

Teste de Hipóteses

Desejamos testar hipóteses que ajudem a comprovar que o medicamento funcione ou não, para tal temos um conjunto de dados de teste, no qual desejamos saber o efeito do remédio.

Para tal, devemos seguir o roteiro para testar hipóteses:

1. Estabelecer as hipótese nula e alternativa;
2. Definir a forma da região crítica, com base na hipótese alternativa;
3. Identificar a distribuição do estimador e obter sua estimativa;
4. Fixar α e obter a região crítica;
5. Concluir o teste com base na estimativa e na região crítica.

Problema

Desejamos testar se um certo tratamento proposto para a doença é eficaz. Uma amostra aleatória de tamanho $n = 30$ é selecionada entre os indivíduos doentes que foram submetidos ao tratamento. Representamos as concentrações dos indivíduos da amostra por X_1, X_2, \dots, X_{30} . Sabe-se que para $i = 1, 2, \dots, 30$, tem-se $X_i \sim N(\mu, 36)$, sendo $\mu = 14$ ou $\mu = 18$ dependendo se o tratamento é eficaz ou não.

```
In [ ]: # Importando os dados
import pandas as pd
dados = pd.read_excel('./data/experimento_medicamento.xlsx')

# Separando o conjunto de dados em sadios, tratamento e placebo.
sadios = dados[['Gênero', 'Grupo de controle']]
tratamento = dados[['Gênero.1', 'Medicamento']]
placebo = dados[['Gênero.2', 'Placebo']]
```

1. Estabelecendo Hipóteses

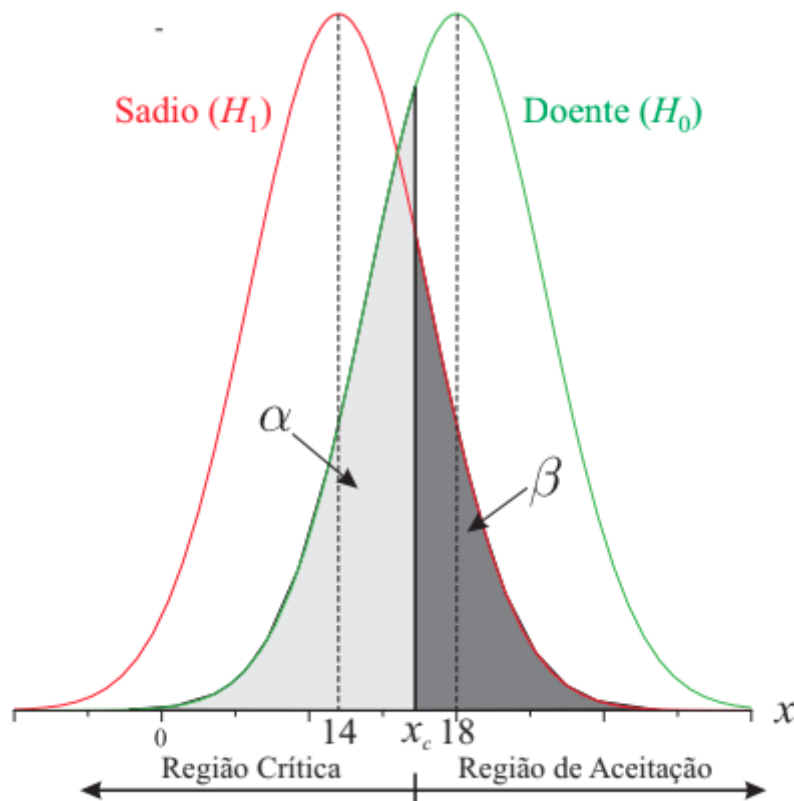
Para a hipótese nula, devemos mostrar que nada mudou desde o uso do remédio, logo H_0 deverá ser o parâmetro para pessoas doentes. Caso o tratamento seja eficaz, iremos estabelecer como uma hipótese alternativa (H_1). Logo, temos as seguintes hipóteses:

$H_0: \mu = 18$ (o tratamento não é eficaz)

$H_1: \mu < 18$ (o tratamento é eficaz)

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

Como montado em sala de aula é necessário estabelecer as regiões críticas da distribuição, conforme a imagem:



3. Identificando a distribuição do estimador e obtendo sua estimativa

Para tal, o problema estabelece uma variável aleatória com distribuição normal $X_i \sim N(\mu, 36)$, que para pessoas sadias $\mu = 14$ e pessoas doentes $\mu = 18$.

4. Fixar μ e obter a região crítica

Para esse teste de hipótese, vamos estabelecer que $\alpha = 3\% = 0.03$, logo precisamos estabelecer a região crítica a partir de α . Vale observar que a região crítica observada da distribuição normal, como a região crítica é unilateral ($H_1: \mu < 18$), logo vamos obter Z_c para a cauda esquerda, nesse caso $P(Z < Z_c) = 0,05$ logo Z_c (tab. Normal de 0.03 da cauda esquerda) = -1.881

```
In [ ]: ZC = -1.88
```

Temos a fórmula da distribuição normal, em que $Z \sim N(0,1)$: $Z_c = X_c - \mu / (\sigma / \sqrt{n})$

Correspondendo à $X_c = 18 + Z_c(6/\sqrt{30})$, em Python, podemos montar uma fórmula geral para quaisquer grupos (com $\alpha = 3\%$):

```
In [ ]: import math

def valor_critico(media, variancia, tam_populacao):
    return media + ZC*(math.sqrt(variancia/tam_populacao))
```

5. Concluir o teste com base na estimativa e na região crítica

Vamos aprofundar mais nos grupos que realizaram o tratamento com o remédio e com o placebo e também com o grupo de controle, para garantir que nossa hipótese pode ser comprovada por pessoas saudáveis.

Para isso, devemos reconhecer o valor crítico para os nossos parâmetros:

```
In [ ]: Xc = valor_critico(18, 36, 30)
```

Grupo Medicamento

Obtendo o valor crítico para nossa população (`Xc`), podemos enfim constatar nossa hipótese

```
In [ ]: True if tratamento['Medicamento'].mean() < Xc else False
```

```
Out[ ]: False
```

Com isso, comprovamos que a Hipótese Alternativa é falsa (`False`), e não conseguiu comprovar a eficiência do medicamento.

Grupo Placebo

```
In [ ]: True if placebo['Placebo'].mean() < Xc else False
```

```
Out[ ]: False
```

Para o grupo de doentes que fizeram o teste com placebo, o teste de hipótese nula se comprova

Grupo de controle

```
In [ ]: True if saudios['Grupo de controle'].mean() < Xc else False
```

```
Out[ ]: True
```

Da mesma forma, ocorre com o grupo de controle, que se comprova verdadeiro para pessoas do Grupo de Controle.

Grupo Medicamento por gênero

Vamos observar os grupos por gênero separadamente, para verificar se algum grupo se destaca para a devida hipótese

```
In [ ]: # Separando os Grupos de Tratamento por Gênero
tratamento_feminino = tratamento.loc[tratamento['Gênero.1'] == 'F']
tratamento_masculino = tratamento.loc[tratamento['Gênero.1'] == 'M']
```

Grupo Feminino

```
In [ ]: # Atualizando o valor crítico para a população masculina
Xc = valor_critico(18, 36, int(tratamento_masculino['Medicamento'].count()))

In [ ]: # Testando se a hipótese Alternativa é verdadeira
True if tratamento_feminino['Medicamento'].mean() < Xc else False

Out[ ]: False
```

Grupo Masculino

```
In [ ]: # Atualizando o valor crítico para uma população de 15 pessoas
Xc = valor_critico(18, 36, int(tratamento_masculino['Medicamento'].count()))

In [ ]: # Testando se a hipótese Alternativa é verdadeira
True if tratamento_masculino['Medicamento'].mean() < Xc else False

Out[ ]: False
```