São Paulo, 30 de maio de 2019

Atividade II - Inferência Estatística

Natália Sales Mesquita; RA 21341077

Exercício 9: O nível de colesterol no sangue é uma variável aleatória com distribuição normal, de média desconhecida μ e desvio-padrão σ = 60 mg/100 ml.

A) Qual um estimador para a média μ?

$\mu \rightarrow \bar{x}$

B)Você conhece a distribuição desse estimador?

Sim, pelo Teorema do Limite Central (TLC) sabe-se que a distribuição de um estimador para a média segue uma curva gaussiana, isto é, curva normal de distribuição.

C) Em uma amostra de 50 pacientes, observou-se uma média amostral $\bar{x}$= 268. Qual seria uma estimativa pontual para μ?

μ=268

D) Considerando a amostra do item anterior, construa um intervalo de confiança para a média desconhecida com nível de confiança de 95%.

Para definir o intervalo de confiança deve-se calcular os limites superior e inferior o qual segue a equação:

$$L = \bar{\boldsymbol{X}} \pm \boldsymbol{Z} * \frac{\boldsymbol{\sigma}}{\sqrt{\boldsymbol{N}}}$$

No caso de 95% de confiança Z assume 1,96. Obtém-se os valores no RStudio:

LI=251.3688

LS=284.6312

```
#CÓDIGO NO RSTUDIO

N1=50; sigma =60;X_barra=268;z= 1.96

LI= X_barra-z*sigma/sqrt(N1);LI

LS = X_barra+z*sigma/sqrt(N1);LS
```

Portanto, o intervalo está entre 251.4 e 284.6 com confiança de 95%.

E) Teste a hipótese de que μ = 260, contra a alternativa de que μ > 260 com base na mesma amostra. Utilize um nível de 5%.

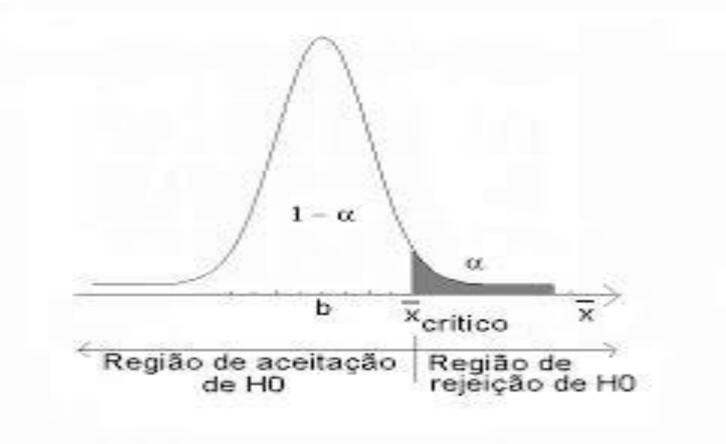

Figura 1- Distribuição normal com aceitações e rejeições

Etapas de um teste de hipóteses:

1. Estabelecer as hipóteses nula e alternativa.

$$\text{Hipóteses}\begin{Bmatrix} HO: \mu = \ 260 \\ H1: \mu > 260 \end{Bmatrix}$$

2. Definir a forma da região crítica, com base na hipótese alternativa.

Para

$$\begin{Bmatrix} a) \boldsymbol{\bar{X}} > \bar{X}_C, \boldsymbol{rejeita} - se\ H0 \\ b) \boldsymbol{\bar{X}} < \bar{X}_C, \boldsymbol{aceita} - se\ H0 \end{Bmatrix}$$

3. Identificar a distribuição do estimador $\boldsymbol{\bar{X}}$.

$$\boldsymbol{\bar{X} \sim N(\mu, \frac{\sigma^2}{\sqrt{N}})}$$

Substituindo valores:

$$\boldsymbol{\bar{X} \sim N(260, \frac{60^2}{\sqrt{50}})}$$

4. Fixar um valor para α e obter a região crítica.

Com α=0,05;a região que tem consequências mais graves se houver um erro na inferência é, <u>por convenção</u>, na situação de rejeitar H0 dado que H0 é verdade . Escrito matematicamente por :

$$\mathbb{P}\ [\text{Rejeitar H0} \mid \text{H0 é verdade}],$$

Substituindo os valores do problema:

$$\mathbb{P}[\bar{X} > \bar{X}\text{c} \mid \mu = 260]$$

Em que $\overline{X}c$ é o valor crítico e $\overline{X}$ é a média da amostra. Calculando com a função *qnorm* no Rstudio:

```
#CÓDIGO NO RSTUDIO
x_critico <- qnorm(0.05, mean = 260, sd =60/sqrt(50), lower.tail = FALSE); x_critico
```

Xc = 273.957

5. Calcular o valor de $\boldsymbol{\overline{X}}$ e concluir o teste com base na região crítica.

$\overline{X}$ = 260 , fornecido pelo enunciado do problema.

Observando na Figura 1 que 260 < 273.9. Portanto pelo *item 2.b* aceita-se H0

F) Qual deve ser o tamanho da amostra para que o intervalo de confiança tenha um erro de 15 unidades? Use 95% de confiança.

```
#CÓDIGO NO RSTUDIO
erro =15; z_95= 1.96; sigma = 60
N2= (z_95*sigma/erro)^2;N2
```

N=62 amostras