# UENF

Universidade Estadual do Norte Fluminense Darcy Ribeiro

**Curso:** Ciência de Computação **Data:** 10 /.06./2024
**Prova:** L7 **Período:** **3º** **Disciplina:** Estatística e Probabilidades
**Professor:** Fermín Alfredo Tang **Turno:** Diurno
**Nome do Aluno: .........................................Matrícula: ......................**

## Lista de exercícios 7 – Teste de Hipóteses

1. A propaganda da companhia de cigarros Tabacox afirma que o teor médio de nicotina dos cigarros da marca Delicious, que ela fabrica, é de no máximo 0,7 mg. Um organismo fiscalizador analisa 16 cigarros dessa marca, obtendo um valor médio para a amostra analisada de 0,708mg de nicotina. Os dados da amostra:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 0,718 | 0,703 | 0,692 | 0,792 | 0,657 | 0,679 | 0,706 | 0,719 |
| 0,673 | 0,682 | 0,665 | 0,684 | 0,770 | 0,761 | 0,699 | 0,728 |

O organismo decide denunciar o fabricante a justiça, que o autua por propaganda enganosa e o condena a pagar uma elevada multa. A companhia decide recorrer.

Construir o teste de hipótese para o teor médio de nicotina, considerando um desvio-padrão populacional igual a 0,04mg. Pode-se concluir que o teor de nicotina excede ao anunciado?. A companhia deve ser multada?

**Resposta1.-**

Considera-se que o teor de nicotina segue uma distribuição normal.

A média amostral é $\bar{x} = 0{,}708$

1. As hipóteses nula $H_0: \mu = 0{,}7$ e alternativa $H_1: \mu > 0{,}7$
2. Podemos considerar o nível de significância em $\alpha = 0{,}05$
3. Como o desvio padrão é conhecido, $\sigma = 0{,}04$, e o tamanho da amostra é $n = 16$ a estatística de teste é:

$$z = \frac{\bar{x} - u_0}{\sigma/\sqrt{n}} = \frac{\bar{x} - 0{,}7}{0{,}04/\sqrt{16}} = \frac{\bar{x} - 0{,}7}{0{,}01} = \frac{0{,}708 - 0{,}7}{0{,}01} = 0{,}8$$

4. Para especificar a região de rejeição, considera-se que quanto maior for a média amostral $\bar{x}$, maiores serão as razões para rejeitar $H_0$.

   Rejeitaremos $H_0$ se $z \geq z_\alpha$ (teste de cauda superior)

   O valor crítico para $z_{0,05} = 1{,}64$

5. Como a estatística de teste $z = 0{,}8 \ngeq z_{0,05} = 1{,}64$

   Não rejeitamos a hipótese $H_0$. Não se pode afirmar que o teor de nicotina real excede significativamente ao anunciado. A companhia não deveria ser multada.

   Observe que

   Rejeitamos $H_0$ se $\bar{x} \geq C$

   Padronizando $P(\bar{x} \geq C) = P\left(\frac{\bar{x} - 0{,}7}{0{,}01} \geq \frac{C - 0{,}70}{0{,}01}\right) = 0{,}05$

Isso acontece quando $P(z \geq 1{,}64)$

Logo $\frac{C - 0{,}70}{0{,}01} = 1{,}64$ resulta em $C = 0{,}70 + 1{,}64 \times 0{,}01 = 0{,}7164$

2. Resolva o exemplo anterior considerando que não se conhece o desvio padrão. Além disso o tamanho da amostra $n = 16$ é pequeno. Logo, o teste de hipótese deverá usar a distribuição t student.com n-1 graus de liberdade.

**Resposta2.-**

Considera-se que o teor de nicotina segue uma distribuição normal.

A média amostral é $\bar{x} = 0{,}708$

1. As hipóteses nula $H_0: \mu = 0{,}7$ e alternativa $H_1: \mu > 0{,}7$
2. Podemos considerar o nível de significância em $\alpha = 0{,}05$
3. Como o desvio padrão é desconhecido, calcula-se o desvio padrão amostral,

$$s^2 = \frac{1}{n-1}\left\{\sum_{i=1}^{n} {x_i}^2 - \left(\sum_{i=1}^{n} x_i\right)^2 /n\right\} = 0{,}001497$$

$$s = 0{,}0387$$

Como o tamanho da amostra $n = 16$ é pequeno a estatística de teste é:

$$t = \frac{\bar{x} - u_0}{s/\sqrt{n}} = \frac{\bar{x} - 0{,}7}{0{,}0387/\sqrt{16}} = \frac{\bar{x} - 0{,}7}{0{,}0387/4} = \frac{0{,}708 - 0{,}7}{0{,}0387/4} = 0{,}827$$

4. Para especificar a região de rejeição, considera-se que quanto maior for a média amostral $\bar{x}$, maiores serão as razões para rejeitar $H_0$.

   Rejeitaremos $H_0$ se $t \geq t_{\alpha,n-1} = t_{0{,}05;\,15} = 1{,}753$ (teste de cauda superior)

   O valor crítico para $t_{0{,}05;\,15} = 1{,}753$

5. Como a estatística de teste $t = 0{,}827 \ngeq t_{0{,}05;\,15} = 1{,}753$

   Não rejeitamos a hipótese $H_0$. Não se pode afirmar que o teor de nicotina real excede significativamente ao anunciado. A companhia não deveria ser multada.

3. Deseja-se testar a hipótese nula $H_0$ de que em média os habitantes de uma certa localidade L vivem pelo menos 80 anos contra a alternativa $H_1$ de que eles vivem em média menos de 80 anos. Com base em levantamentos anteriores sabe-se que o tempo de vida das pessoas dessa localidade segue distribuição normal com desvio padrão populacional de 20 anos. Foi extraída uma amostra com 16 casos dos registros de óbitos da localidade, no ultimo ano, a média amostral foi de 70 anos.

a) Qual seria a conclusão sobre o tempo médio de vida, ao nível de significância de 5%.

b) Qual é o nível crítico?

**Resposta3.-**

a) Considera-se que o tempo de vida segue uma distribuição normal.

A média amostral é $\bar{x} = 70$ o desvio padrão $\sigma = 20$, amostra pequena $n = 16$

1. As hipóteses nula $H_0: \mu = 80$ e alternativa $H_1: \mu < 80$

2. Podemos considerar o nível de significância em $\alpha = 0{,}05$
3. Como o desvio padrão é conhecido, $\sigma = 20$, e o tamanho da amostra é $n = 16$ a estatística de teste é:

$$z = \frac{\bar{x} - 80}{\sigma/\sqrt{n}} = \frac{\bar{x} - 80}{20/\sqrt{16}} = \frac{\bar{x} - 80}{5} = \frac{70 - 80}{5} = -2$$

4. Para especificar a região de rejeição, considera-se que quanto menor for a média amostral $\bar{x}$, maiores serão as razões para rejeitar $H_0$.

   Rejeitaremos $H_0$ se $z \leq z_\alpha$ (teste de cauda inferior)

   O valor crítico para $z_{0,05} = -1{,}64$
5. Como a estatística de teste $z = -2 \leq z_{0,05} = -1{,}64$

   Rejeitamos a hipótese $H_0$. A partir da amostra e com o desvio $\sigma$ não se pode afirmar que a média de vida é 80 anos ou mais.

Observe que existe um valor de corte C, para o qual rejeitamos $H_0$:

Rejeitaremos $H_0$ se $\bar{x} < C$

$\alpha = 0{,}05 = P(Erro\ I) = P(Rejeitar\ H_0)$ quando $H_0$ verdadeira

$0{,}05 = P(\bar{x} < C)$, onde $\bar{x} \sim N(80,\ 5^2)$

Padronizando $P(\bar{x} < C) = P\left(\frac{\bar{x} - 80}{5} < \frac{C - 80}{5}\right) = 0{,}05$

Isso acontece quando $P\left(\frac{\bar{x} - 80}{5} < -1{,}64\right)$

Logo $\frac{C - 80}{5} = -1{,}64$ resulta em $C = 80 - 1{,}64 \times 5 = 71{,}80$

Rejeitaremos $H_0$ se $\bar{x} < 71{,}80$

b) Calculo do nível crítico:

Se tivéssemos usado $\alpha = 0{,}01$ Repetindo o procedimento anterior teríamos:

Rejeitaremos $H_0$ se $z \leq z_\alpha$ (teste de cauda inferior)

O valor crítico para $z_{0,01} = -2{,}33$

Como a estatística de teste $z = -2 \nleq z_{0,01} = -2{,}33$

Não rejeitamos a hipótese $H_0$. A partir da amostra e com o desvio $\sigma$ se pode afirmar que a média de vida é 80 anos ou mais.

Logo $\frac{C - 80}{5} = -2{,}33$ resulta em $C = 80 - 2{,}33 \times 5 = 68{,}35$

Rejeitaremos $H_0$ se $\bar{x} < 68{,}35$

Reduzindo o valor de $\alpha$ de 0,05 até 0,01 vai existir um ponto de $\alpha$ que passa da rejeição para aceitação. Para isso fixamos o valor de corte igual a média amostral:

$\hat{\alpha} = P(\bar{x} < 70)$, onde $\bar{x} \sim N(80,\ 5^2)$ normalizando

$$\hat{\alpha} = P\left(z < \frac{70 - 80}{5}\right) = P(z < -2) = 0{,}0228$$

Com isso, o p-valor o nível crítico é 2,3%.

4. Na questão dos cigarros determine o poder do teste, para rejeitar corretamente a hipótese nula. Ou seja o valor do $\beta$ erro tipo II, considerando a hipóteses nula $H_0: \mu = 0{,}7$ e sendo que o verdadeiro conteúdo médio de nicotina dos cigarros na população é $\mu = 0{,}71$.

**Resposta4.-**

Neste caso,

$\beta = P(Erro\ II) = P(Aceitar\ H_0: \mu = 0{,}7)$ quando $\mu = 0{,}71$, ou seja $H_0$ Falsa.

Lembre que foi considerado $\alpha = 0{,}05$

E encontramos o valor crítico $C = 0{,}7164$ mg

$\beta = P(\bar{x} < 0{,}7164)$ quando $\mu = 0{,}71$ normalizando

$\beta = P\left(z < \frac{0{,}7164-0{,}71}{\sigma/\sqrt{n}}\right) = P\left(z < \frac{0{,}7164-0{,}71}{0{,}04/\sqrt{16}}\right) = P(z < 0{,}64) = 0{,}7389$

Com isso o poder do teste é de $1 - \beta = 0{,}2611$

Isso significa que se o verdadeiro valor de $\mu = 0{,}71$ a hipótese $H_0: \mu = 0{,}7$ será corretamente rejeitada em 26,11% dos casos.

5. Queremos testar a hipótese nula de que a proporção $p$ de psicólogos atuando na área de clínica é menor ou igual a 60% contra a hipótese alternativa de que ela é superior a 60%, em uma certa cidade. Foi usada uma amostra aleatória de $n = 80$ psicologos selecionado ao acaso nessa cidade. Suponhamos que, como resultado do experimento, verificou-se que 61 entre 80 psicologos sorteados atuam na área clínica.

Qual decisão deve ser tomada, ao nível de significância de 0,01?

**Resposta5.-**

Neste caso temos: $\hat{p} = \frac{x}{n} = \frac{61}{80} = 0{,}763$ Proporção amostral

$p_0 = 0{,}60$ $\quad p_0 = 0{,}40$

1. Hipótese nula: $H_0: p = p_0$

   Hipótese alternativa: $H_a: p > p_0$ $\quad z \geq z_\alpha$ (teste de cauda superior)
2. Considerase o nível de significância em $\alpha = 0{,}01$
3. Calcula-se a estatística de teste: Valor da estatística de teste:

   $z = \frac{\hat{p} - p_0}{\sqrt{p_0 q_0/n}} = \frac{0{,}763 - 0{,}6}{\sqrt{0{,}60 \times 0{,}40/80}} = \frac{0{,}163}{0{,}0548} = 2{,}974$
4. Para especificar a região de rejeição, considera-se que quanto maior for a média amostral $\hat{p}$ , maiores serão as razões para rejeitar $H_0$.

   Rejeitaremos $H_0$ se $z \geq z_\alpha$ (teste de cauda superior)

   O valor crítico para $z_{0,01} = 2{,}33$
5. Como a estatística de teste $z = 2{,}974 \geq z_{0,01} = 2{,}33$

   Rejeitamos a hipótese $H_0$. Conclui-se que a proporção de psicólogos é maior que 60%.

6. Sergio afirma que Raquel dirige seu carro na estrada a uma velocidade média superior a 100km/h, enquanto Raquel discorda, afirmando dirigir na estrada a uma velocidade média menor ou igual a 100km/h. Para dirimir essa controvérsia, Sergio resolve cronometrar o tempo (em minutos) que ela gasta ao volante em 10 viagens, sempre pelo mesmo percurso que liga duas cidades 120km distantes:

73, 68, 73, 61, 70, 78, 63, 64, 74, 62.

a) Quem parece ter razão, ao nível de significância de 5%?
b) Qual o nível crítico?

**Resposta 6.-**

a) As hipóteses a serem consideradas são $H_0: \mu \leq 100$ (Raquel) e $H_1: \mu > 100$ (Sergio), em que $\mu$ é a média populacional da velocidade (em km/h) com que Raquel dirige na estrada. Como o desvio padrão populacional $\sigma$ é desconhecido, teremos que utilizar a distribuição t de *Student* com $\nu = 10 - 1 = 9$ graus de liberdade.

O critério de decisão a ser utilizado é:

rejeitar $H_0$, se $T_{obs} = \dfrac{\bar{x} - 100}{s/\sqrt{10}} > t_{0,95} = 1{,}83$. Caso contrário, aceitar $H_0$.

Os dados com os quais devemos trabalhar neste caso são as velocidades médias (em km/h) correspon-ntes a cada uma das 10 viagens. Eles podem ser obtidos por meio da expressão:

$$\text{velocidade}(\text{km/h}) = \frac{120\text{km}}{\text{tempo}(\text{min})} \times 60\left(\frac{\text{min}}{\text{h}}\right)$$

Os resultados obtidos em km/h são:

98,63 105,88 98,63 118,03 102,86 92,31 114,29 112,50 97,30 116,13

Com base nesses dados, temos: $\bar{x} = 105{,}66$ ; $s = 9{,}06$ .

Daí, $T_{obs} = \dfrac{105{,}56 - 100}{9{,}06/\sqrt{10}} = 1{,}97 > 1{,}83$, o que implica que $H_0$ deve ser rejeitada.

Isso significa que, ao nível de significância $\alpha = 0{,}05$, Sergio parece estar com a razão.

b) O p-valor no caso é $\tilde{\alpha} = P(T > 1{,}97) = 0{,}0399$

**7.-** Queremos testar a hipótese $H_0$: $\mu = 50$ contra a alternativa $H_1$: $\mu = 52$, onde $\mu$ é a média populacional de uma distribuição Normal com variância conhecida $\sigma = 16$. Temos então $n$ variáveis aleatórias iid $X_1, X_2, \dots, X_n$, que seguem essa lei de probabilidade, e a média amostral $\bar{X}$ será usada como estatística de teste. Como 50 < 52, e claro que quanto maior for $\bar{X}$ maiores serão as razões para se rejeitar $H_0$ em favor de $H_1$. Então o critério de decisão adequado será da forma: rejeitar $H_0$, $H_0$se

$\bar{X}_{obs} > \bar{X}_C$, e aceitar $H_0$, se $\bar{X}_{obs} \le \bar{X}_C$; onde $\bar{X}_C$ é uma constante a determinar em funcao de outras condicoes a serem especificadas.

a) Se $n$ = 30 e fixarmos P(Erro I) em α = 0,01, quais devem ser o critério de decisão e $\beta$ = P(Erro II)?
b) Mostre que se o tamanho da amostra for mantido em $n$ = 30 , ao fazermos com que o ponto de corte $\bar{X}_C$ se mova para a esquerda de forma que $\beta$ diminua, $\alpha$ necessariamente aumentara.
c) Quais devem ser o critério de decisão e o tamanho $n$ da amostra para que a probabilidade do Erro II se reduza a β = 0,05, mantendo a probabilidade do Erro I em α = 0,01?

**Resposta 7.-**

a) Sabemos que $0{,}01 = \alpha = P(\text{Erro I}) = P(\text{Rejeitar } H_0)$, se $H_0$ é verdadeira.

Então $0{,}01 = P\left[\bar{X} > \bar{x}_C\right]$, se $\mu = 50 \Rightarrow 0{,}01 = P\left[\dfrac{\bar{X}-50}{4/\sqrt{30}} > \dfrac{\bar{x}_C - 50}{4/\sqrt{30}}\right] = P\left[Z > \dfrac{\bar{x}_C - 50}{4/\sqrt{30}}\right]$, onde Z ~ N(0;1).

Logo, $\dfrac{\bar{x}_C - 50}{4/\sqrt{30}} = 2{,}33$ (valor obtido da tabela da Normal Padrão) $\Rightarrow \bar{x}_C = 51{,}70$ .

Por outro lado, $\beta = P(\text{Aceitar } H_0)$, se $H_0$ é falsa, ou seja, $\beta = P\left[\bar{X} \le 51{,}70\right]$, se $\mu = 52$ .

Então, $\beta = P\left[\dfrac{\bar{X}-52}{4/\sqrt{20}} \le \dfrac{51{,}70 - 52}{4/\sqrt{20}}\right] = P(Z \le -0{,}412) = 0{,}340$ , o que é um valor excessivamente alto para uma probabilidade de erro.

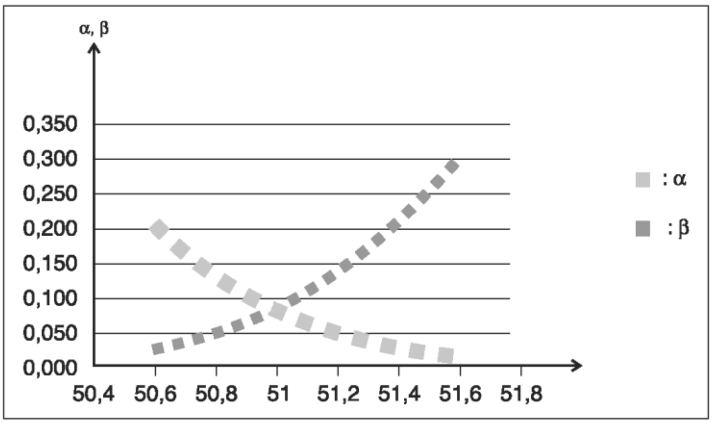
Assim, para que $\beta = P(\text{Erro II})$ diminuísse (mantendo $n = 30$), a única alternativa possível seria deslocar o ponto de corte $\bar{x}_C$ para a esquerda, e desse modo reduzir a região de aceitação. Isso é o que faremos no item (b) a seguir.

b) Vamos fazer com que o ponto de corte $\overline{x}_c$ decresça de 0,1 em 0,1, desde 51,6 até 50,6.

Assim, teremos, para cada valor possível de $\overline{x}_c$, $\alpha = P\left[Z > \frac{\overline{x}_c - 50}{4/\sqrt{30}}\right]$ e $\beta = P\left[Z \leq \frac{\overline{x}_c - 52}{4/\sqrt{20}}\right]$.
Fazendo os cálculos, montamos a tabela a seguir:

| $\overline{x}_c$ | $\beta$ | $\alpha$ | | $\overline{x}_c$ | $\beta$ | $\alpha$ |
|---|---|---|---|---|---|---|
| 51,6 | 0,292 | 0,014 | | 51,0 | 0,085 | 0,085 |
| 51,5 | 0,247 | 0,020 | | 50,9 | 0,066 | 0,109 |
| 51,4 | 0,206 | 0,028 | | 50,8 | 0,050 | 0,137 |
| 51,3 | 0,169 | 0,038 | | 50,7 | 0,038 | 0,169 |
| 51,2 | 0,137 | 0,050 | | 50,6 | 0,028 | 0,206 |
| 51,1 | 0,109 | 0,066 | | | | |

que dá origem ao seguinte gráfico:




c) Temos $0{,}01 = \alpha = P(\text{Erro I}) = P(\text{Rejeitar } H_0)$, se $H_0$ é verdadeira.

Então, por um raciocínio análogo ao dos itens anteriores, chegamos a

$$0{,}01 \quad P\left[\ \frac{-\quad 50}{/\sqrt{\ }}\right] \Rightarrow \frac{\overline{x}_c - 50}{4/\sqrt{n}} = 2{,}33. \qquad \text{(I)}$$

Analogamente, $0{,}05 = \beta = P(\text{Erro II}) = P(\text{Aceitar } H_0)$, se $H_0$ é falsa. Daí,

$$0{,}05 = P\left[Z \leq \frac{\overline{x}_c - 52}{4/\sqrt{n}}\right] \Rightarrow \frac{\overline{x}_c - 52}{4/\sqrt{n}} = -1{,}64. \qquad \text{(II)}$$

Resolvendo o sistema de duas equações a duas incógnitas formado pelas equações (I) e (II), obtemos $\overline{x}_c = 51{,}17$ e $n = 63{,}08 \cong 63$. Isso mostra que, de fato, para podermos reduzir o valor de $\beta$ sem que $\alpha$ cresça, é necessário aumentar o tamanho n da amostra.

**8.-** Em uma pesquisa eleitoral referente ao primeiro turno de uma eleição para governador foram ouvidos $n$ = 1.000 eleitores selecionados aleatoriamente e entre eles $m$ = 510 declararam-se favoráveis ao candidato A. Deseja-se testar a hipótese

$H_0$ de que a proporção $p$ de eleitores do candidato A é menor ou igual a 0,5 contra a alternativa de que A venceria direto, sem a necessidade do segundo turno.

a) Qual seria a sua decisão ao nível de significância de 5%? Por quê?

b) Qual é o $p$-valor?

c) Se na realidade $p$ = 0,55, qual seria a probabilidade de ser cometido o Erro de tipo II ao ser usada essa mesma regra de decisão? Qual o poder do teste neste caso?

**Resposta 8.-**

a) Queremos testar $H_0 : p \leq 0,5$ contra $H_1 : p > 0,5$ ao nível de significância $\alpha = 0,05$.

Então, neste caso, temos $p_0 = 0,5$ e $\sigma_0 = \sqrt{\dfrac{0,5 \times 0,5}{1000}} = 0,0158$

Por outro lado, $1 - \alpha = 0,95$ e $z_{0,95} = 1,645$.

Assim, o critério de decisão é:
Rejeitar $H_0$ se $\hat{p} > 0,5 + 1,645 \times 0,0158 = 0,526$.

Como $\hat{p}_{obs} = \dfrac{510}{1000} = 0,51 < 0,526$, a hipótese nula $H_0$ deve ser aceita.

Isso significa que, ao nível de significância $\alpha = 0,05$, não há evidências suficientes de que o candidato A já ganharia a eleição no primeiro turno.

b) O p-valor $\tilde{\alpha}$ é o menor valor de $\alpha$ para o qual ainda rejeitaríamos $H_0$, com os dados disponíveis. Então $\tilde{\alpha} = P\left(\hat{p} > 0,51\right)$, se $p = 0,5$. Ou seja, padronizando, temos

$$\tilde{\alpha} = P\left(\frac{\hat{p} - 0,5}{0,0158} > \frac{0,51 - 0,5}{0,0158}\right) = P(Z > 0,632) = 0,264.$$

Como esse p-valor é excessivamente grande, isso reforça a nossa decisão de não rejeitar a hipótese nula $H_0$ de que há necessidade de um segundo turno.

c) Se $p_1 = 0,55$, temos $\sigma_1 = \sqrt{\dfrac{0,55 \times 0,45}{1000}} = 0,0157$. Então,

$$\beta = P\left(\text{Aceitar } H_0\right) = P\left(\hat{p} \leq 0,526\right) = P\left(\frac{\hat{p} - 0,55}{0,0157} \leq \frac{0,526 - 0,55}{0,0157}\right) = P(Z \leq -1,525) = 0,064.$$

Nesse caso, o poder do teste seria 1 – 0,064 = 0,936 ou 93,6%.