Engenharias

# M109

# Estatística

Aula 5 – Testes de Hipóteses – Parte I

Profª. Rosimara Beatriz Arci Salgado

Créditos:
Profª. Karina Perez Mokarzel Carneiro
Prof. Rausley Adriano Amaral de Souza

Inatel

CAMINHOS QUE CONECTAM COM O FUTURO

## Bibliografia básica:

Arango HG. Bioestatística: teórica e computacional. 3ªed. Rio de Janeiro: Guanabara Koogan; 2011.

SPIEGEL, Murray Ralph; FARIA, Alfredo Alves De Probabilidade e estatística. São Paulo, SP: McGraw-Hill, 1978.

# TESTE DE HIPÓTESES

Inferência estatística é um ramo da Estatística cujo objetivo é fazer afirmações a partir de um conjunto de valores representativo (amostra) sobre um universo (população), assume-se que a população é muito maior do que o conjunto de dados observados, a amostra. Tal tipo de afirmação deve sempre vir acompanhada de uma medida de precisão sobre sua veracidade. Para realizar este trabalho, o estatístico coleta informações de dois tipos, experimentais (as amostras) e aquelas que obtém na literatura.

Teste de Hipóteses trata-se de uma técnica para se fazer a inferência estatística sobre uma população a partir de uma amostra

Source:https://pt.wikipedia.org/wiki/Infer%C3%AAncia_estat%C3%ADstica#:~:text=Infer%C3%AAncia%20estat%C3%ADstica%20%C3%A9%20um%20ramo,de%20dados%20observados%2C%20a%20amostra.

# CONCEITOS BÁSICOS

**Hipótese:** é uma pressuposição a respeito de um determinado problema.

O mecanismo de comprovação (verificação) é denominado **teste de hipóteses**. Assim, testar uma hipótese quer dizer verificar se um pressuposto é verdadeiro ou não.

A veracidade ou não do pressuposto é chamada de **conclusão**.

# POR EXEMPLO:

Níveis elevados de bilirrubina em recém-nascidos afetam a capacidade auditiva deles?

Para se chegar a uma conclusão sobre essa questão é necessário **formular uma hipótese e testá-la**.

A formulação da hipótese está relacionada com a forma de conduzir a experiência (**desenho experimental**).

# TEORIA DOS TESTES DE HIPÓTESES

Neste exemplo, poderia ser selecionada uma amostra de *n* recém-nascidos e:

- Efetuar uma dosagem dos níveis de bilirrubina.
- Aferição da capacidade auditiva (quantitativa).

A partir dos dados, seriam constituídos dois grupos:

- A = taxa de bilirrubina normal
- B = taxa de bilirrubina elevada

Seriam então comparadas as capacidades auditivas médias dos grupos A e B.

# TEORIA DOS TESTES DE HIPÓTESES

## Construção das Hipóteses

- A = taxa de bilirrubina normal
- B = taxa de bilirrubina elevada

| | |
|---|---|
| **Hipótese de Nulidade → $H_0$** $H_o \rightarrow \mu_A = \mu_B$ | **Igualdade entre os grupos** |
| **Hipótese Alternativa → $H_1$** $H_1 \rightarrow \mu_A \neq \mu_B$ | Diferença (não igualdade) entre os grupos |

Média da capacidade auditiva do grupo A

Média da capacidade auditiva do grupo B

# TEORIA DOS TESTES DE HIPÓTESES

Assim:

**Aceitar $H_0$** ou dizer que ela é verdadeira por meio de um teste significa afirmar que níveis de bilirrubina **não** estão relacionados com a perda da capacidade auditiva.

**Rejeitar $H_0$** ou dizer que ela é falsa, implica comprovar que os níveis de bilirrubina afetam a capacidade auditiva.

Naturalmente, aceitar $H_1$ implica rejeitar $H_0$ e vice-versa.

**Hipótese de Nulidade → $H_0$**

$$H_o \rightarrow \mu_A = \mu_B$$

**Hipótese Alternativa → $H_1$**

$$H_1 \rightarrow \mu_A \neq \mu_B$$

# TEORIA DOS TESTES DE HIPÓTESES

As hipóteses devem ser **binárias** e **excludentes**.

Permitem que apenas uma das hipóteses seja testada para se obter uma conclusão consistente.

**Por convenção, testa-se sempre $H_0$**. Dessa forma, aceitar $H_0$ implica comprovar a igualdade (nulidade das diferenças); rejeitar $H_0$ significa comprovar a diferença entre os grupos testados.

# TEORIA DOS TESTES DE HIPÓTESES

## Regra de decisão

É um procedimento estatístico para se decidir entre **aceitar ou rejeitar** a hipótese de nulidade ($H_0$).

Normalmente, formula-se a regra de decisão para testar uma hipótese a partir do resultado de um **teste estatístico**.

Ao se tomar uma decisão sobre se a hipótese de nulidade deve ser aceita ou não, existirá a possibilidade de se estar cometendo um **erro**. Por esse motivo, as regras de decisão são construídas seguindo **critérios** que permitam reduzir os erros a elas associados.

# TEORIA DOS TESTES DE HIPÓTESES

## Erros de decisão

| | |
|---|---|
| Erro tipo I → EI | Rejeitar $H_0$ quando ela é verdadeira |
| Erro tipo II → EII | Aceitar $H_0$ quando ela é falsa |

| |
|---|
| **Hipótese de Nulidade → $H_0$** <br> $H_o \rightarrow \mu_A = \mu_B$ |
| **Hipótese Alternativa → $H_1$** <br> $H_1 \rightarrow \mu_A \neq \mu_B$ |

**Nível de significância (α)** → probabilidade de cometer EI

Portanto EI: Erro ao se afirmar que existem diferenças entre os grupos que estão sendo comparados

# TEORIA DOS TESTES DE HIPÓTESES

## Erros de decisão

| | |
|---|---|
| **Erro tipo I → EI** | **Rejeitar $H_0$ quando ela é verdadeira** |
| **Erro tipo II → EII** | **Aceitar $H_0$ quando ela é falsa** |

| | | **Fato** | |
|---|---|---|---|
| | | Verdadeiro(+) | Falso(-) |
| **Decisão** | Aceitação (+) | Decisão Correta(+,+) | **Erro II** |
| | Rejeição (-) | **Erro I** | Decisão correta(-,-) |

## Exemplo:

Suponha que a proporção da população com mais de 60 anos afetada pela doença de Alzheimer seja, segundo a literatura, de 20%.

Se um pesquisador efetuasse um levantamento de dados na região onde ele trabalha e verificasse que, dentre 50 pessoas na faixa etária considerada, os portadores da síndrome seria de 16.

## Exemplo:

Como a proporção por ele encontrada na experiência (16/50 = 32%) difere da proporção citada na literatura, o pesquisador pode imaginar que, por alguma razão, a proporção de indivíduos afetados pela síndrome na sua região de atuação é diferente/maior que a da população em geral.

Contudo, essa diferença pode ser apenas casual, decorrente do fato de se estar trabalhando com uma amostra. Isto é, pode tratar-se de um erro de amostragem, o que levaria a tirar uma conclusão errada, fazer uma afirmação errada.

# Exemplo:

O problema é decidir:

**A partir de que valor é razoável começar a pensar que as taxas sejam diferentes.**

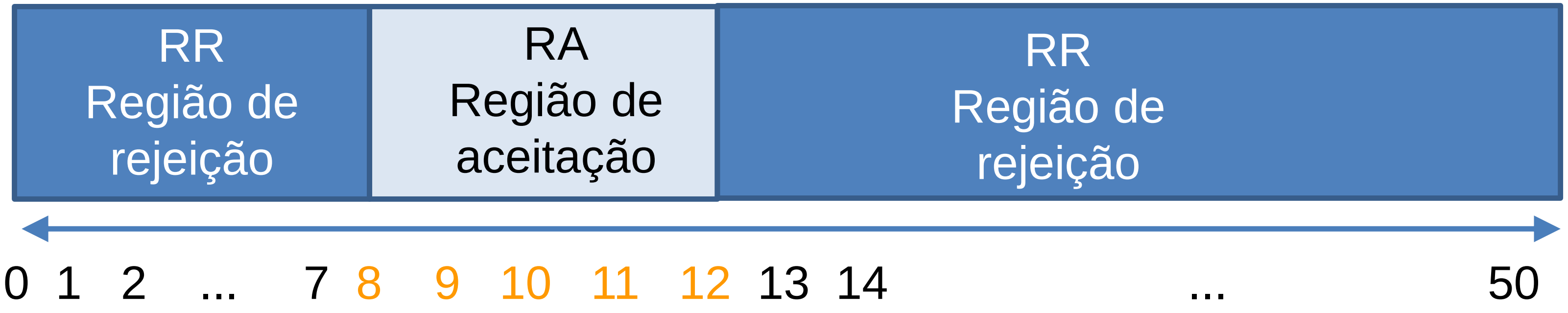
Inatel | CAMINHOS QUE CONECTAM COM O FUTURO

# Exemplo:

Na amostra com 50 indivíduos, com base na literatura, o valor esperado de doentes de Alzheimer é de 10 indivíduos.

$$\mu = E = n.p = 50 \times 0,2 = 10$$

Podemos criar uma regra que aceite $H_0$ quando o número de indivíduos com a síndrome esteja entre 8 e 12 inclusive, porém não é um critério de referência (conhecido).



# Exemplo:

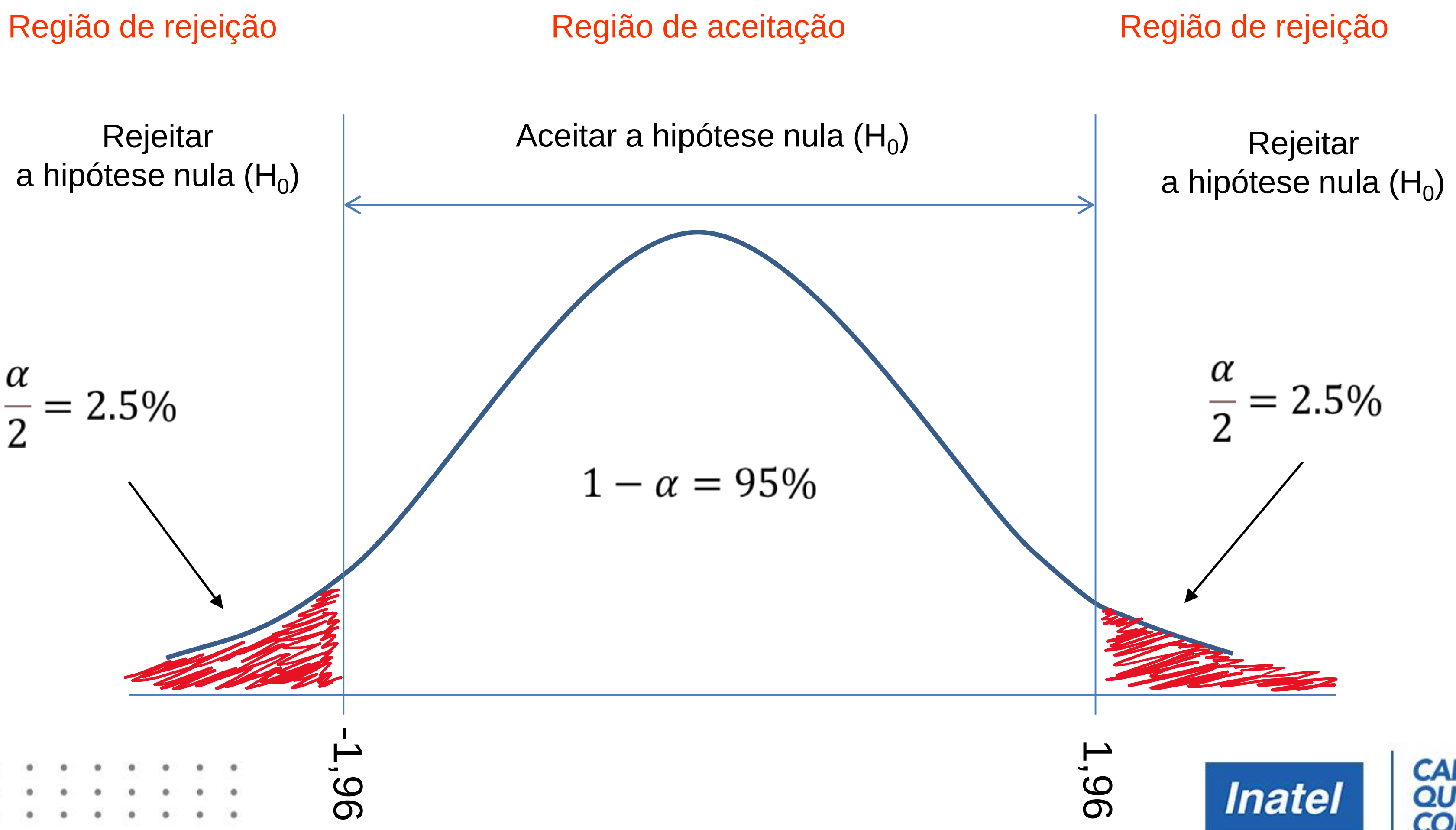
Na prática, define-se o erro máximo tolerado.

Por exemplo: se for definido um nível de significância (probabilidade de cometer EI) de 5%, a decisão tomada aceitará um erro tipo I de no máximo 5% e a confiança (grau de confiança) desta decisão será de 95%.

No exemplo dos portadores de Alzheimer, quais seriam os limites de aceitação da hipótese de nulidade ao nível de significância de 5%?

**Exemplo:** Nível de significância $\alpha = 5\%$

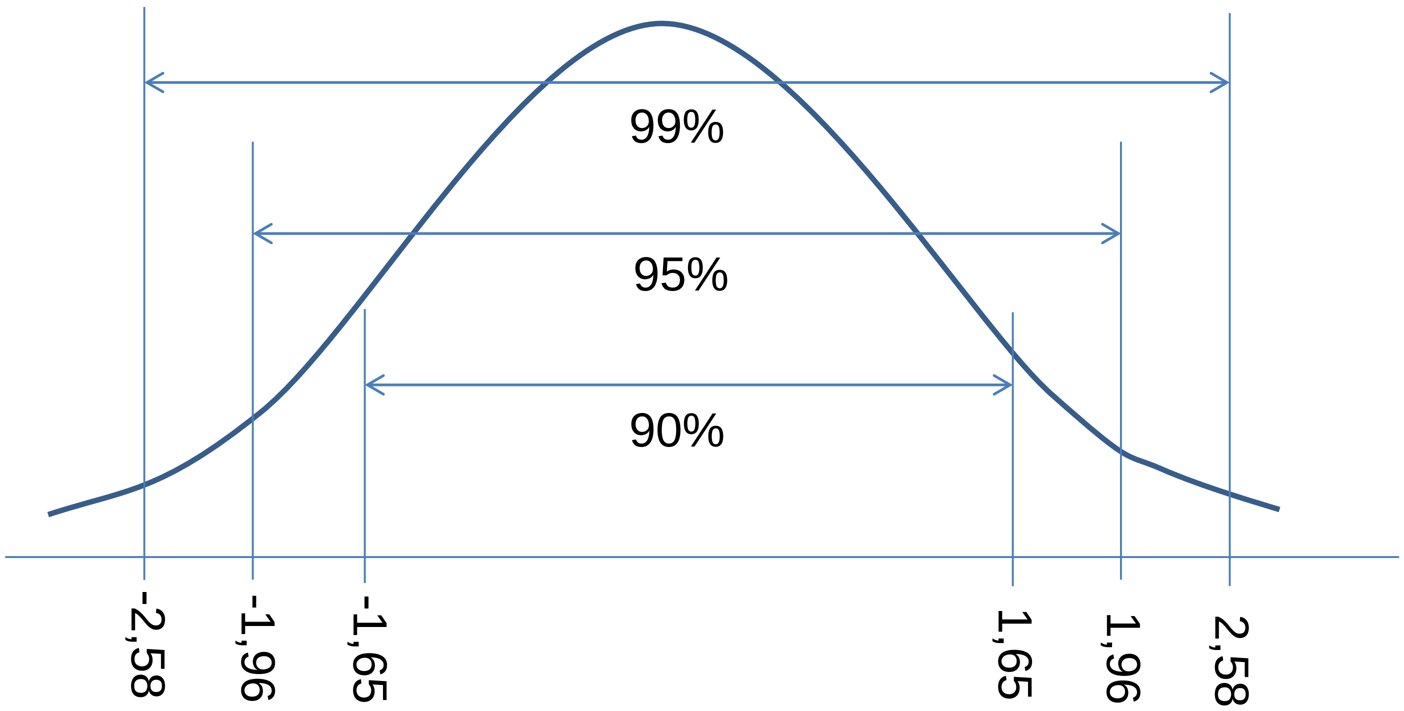
Admitir que a distribuição amostral das médias é Gaussiana



# Exemplo:

5% = 0,05

0,05/2 = 0,025



## Exemplo:

No exemplo dos portadores de Alzheimer, quais seriam os limites de aceitação da hipótese de nulidade ao nível de significância de 5%?

5% = 0,05

0,05/2 = 0,025

Na tabela padronizada, para A = 0,025, tem-se z = -1,96

$$z = \frac{x - \mu}{\sigma}$$

$$\mu = n.p = 50.0,2 = 10$$

$$\sigma = \sqrt{n.p.q} = 2,8284$$

$$\boxed{q = 1\text{-}p = 1\text{-}0,2 = 0,8}$$

# Exemplo:

5% = 0,05

0,05/2 = 0,025

Na tabela, para A = 0,025, tem-se z = -1,96

$$z = \frac{x - \mu}{\sigma}$$

$$\mu = n.p = 50.0,2 = 10$$

$$\sigma = \sqrt{n.p.q} = 2,8284$$

$$\pm 1,96 = \frac{x - 10}{2,8284}$$

$$4,46 \le x \le 15,54$$

# Exemplo:

5% = 0,05

0,05/2 = 0,025

Na tabela, para A = 0,025, tem-se z = -1,96

$$z = \frac{x - \mu}{\sigma}$$

$$\mu = n.p = 50.0,2 = 10$$

$$\sigma = \sqrt{n.p.q} = 2,8284$$

$$\pm 1,96 = \frac{x - 10}{2,8284}$$

$$4,46 \leq x \leq 15,54$$

16 $\neq$ 10,
rejeito $H_0$

**Conclusão**:

Para nível de significância de 5%, $H_o$ deve ser aceita para $5 \leq x \leq 15$

# Rosimara Salgado

Professora
Coordenadora do NEaD

rosimara@inatel.br