AULA 19 - Testes de Hipóteses para uma população

Professor: Jorge L. Bazán Monitora: Patrícia Stülp

21/06/2023

Exemplo: Teste para a média de uma população com variância conhecida

Em uma fábrica de cerveja a quantidade de cerveja, em latas, seguia uma distribuição normal com média 350 e desvio padrão 3ml. Após alguns problemas na linha de produção, suspeita-se que houve alteração na média. Uma amostra de 20 latas forneceu uma média de 346 ml.

Usando a informação fornecida podemos encontrar o seguinte intervalo de confiancia. Verifique!

$n = 20$ IC95

Mas essa informação não é suficiente para responder se a suspeita é certa ou errada. Para respondê-la, iremos testar as seguintes hipóteses.

PASSO 1:Especificar as hipóteses

- $H0 : \mu = 350$
- $Ha : \mu \neq 350$

PASSO 2: Especificar a estatística do teste

$$\frac{\overline{X} - \mu_0}{\sqrt{\sigma^2/n}} \sim N(0, 1)$$

PASSO 3:

O erro tipo 1, o nível de significância do teste α neste problema é 0.05. Como a hipótese é de duas caudas (a hipótese Ha corresponde ao caso maior que e menor que), então o correspondente valor da estatística para este nível de significância é obtido da seguinte forma

```
# erro tipo 1 = alfa
alfa = 0.05

# hipoteses de duas caudas
# valor de ztab para o erro tipo 1
```

```
qnorm(alfa/2)

## [1] -1.959964

qnorm(1-alfa/2)

## [1] 1.959964
```

PASSO 4: Calcular o p-valor (ou a região crítica do teste)

Este passo pode ser feito de duas formas.

a) Encontrando a região crítica do teste ou ponto de corte

Devemos encontrar o ponto de corte ou região crítica tal que

$$P(\overline{x} \neq xcorte) = \alpha$$
$$P(\mid \overline{x} \mid > xcorte) = \alpha$$

equivale a encontrar xcorte tal que

$$2P(zest < ztab) = \alpha$$

onde

$$ztab = \frac{xcorte - \mu_0}{\sqrt{\sigma^2/n}}$$

ou

$$\begin{aligned} xcorte1 &= \mu_0 - \sqrt{\sigma^2/n} \times ztab \\ xcorte2 &= \mu_0 + \sqrt{\sigma^2/n} \times ztab \end{aligned}$$

Neste caso, a região crítica é

```
# encontrando ponte de corte
n = 20
xcorte1 = 350 + sqrt(3^2/n)*(-1.96)
xcorte1

## [1] 348.6852

xcorte2 = 350 + sqrt(3^2/n)*(1.96)
xcorte2

## [1] 351.3148
```

b) Encontrando o valor p

Necessitamos encontrar a probabilidade de rejeitar H0 usando a proporção amostral, isto é $P(\mid \overline{X} - 346 \mid)$ quando de fato a hipótese nula é verdadeira, isto é, $\mu_0 = 350$.

$$2P(\overline{X} < \overline{x}obs \mid \mu_0)$$

o qual equivale a

$$P(zest < ztab), ztab = \frac{\overline{x}obs - \mu_0}{\sqrt{\sigma^2/n}}$$

```
n = 20
mu = 350
sigma = 3
xbarraobs = 346

ztab = (xbarraobs - mu)/sqrt(sigma^2/n)
ztab

## [1] -5.962848

2*pnorm(-5.96)

## [1] 2.522379e-09
```

PASSO 5: Decidir entre H0 e Ha

Este paso pode ser feito de duas formas dependendo do passo anterior.

- Usando região crítica como $\overline{x} \leq 348.6852$ ou $\overline{x} \geq 351.3148$, vemos que se cumpre $\overline{x} = 346 \leq 348.6852$ e, portanto, rejeitamos H0. Concluímos que teve alteração da média com um nível de confiança de $1 - \alpha = 0.95$ ou probabilidade de erro tipo 1 de 0.05.
- Usando valor p: neste caso encontramos que o $pvalor = 0.522379e - 09 < \alpha = 0.05$ e, portanto rejeitamos H0. Concluímos que teve alteração da média com um nivel de confiança de $1 - \alpha = 0.95$ ou probabilidade de erro tipo 1 de 0.05.

Exemplo: Teste para a média de uma população com variância desconhecida

Em uma fábrica de cerveja a quantidade de cerveja em latas seguia uma distribuição Normal com média 350. Após alguns problemas na linha de produção, suspeita-se que houve alteração na média. Uma amostra de 20 latas forneceu uma média de 346 ml e desvio padrão 5.

Usando a informação fornecida podemos podemos encontrar o seguinte intervalo de confiancia. Verifique!

$n = 20$, $\overline{x} = 346$, $s^2 = 25$

Mas essa informação não é suficiente para responder se a suspeita é certa ou errada. Para respondê-la, iremos testar as seguintes hipóteses:

PASSO 1:Especificar as hipóteses

- $H0 : \mu = 350$
- $Ha : \mu \neq 350$

PASSO 2: Especificar a estatística do teste

$$\frac{\overline{X} - \mu_0}{\sqrt{s^2/n}} \sim t(n-1)$$

PASSO 3:

O erro tipo 1, o nível de significância do teste α neste problema é 0.05. Como a hipótese é de duas caudas (a hipótese Ha corresponde ao caso maior que e menor que), então o correspondente valor da estatística para este nível de significância é obtido da seguinte forma

```
# erro tipo 1 = alfa
alfa = 0.05
```

PASSO 4: Calcular o p-valor (ou a região crítica do teste)

b) Encontrando o valor p

Necessitamos encontrar a probabilidade de rejeitar H0 usando a proporção amostral, isto é $P(\mid \overline{X} - 346 \mid)$ quando de fato a hipótese nula é verdadeira, isto é, $\mu_0 = 350$.

$$2P(\overline{X} < \overline{x}obs \mid \mu_0)$$

o qual equivale a

$$P(test < ztab), ttab = \frac{\overline{x}obs - \mu_0}{\sqrt{s^2/n}}$$

```
n = 20
mu = 350
s = 5
xbarraobs = 346

ttab = (xbarraobs - mu)/sqrt(s^2/n)
ttab

## [1] -3.577709
```

```
2*pt(-3.6, n-1)

## [1] 0.001908447
```

PASSO 5: Decidir entre H0 e Ha

Este paso pode ser feito de duas formas dependendo do passo anterior.

- Usando valor p: neste caso encontramos que o $pvalor = 0.0019 < \alpha = 0.05$ e, portanto, rejeitamos H0. Concluímos que teve alteração da média com um nível de confiança de $1 - \alpha = 0.95$ ou probabilidade de erro tipo 1 de 0.05.