end{aligned}
$$

como $\alpha = 1\%$, então $\frac{(n-1)b}{\sigma^2} = \chi_2^2$ e $\chi_2^2 : P(\chi^2_{14} \geq \chi_2^2) = 0,005\%$, logo

$$
\begin{aligned}
b &= \frac{\sigma^2}{(n-1)}\chi_2^2 \\
&= \frac{1,5}{14}31,319 \\
&= 3,3556
\end{aligned}
$$

assim, RC $= \{S^2 : S^2 \leq 0,4366 \text{ ou } S^2 \geq 3,3556\}$, como $s^2 = 1,96 \notin$ RC, logo podemos concluir ao nível de significância de 1% que a variância do tempo de permanência de economistas recém-formados no primeiro emprego é de 1,5 anos.

Considerando $\alpha = 2\%$ tem–se que $\chi_1^2 = 4,660$ e $\chi_2^2 = 29,142$, assim

$$
a = \frac{1,5}{14}4,660 = 0,4933 \text{ e } b = \frac{1,5}{14}29,142 = 3,1223,
$$

assim, RC $= \{S^2 : S^2 \leq 0,4933 \text{ ou } S^2 \geq 3,1223\}$, como $s^2 = 1,96 \notin$ RC, logo podemos concluir ao nível de significância de 2% que a variância do tempo de permanência de economistas recém-formados no primeiro emprego é de 1,5 anos.

Exercício 4:

Certa instituição financeira tem uma determinada linha de crédito que 65% dos clientes conseguem pagar, sem se tornarem inadimplentes. Para diminuir a taxa de inadimplência, um estudo foi feito e novas regras de crédito foram estabelecidas. Com essas novas regras, espera-se que a porcentagem de bons clientes aumente. Para verificar se essas mudanças são efetivas, a instituição decide fazer um teste, escolhendo 22 clientes que receberam o crédito segundo as novas regras.

(a) Formule este problema como um problema de teste de hipóteses.

(b) Interprete os erros de tipo I e de tipo II.

(c) Se, dentre os 22 clientes que receberam o crédito, 15 deles conseguiram liquidar todo o pagamento, qual é o nível descritivo e qual é a decisão a ser tomada?

Solução:
Seja p a proporção de clientes que conseguem pagar, sem se tornarem inadimplentes.

(a) Se deseja testar

$$\begin{cases} H_0: & p \leq 0,65, \text{ a proporção de bons clientes não aumentou.} \\ \\ H_1: & p > 0,65, \text{ a proporção de bons clientes aumentou.} \end{cases}$$

(b) os erros de tipo I e de tipo II são

Erro tipo I: (Rejeitar H_0 quando H_0 é verdadeira.) Afirmar que a proporção dos clientes que conseguem pagar, sem se tornarem inadimplentes aumentou, (é maior que 0,65) quando na realidade essa proporção não aumentou (é menor).

Erro tipo II: (Não rejeitar H_0 quando H_0 é falsa.) Afirmar que a proporção dos clientes que conseguem pagar, sem se tornarem inadimplentes não aumentou (é menor que 0,65), mas na realidade essa proporção aumentou (é maior).

(c) A proporção estimada de clientes que conseguem pagar, sem se tornarem inadimplentes é $\widehat{p}_{obs} = \frac{15}{22} = 0,682$. O nível descritivo é

$$\begin{aligned} \widehat{\alpha} &= P(\widehat{p} > 0,682 | p_0 = 0,65) \\ &= P\left(\frac{\widehat{p} - 0,65}{\sqrt{\frac{0,65(1-0,65)}{22}}} > \frac{0,682 - 0,65}{\sqrt{\frac{0,65(1-0,65)}{22}}}\right) \\ &= P\left(Z > \frac{0,032}{0,102}\right) \\ &= P(Z > 0,315) \\ &= 1 - P(Z \leq 0,315) = 1 - 0,624 = 0,3764. \end{aligned}$$

Com base no nível descritivo pode–se concluir que não se rejeita H_0 aos níveis de significância usuais, isto é, não existem evidências para afirmar que a proporção de bons clientes aumentou, e que as mudanças foram efetivas.

Exercício 5:

Um levantamento de opinião mostrou que nos últimos meses a proporção p de habitantes de certo país que desaprovam a política de economia de energia do governo federal é igual a 75%. O presidente do país introduz uma série de mudanças na política de economia de energia e seus assessores garantem que essa proporção diminuiu. Para isso, 60 pessoas são entrevistadas depois da introdução das mudanças.

(a) Formule este problema como um problema de teste de hipóteses.

(b) Qual é o significado do erro do tipo I e do tipo II para o problema?

(c) Qual é a região crítica para um nível de significância de 5%?

(d) Se 38 das 60 pessoas entrevistadas desaprovam a política de economia de energia do governo federal, qual é a conclusão?

(e) Calcule o nível descritivo e conclua.

Solução:

Seja p : "proporção de habitantes que desaprovam a política de economia de energia do governo federal".

(a) Se deseja testar

$$\begin{cases} H_0: & p \geq 0,75, \text{ ou } p = 0,75, \text{ a proporção é a mesma ou maior.} \\ \\ H_1: & p < 0,75, \text{ a proporção diminuiu.} \end{cases}$$

(b) os erros de tipo I e de tipo II são

Erro tipo I: (Rejeitar H_0 quando H_0 é verdadeira.) Afirmar que a proporção de habitantes que desaprovam a política de economia de energia do governo federal é menor que 0,75 quando na realidade essa proporção é maior.

Erro tipo II: (Não rejeitar H_0 quando H_0 é falsa.) Afirmar que a proporção de habitantes que desaprovam a política de economia de energia do governo federal é maior que 0,75, mas na realidade essa proporção é menor.

(c) Quando $p_0 = 0,75$, $\frac{\widehat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \sim N(0,1)$, como $\alpha = 5\%$, $z = 1,645$, assim

$$\text{RC} = \{\widehat{p} : \widehat{p} \leq c\},$$

em que

$$\begin{aligned}
c &= p_0 - z_{(1-\alpha)}\sqrt{\frac{p_0(1-p_0)}{n}}\\
&= 0,75 - 1,645\sqrt{\frac{0,75(1-0,75)}{60}}\\
&= 0,658.
\end{aligned}$$

A região crítica é RC $= \{\widehat{p} : \widehat{p} \leq 0,658\}$.

(d) A estimativa para p é $\widehat{p}_{obs} = \frac{38}{60} = 0,633$. Ao nível de significância do 5% como $\widehat{p}_{obs} \in$ RC pode–se concluir que a proporção de habitantes que desaprovam a política de economia de energia do governo federal diminuiu.

(e) O nível descritivo é

$$\begin{aligned}
\widehat{\alpha} &= P\left(\widehat{p} < 0,633 | p_0 = 0,75\right)\\
&= P\left(Z < \frac{0,633-0,75}{\sqrt{\frac{0,75(1-0,75)}{60}}}\right)\\
&= P\left(Z < \frac{-0,117}{0,056}\right)\\
&= P(Z < -2,093)\\
&= 1 - P(Z < 2,093) = 0,018.
\end{aligned}$$

Como $\widehat{\alpha} = 0,018 < 0,05 = \alpha$, ao nível de significância do 5% pode–se concluir que se rejeita a hipótese nula, isto é, existe evidência de que a proporção de habitantes que desaprovam a política de economia de energia do governo federal diminuiu. Esta conclusão é a mesma que a obtida no item (d).