Teste de Hipóteses

- estimadores
- intervalos de confiança

A partir de agora: situações em que temos que

(a) escolher ou refutar uma hipótese sobre o parâmetro desconhecido

(b) decidir entre duas hipóteses sobre o parâmetro desconhecido

Suponha que um conjunto de dados seja modelado como a realização das v.a.'s $X_1, X_2, \ldots X_n$.

A distribuição de X_i é conhecida, mas tem parâmetros desconhecidos. Vamos propor hipóteses sobre esses parâmetros desconhecidos.

Em particular vamos estudar distribuições com apenas um parâmetro, θ, desconhecido, e propor hipóteses para θ.

- **Hipótese (Estatística)**: qualquer afirmação que se faça sobre um parâmetro desconhecido.

$\Longrightarrow$ Modelamento Probabilístico do Problema $\Longrightarrow$

$\Longrightarrow$ Identificação da A.A.S. $X_1, X_2, \ldots X_n$

Qual a distribuição de X_i?

Sabemos qual é, mas há um ou mais parâmetros desconhecidos.

Se X_i tiver distribuição de Bernoulli de parâmetro p desconhecido, as hipóteses abaixo são plausíveis.

Podemos interpretá-las, pensando em uma moeda que não sabemos se é honesta.

A hipótese nula afirma que ela é honesta (p=0.5 de sair cara ou coroa) enquanto a hipótese alternativa diz que ela não é honesta, e teria p=0.8 de sair coroa.

$$H_0 : p = 0,5$$

contra a hipótese alternativa

$$H_1 : p = 0,8$$

Como julgar se a Hipótese é verdadeira ou não?

Uma pergunta mais adequada: Aceito ou Rejeito a Hipótese?

Com que Probabilidade Aceito ou Rejeito a Hipótese?

Para aceitar ou rejeitar uma Hipótese proposta precisamos de dois conceitos:

(2) **Teste Estatístico.**

(3) **Regra de Decisão.**

(2)**Teste Estatístico: é qualquer função da amostra, (É uma Estatística ou um Estimador)**

$$T = h(X_1, X_2, \ldots X_n),$$

cujo valor numérico (uma estimativa $t = h(x_1, x_2, \ldots x_n)$) pode ser usado para decidir sobre rejeitar ou aceitar uma Hipótese.

(3) **Qual é a Regra de Decisão?**

(5)**Nível de Significância (α) do Teste de Hipóteses .**

A probabilidade α, de rejeitar uma hipótese verdadeira, é também chamada probabilidade de cometer o *Erro Tipo I*

Testes sobre a Média de uma População com Variância Conhecida

Exemplo 3- De uma população normal com variância 36 toma-se uma A.A.S. de tamanho 16, obtendo~~s~~ $\bar{X} = 43$. Ao nível de significância de 10%, testar as hipóteses.

$$\begin{cases} H_0: & \mu = 45 \\ H_1: & \mu \neq 45 \end{cases}$$

(1) **Identificação da A.A.S.:**

(1.a) $n = ?$

(1.b) **Vc tem em mãos uma realização das v.a.'s** $X_1, X_2, \ldots X_n$

(1.c) **Qual a distribuição de** X_i**?**

(2) **Teste Estatístico:**

(2.a) **Quem é o Teste Estatístico?**

(2.b) **Qual é a distribuição do Teste Estatístico?**

(3) **Qual é a Regra de Decisão?**

(3) **Qual é a Região Crítica?**

(1) **Identificação da A.A.S.:**

(1.a) $n = 16$

(1.b) **Vc tem em mãos uma realização das v.a.'s** $X_1, X_2, \dots X_{16}$

(1.c) **Qual a distribuição de** X_i**?**

$$X_i \sim N(\mu, \sigma^2) \text{ com } \sigma^2 = 36$$

(2) **Teste Estatístico:**

(2.a) **Quem é o Teste Estatístico?**

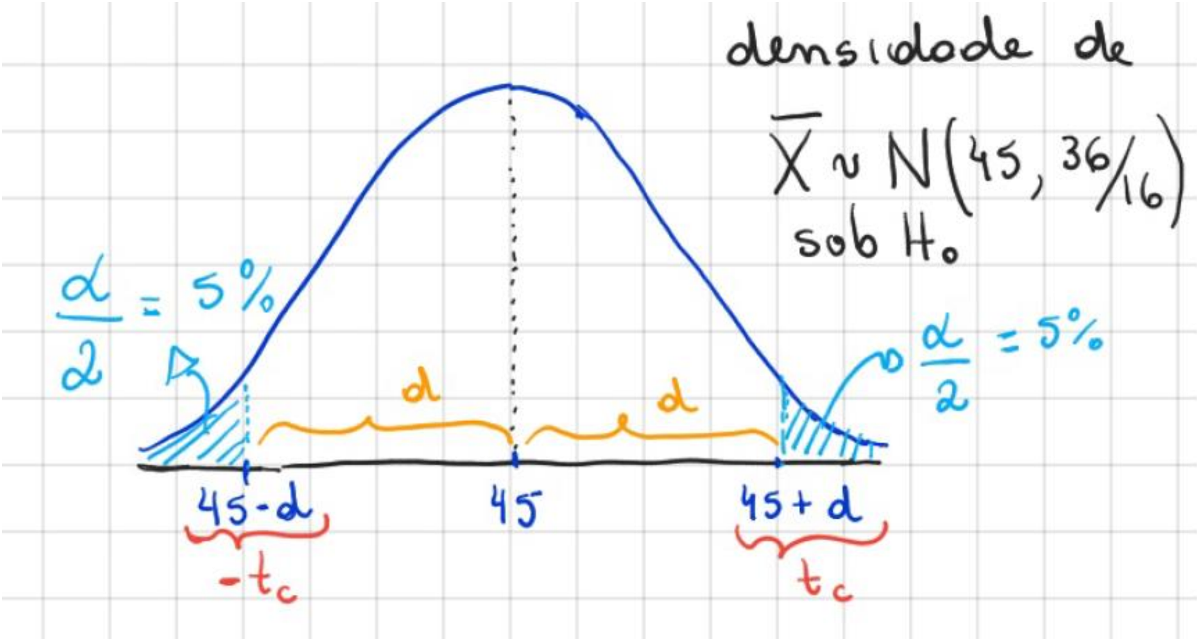
$$\bar{X} = \frac{\sum_{i=1}^{16} X_i}{16}$$

(2.b) **Qual é a distribuição do Teste Estatístico?**

$$\bar{X} \sim N(\mu, \sigma^2/n) \text{ com } \sigma^2 = 36/16$$

(3) **Qual é a Regra de Decisão?**

A encontramos a partir do nível de significância $\alpha = 10\%$


densidade de
$\bar{X} \sim N(45, 36/16)$
sob H_0
$\frac{\alpha}{2} = 5\%$
$\frac{\alpha}{2} = 5\%$
d
d
45-d
45
45+d
$-t_c$
t_c

Concluindo

$$-t_c = 42,54 \ \ \textbf{e} \ \ t_c = 47,46$$

A Região de Aceitação de H_0 é

$$\{\bar{X} : 42,54 < \bar{X} < 47,46\}$$

A amostra forneceu $\bar{X} = 43$ que pertence à Região de Aceitação de H_0.

Logo, aceitamos H_0 com 10% de chance de aceitarmos uma Hipótese Falsa.

Ou ainda, aceitamos H_0 com 10% de chance de cometermos Erro Tipo I (Aceitar H_0 quando ela é falsa).

Poderíamos ter usado o conceito de p-valor:

$$p\text{-valor} = \mathbb{P}\,(X \geq \overline{x}\,;\ \text{sob } H_0)$$

→ $\overline{x}$ é o valor encontrado após a amostragem das X_i's

→ $\overline{x}$ é a estimativa de $\overline{X}$

Se p-valor < α então rejeitamos H_0, pois o valor de $\overline{x}$ encontrado não tem como ocorrer dentro da hipótese H_0.

Um exemplo extremo:

Na Antártida até bem pouco tempo atrás as temperaturas médias eram de $-50°C$.

Digamos que foram encontradas temperaturas médias em 2020, de $-5°C$.

Supondo que as temperaturas $X_1, X_2, \ldots, X_n$ tenham distribuição normal c/ parâmetros μ (desconhecida) e σ^2 (conhecida), temos que

$\bar{X} \sim N(\mu, \sigma^2)$ e $H_0: \mu = -50°C$

$H_1: \mu > -50°C$

p-valor $= \mathbb{P}(\bar{X} \geq -5;\ \underbrace{\mu = -50}_{})$

esta circunstância
não é mais geradora
de temperaturas
tão "altas".

p-valor $< \alpha$

e rejeitamos H_0