



Teste de Hipótese para Médias, Desvio padrão e Variâncias

Universidade Estadual de Santa Cruz

Apresentado e
editado por:
Ângela Marim
Marcella Cunha

Introdução

- Por vezes, levantamos *hipóteses* a respeito de uma característica da população, seja por modificações ocasionadas por fenômenos da natureza, seja por modificações ocasionadas pela implementação de uma nova técnica, dentre outros fatores.



Conceitos fundamentais

- Deste modo, precisamos de técnicas para nos dizer se refutamos ou não a hipótese levantada.
- Naturalmente surgem duas hipóteses:

A **hipótese nula (ou da existência)** que é uma declaração sobre os valores de um ou mais parâmetros. Esta hipótese representa o *status quo* (estado atual) e, geralmente, não é rejeitada a menos que o resultado da amostra implica fortemente que ela é falsa.

A **hipótese alternativa** é uma declaração que contradiz a hipótese nula. Muitas vezes esta hipótese é chamada de *hipótese de pesquisa*, ou seja, é aquilo que o pesquisador quer provar.

Formulação das hipóteses

- Genericamente podemos estar interessados nas seguintes hipóteses:

HIPÓTESE BILATERAL

$$H_0: \mu = \mu_0$$
$$H_a: \mu \neq \mu_0$$

$$H_0: \sigma_1^2 / \sigma_2^2 = \sigma_0^2$$
$$H_a: \sigma_1^2 / \sigma_2^2 \neq \sigma_0^2$$

HIPÓTESE UNILATERAL Á DIREITA

$$H_0: \mu = \mu_0$$
$$H_a: \mu > \mu_0$$

$$H_0: \sigma_1^2 / \sigma_2^2 = \sigma_0^2$$
$$H_a: \sigma_1^2 / \sigma_2^2 > \sigma_0^2$$

HIPÓTESE UNILATERAL Á ESQUERDA

$$H_0: \mu = \mu_0$$
$$H_a: \mu < \mu_0$$

$$H_0: \sigma_1^2 / \sigma_2^2 = \sigma_0^2$$
$$H_a: \sigma_1^2 / \sigma_2^2 < \sigma_0^2$$



Regra para tomadas de decisões

- Um vez formulado as hipóteses, o teste estatístico irá definir apropriadamente quais regiões representam cada uma das hipóteses formuladas. Geralmente definimos tais regiões como:

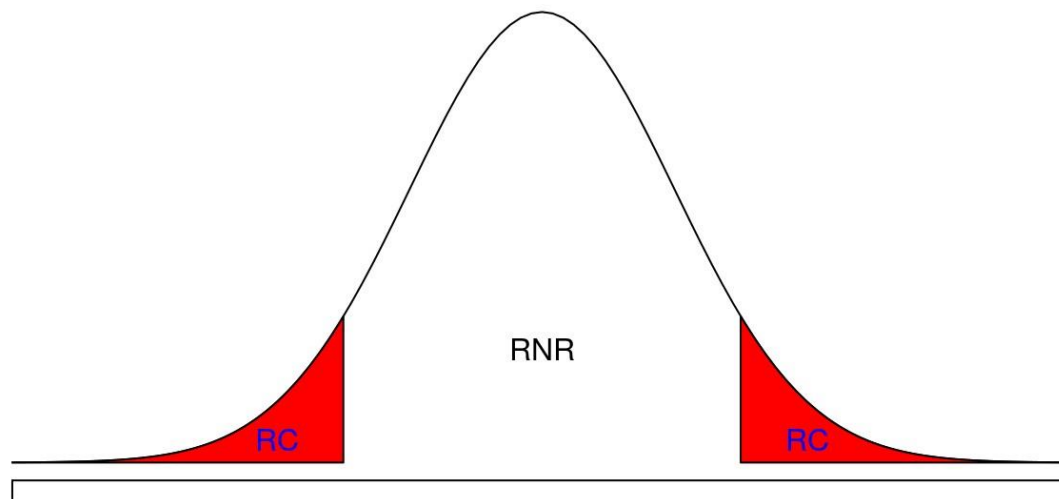
Região de não rejeição - RNR é o intervalo de valores de uma *estimativa* em que não há evidências contra a hipótese nula (H_0).

Região de rejeição (ou região crítica - RC) é o intervalo de valores de uma *estimativa* em que há evidências a favor da hipótese alternativa (H_1).

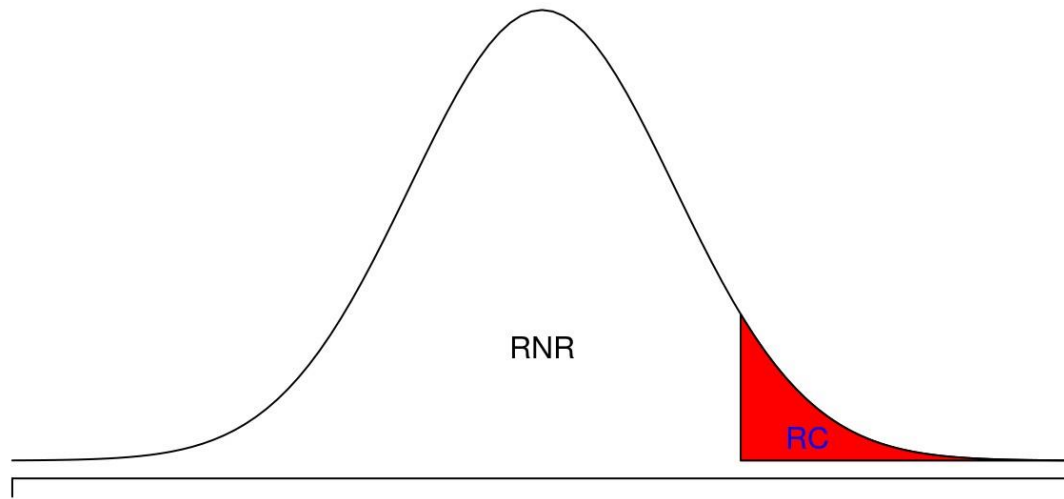
Visualização gráfica das hipóteses

- Lembrem-se que tanto o estimador x quanto p tem distribuição normal de probabilidade.
- Portanto, as hipóteses são visualizadas da seguinte maneira:

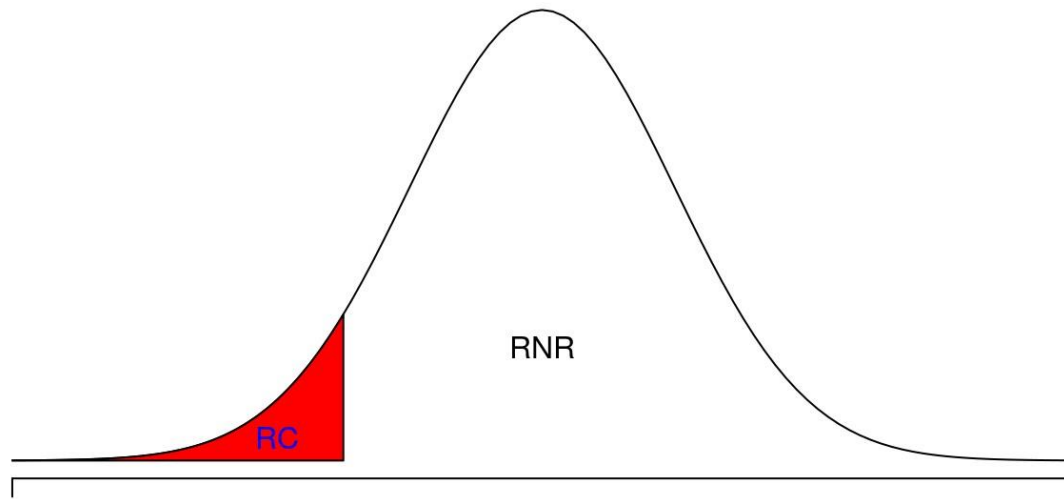
Hipótese bilateral



Hipótese unilateral à direita



Hipótese unilateral à esquerda



Possíveis erros de decisão

- Quando realizamos teste de hipótese, podemos cometer dois tipos de erros:

O **erro tipo I (ou de 1ª espécie)** que ocorre quando nós incorretamente rejeitamos H_0 , ou seja, H_0 é verdadeira e nosso procedimento inferencial baseado na amostra a rejeita.

O **erro tipo II (ou de 2ª espécie)** que ocorre quando nós incorretamente não rejeitamos H_0 , ou seja, H_0 não é verdadeira e nosso procedimento inferencial não consegue detectar esse fato.

- A tabela abaixo resume os possíveis resultados de um teste de hipótese.

Decisão	Na população	
	H_0 é verdadeira	H_0 não é verdadeira
H_0 não é rejeitada	Decisão está correta	Erro do tipo II
H_0 é rejeitada	Erro do tipo I	Decisão está correta

Quantificando os tipos de erros

- O erro do tipo I é geralmente representada pela letra grega α , a mesma utilizada em intervalos de confiança.
 - A probabilidade de cometermos este erro é fornecida por nós mesmos, mas geralmente na literatura costuma-se usar 0,01; 0,05 e 0,10.
 - No entanto, só você que planejou e executou a pesquisa sabe o custo de se cometer este tipo de erro. Portanto, não se prenda a receita de bolos encontrados na literatura (0,01, 0,05 e 0,10).

- O erro do tipo II é geralmente representada pela letra grega β , e o seu cálculo não será apresentado neste material por duas razões a saber:
 - Para quantificar este erro, é necessário supor que H_a seja verdadeira, condição um pouco irreal para o nosso público.



Testes de Hipóteses para Média de Populações Normais - Variância conhecida e desconhecida

Universidade Estadual de Santa Cruz

Ivan BezerraAllaman

Considerando variância conhecida

Introdução

- Nestes casos utiliza-se a seguinte estatística de teste:

$$Z = \frac{\underline{X} - \underline{\mu_0}}{\sigma / \sqrt{n}}$$

- Calculada a estatística de teste, chegou a hora de tomarmos uma decisão.
- Há duas regras:
 - a regra de rejeição baseada no critério do valor crítico;
 - a regra de rejeição baseada no critério do *p-valor*.
 - Você pode optar por uma delas.

Regra baseada no critério do valor crítico

- Obviamente que a tomada de decisão irá depender do tipo de hipótese estabelecida (bilateral, unilateral à direita ou a esquerda).
- Logo, teremos as seguintes regras de acordo com as hipóteses estabelecidas:

Hipótese bilateral

Rejeitar H_0 se $Z \leq -Z_{\alpha/2}$ ou se $Z \geq Z_{\alpha/2}$

Hipótese unilateral à direita

Rejeitar H_0 se $Z \geq Z_{\alpha}$

Hipótese unilateral à esquerda

Rejeitar H_0 se $Z \leq -Z_{\alpha}$

Regra baseada no *p*-valor

- O *p*-valor é a probabilidade de se obter uma estatística de teste igual ou mais extrema que aquela observada em uma amostra, sob a hipótese nula. Em termos coloquiais, ele avalia o quão bem os dados da amostra apoiam o argumento de que a hipótese nula é verdadeira, ou seja, o quão compatíveis os seus dados são com a hipótese nula.
- Logo, tem-se a seguinte regra de decisão:

Hipótese bilateral

Rejeitar H_0 se o p -valor $\leq \alpha/2$

Hipótese unilateral à direita

Rejeitar H_0 se o p -valor $\leq \alpha$

Hipótese unilateral à esquerda

Rejeitar H_0 se o p -valor $\leq \alpha$

**Considerando variância
desconhecida**

Introdução

- Neste caso só irá mudar a função densidade de probabilidade que será a *t-Student*, pois todo o raciocínio anterior é válido.
- Lembrem-se que além da média você deverá também estimar a variância da amostra.
- Logo, tem-se a seguinte estatística de teste:

$$t = \frac{X - \mu_0}{s/\sqrt{n}}$$

- Lembrando que a distribuição t tem ν graus de liberdade que é $n - 1$.

Aplicação

1. Uma máquina de misturar fertilizantes é adaptada para fornecer 10 g de nitrato para cada 100 g de fertilizante. 10 porções de 100 g são examinadas, com as seguintes percentagens de nitrato:

9 12 11 10 11 9 11 12 9 10

Há razões para crer que a percentagem de nitrato não é 10g ao nível de 10%?

Elaborando as hipóteses tem-se:

$$H_0 : \mu = 10$$

$$H_a : \mu \neq 10$$

Vamos aos cálculos das estatísticas.

$$X = 10,$$

$$s = 1,174$$

$$n = 10$$

$$s_x = \frac{1,174}{\sqrt{10}} = 0,371$$

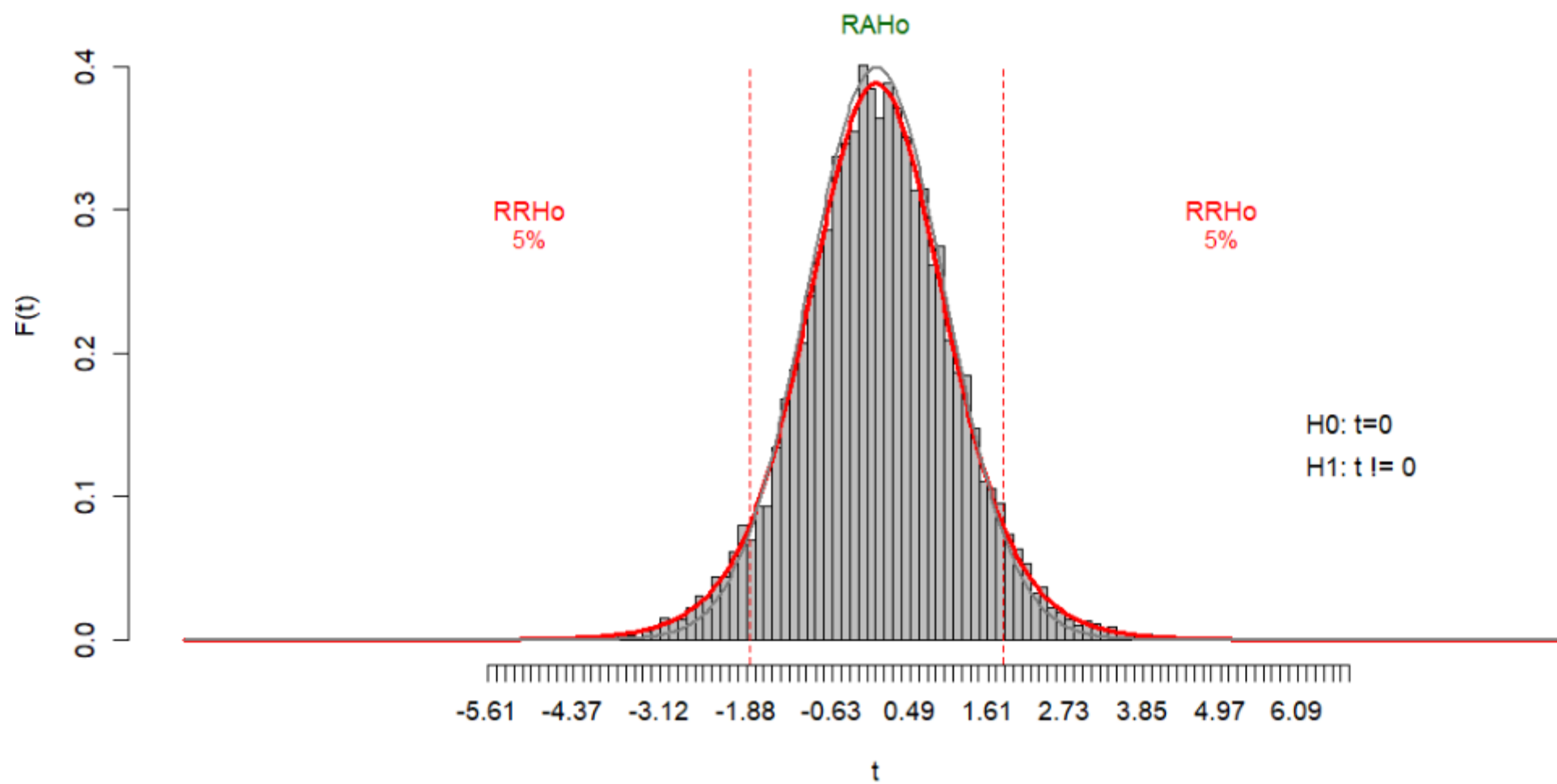
$$t_{calc} = \frac{10,4 - 10}{0,371} = 1,078$$

$$v = n - 1 = 10 - 1 = 9$$

$$p\text{-valor} = pt(t_{calc}, v, lower.tail = F) * 2 = 0,309$$

Logo, como o p-valor (30,9%) é maior que 10%, não rejeitamos H_0 , ou seja, não há razões para crer que a porcentagem de nitrato não é 10%.

Histograma e FDP



**Para a razão de variâncias de
duas populações**

Considerando duas populações Y_1 e Y_2 com médias μ_1 e μ_2 , e variâncias σ_1^2 e σ_2^2 , ou seja

$$Y_1 \sim \mathbf{N}(\mu_1, \sigma_1^2) \quad \text{e} \quad Y_2 \sim \mathbf{N}(\mu_2, \sigma_2^2).$$

Já vimos que a distribuição amostral da razão de duas variâncias amostrais (s_1^2 e s_2^2) possui distribuição F com $n_1 - 1$ graus de liberdade no numerador e $n_2 - 1$ graus de liberdade no denominador.

Intuitivamente:

- ▶ **Se a razão das duas variâncias for próxima de 1, então elas são aproximadamente iguais.**
- ▶ **Em um teste de hipótese para a igualdade de variâncias entre duas populações, verifica-se então se a razão das variâncias está ou não próxima de 1.**

Para o teste é necessário os seguintes requisitos:

- ▶ **Temos uma AAS.**
- ▶ **As duas populações serem independentes.**
- ▶ **As duas populações terem, cada uma, distribuição Normal (essa é uma exigência estrita).**

Sendo assim, usamos a estatística de teste

$$F = \frac{s_1^2}{s_2^2} \sim F_{v_1, v_2},$$

em que $v_1 = n_1 - 1$ graus de liberdade no numerador e $v_2 = n_2 - 1$ graus de liberdade no denominador.

Importante: s_1^2 deve ser sempre a maior das duas variâncias amostrais.

Procedimentos gerais para um teste de hipótese para a diferença de duas variâncias

1. Definir a hipótese nula (H_0) e a alternativa (H_a).
2. Definir o nível de significância α .
- Σ. Definir o tipo de teste, com base na hipótese alternativa.
- {. Calcular a estatística de teste

$$F = \frac{s_1^2}{s_2^2}$$

- \$. Determinar a região crítica (região de rejeição), com base no nível de significância α .
- õ. Conclusão do teste.