Inferência Estatística I

Testes de Hipóteses

Prof. Paulo Cerqueira Jr
Faculdade de Estatística - FAEST
Instituto de Ciências Exatas e Naturais - ICEN

https://github.com/paulocerqueirajr

Introdução

Introdução

- Em estatística, uma hipótese é uma afirmativa sobre um propriedade da população (ex.: média).
- Um teste de hipóteses é um procedimento padrão (regra de decisão) para se testar uma afirmativa sobre uma propriedade da população.

Exemplos:

- A produtividade média de milho em Santa Catarina é de 2300kg/ha. (teste para a média);
- A proporção de alevinos de tilápia do Nilo que atingem o peso adequado em 120 dias é de 54%. (teste para a proporção)
- A sobrevivência de mudas não dependem da época do plantio. (teste qui-quadrado);
- A proporção de fixação de fitoplâncton em dois tipos de solos é a mesma (Teste de comparação de porporções).

Testes de hipóteses

- São ferramentas estatísticas que quantificam quão plausíveis são os resultados observados em uma amostra, podem ou não, ser verdadeiro.
- Além disto, um teste também define um ponto de corte (regra de decisão) para tomarmos a decisão de aceitar ou rejeitar a hipótese testada.
- Veremos testes para afirmações sobre a média e proporção de uma população.

Introdução

> Um criador do *Colossoma macropomum* (tambaqui), criado à densidade de $1,0 \text{ peixe}/m^2/120 \text{ dias}$, afirma que os mesmos tem peso médio de $360,7g$ e uma variância de $30,7g^2$.

Uma amostra com 22 peixes foi formada e vetificou-se que o peso médio foi igual à $358,2g$. Dessa forma, temos duas situações:

1. H_0: O criador está correto. (Hipótese nula)
2. H_1: O criador está errado. (Hipótese alternativa)

O que de fato queremos saber?

```
Se a média é igual a $360,7g$ ou diferente!!
```

Testes de hipóteses

Definição 1 (Hipóteses estatística) É uma afirmação ou conjetura sobre o parâmetro, ou parâmetros, da distribuição de probabilidades de uma característica, X, da população ou de uma v.a.

Definição 2 (Teste de hipóteses) Um teste de hipóteses estatística é o procedimento ou regra de decisão que nos possibilita decidir por H_0 (Hipótese Nula) ou H_1 (Hipótese Alternativa), com base a informação contida na amostra.

Procedimentos gerais

- População: X com f.d ou f.p ($f(x \mid \theta)$).
- θ é um parâmetro desconhecido, e temos alguma hipótese sobre o valor verdadeiro de θ, por exemplo, afirmamos que seu valor é θ_0.
- Observamos uma a.a. de X, e com ela desejamos comprovar ou não tal hipótese.
- Assim, queremos testar

$$H_0 : \theta = \theta_0$$

- Temos também que explicitar a hipótese que aceitaremos caso H_0 seja rejeitada,

$$\underbrace{H_1 : \theta \neq \theta_0}_{\text{Bilateral}} \quad \text{ou} \quad \underbrace{H_1 : \theta < \theta_0}_{\text{Unilateral à esquerda.}} \quad \text{ou} \quad \underbrace{H_1 : \theta > \theta_0}_{\text{Unilateral à direita.}},$$

que dependerá das informações que o problema traz.

Erros associados aos testes de hipóteses

- Devemos tomar como H_0 aquela hipótese que, rejeitada, conduza a um erro de tipo I mais importante de evitar.
- Por exemplo, suponha um experimento para determinar se um produto A é ou não cancerígeno. Após realizado o teste, podemos concluir:

$$\text{(i) A é cancerígeno ou (ii) A não é cancerígeno.}$$

- Cada uma dessas conclusões pode estar errada e temos os dois tipos de erro:

1. concluir que o produto é cancerígeno, quando ele não é.
2. concluir que o produto não é cancerígeno, quando ele é.

Qual o pior erro?

O **segundo erro** é pior, este deve ser o erro tipo I (rejeitar H_0, quando ela é verdadeira), portanto,

$$H_0 : \text{A é cancerígeno.}$$

Erros associados aos testes de hipóteses:

- Os dois erros que podem ser cometidos ao se realizar um teste de hipóteses são:
  - Rejeitar a hipótese nula (H0), quando tal hipótese é verdadeira;
  - Não rejeitar a hipótese nula (H0) quando ela deveria ser rejeitada.
- De forma mais simplificada temos:

Decisão/ Situação	H_0 **verdade**	H_0 **falso**
Rejeita H_0	Erro I	Certo
Não rejeita H_0	Certo	Erro II

Erros associados aos testes de hipóteses

- Tais erros são expressos em termos de probabilidade.
- O nível de significância ou probabilidade do Erro I é dada por

$$\alpha = P(\text{Rejeitar } H_0 \mid H_0 \text{ verdadeira}).$$

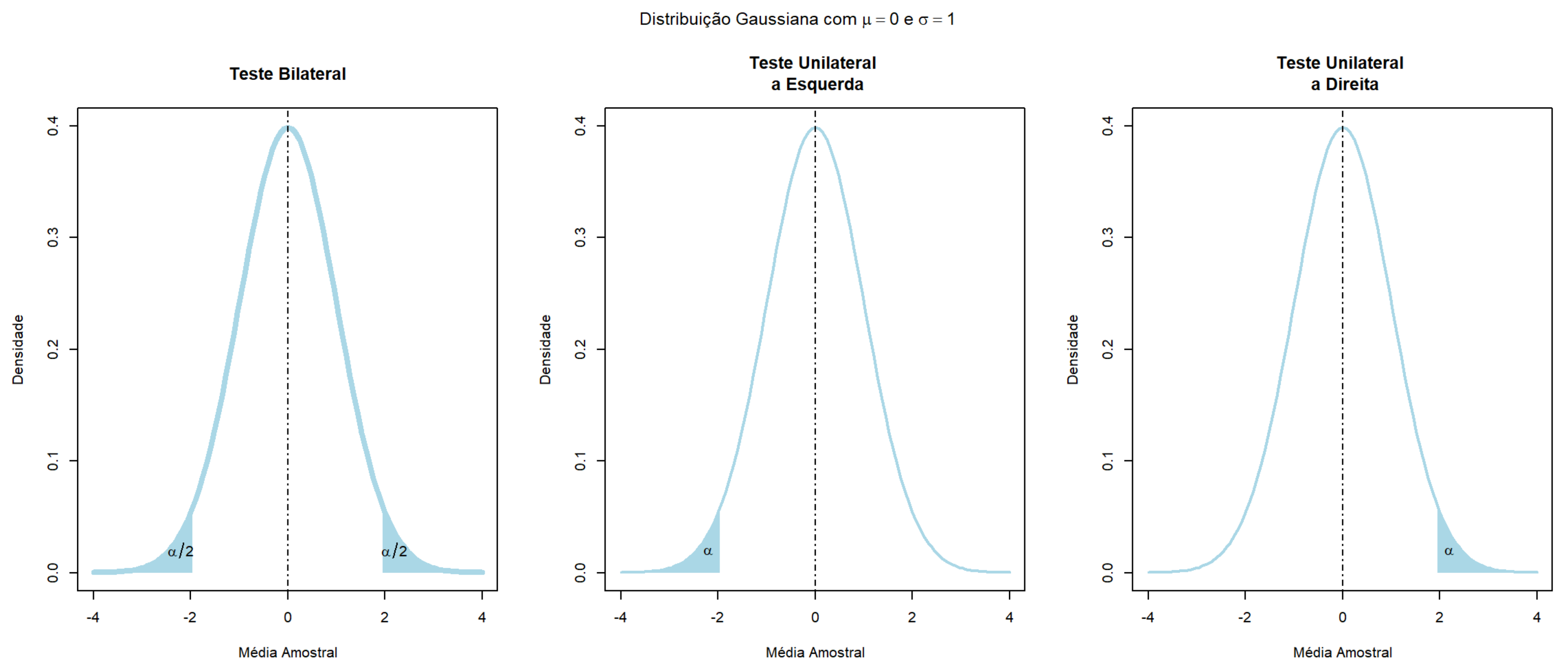
- Em geral o `nível de siginificância` gira em torno de $1\%, 5\%, 10\%$.
- A probabilidade do Erro II é dada por

$$\beta = P(\text{Não rejeitar } H_0 \mid H_0 \text{ falsa}).$$

- O poder o teste é dado por:

$$\text{Poder} = 1 - \beta = 1 - P(\text{Não rejeitar } H_0 \mid H_0 \text{ falsa}) = P(\text{Rejeitar } H_0 \mid H_0 \text{ falsa}).$$

Erros associados aos testes de hipóteses



$H_0 : \mu = \mu_0$ $\quad$ $H_0 : \mu = \mu_0$ $\quad$ $H_0 : \mu = \mu_0$

$H_a : \mu \neq \mu_0$ $\quad$ $H_a : \mu < \mu_0$ $\quad$ $H_a : \mu > \mu_0$

Construindo testes de hipóteses:

1. Estabelecer a hipótese nula. A hipótese alternativa é complementar à hipótese nula.
2. Definir a forma da região de aceitação com base na hipótese nula.
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.

Testes para uma amostra

Testes para uma amostra

Existem uma série de testes de hipóteses, cada um com sua finalidade.

Para uma amostra, temos alguns testes:

- Teste para a proporção;
- Teste para média populacional com variância conhecida;
- Teste para média populacional com variância desconhecida (Teste t de **Student**).

Teste para proporção populacional

Considere o problema de testar a hipótese de que a proporção de sucessos de um ensaio de Bernoulli é igual a um valor específico, p_0.

1. Definição das hipóteses

$$
\begin{aligned}
H_0: p = p_0 \quad &\text{vs} \quad H_1: p \neq p_0 \quad \text{(bilateral)} \\
H_0: p = p_0 \quad &\text{vs} \quad H_1: p > p_0 \quad \text{(unilateral direita)} \\
H_0: p = p_0 \quad &\text{vs} \quad H_1: p < p_0 \quad \text{(unilateral esquerda)}
\end{aligned}
$$

Teste para proporção populacional

2. Estatística de teste

A estatística escolhida é a proporção amostral:

$$\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right) \quad \Rightarrow \text{ Sob } \mathrm{H}_0 \Rightarrow \hat{p} \sim N\left(p_0, \frac{p_0(1-p_0)}{n}\right).$$

3. Região Crítica (RC)

A RC depende da hipótese alternativa considerada:

- Para $H_1 : p \neq p_0 : \; RC = \{\hat{p} \in \mathbb{R} : \hat{p} < p_1 \text{ ou } \hat{p} > p_2\}$
- Para $H_1 : p > p_0 : \; RC = \{\hat{p} \in \mathbb{R} : \hat{p} > p_2\}$
- Para $H_1 : p < p_0 : \; RC = \{\hat{p} \in \mathbb{R} : \hat{p} < p_1\}$

Os passos 4 e 5 dependerão dos valores da amostra, vamos descrevê-los num exemplo.

Exemplo

Um estudo foi realizado para determinar a relação entre uma certa droga e uma anomalia em embriões de frango. Foram injetados 50 ovos fertilizados com a droga no quarto dia de incubação. No vigésimo dia, os embriões foram examinados e 7 apresentaram a anomalia. Deseja-se verificar se a proporção verdadeira é inferior a 25% com um nível de significância de 0,05.

> Solução:

Passo 1: As hipóteses são:

$$H_0 : p = 0.25 \quad \text{vs} \quad H_1 : p < 0.25.$$

Passo 2: Estatística de teste:

$$\hat{p} \sim N\left(p, \frac{p(1-p)}{50}\right), \quad n = 50.$$

Exemplo

Passo 3: Fixamos $\alpha = 0.05$. Sob H_0:

$$\hat{p} \sim N\left(0.25, \frac{0.25(1-0.25)}{50}\right).$$

A região de rejeição será dada por $RC = \{\hat{p} \in \mathbb{R} : \hat{p} < p_1\}$, com p_1 sendo tal que:

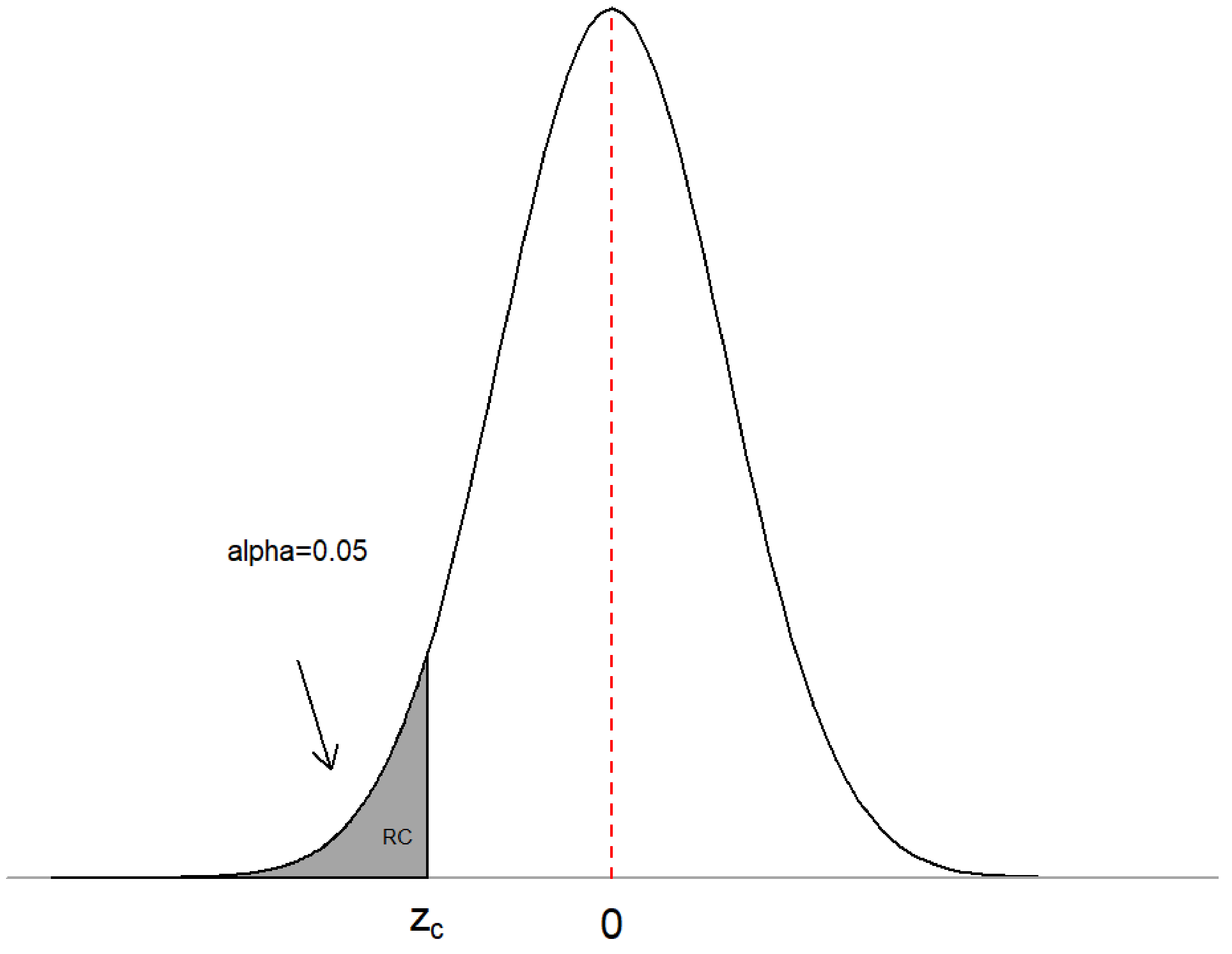
$$\alpha = P(\text{Erro I}) = P(\hat{p} \in RC \mid \hat{p} \sim N\,(0.25, 0.0037)).$$

Ou seja,

$$0.05 = P(\hat{p} < p_1 \mid \hat{p} \sim N\,(0.25, 0.0037)) \iff 0.05 = P\left(Z < \frac{p_1 - 0.25}{\sqrt{0.0037}}\right)$$

Exemplo

Graficamente,



Exemplo

Temos então,

$$z_c = \frac{p_1 - 0.25}{\sqrt{0.0037}}.$$

Consultando o valor tabelado, temos que $z_c = -1.64$. Logo,

$$-1.64 = \frac{p_1 - 0.25}{\sqrt{0.0037}} \iff p_1 = -1.64\sqrt{0.0037} + 0.25 = 0.1963$$

Portanto, a regra de decisão consiste em:

$$\text{Rejeitar } H_0 \text{ se } \hat{p} < 0.1963$$

Passo 4: Com os valores amostrais temos que $\hat{p}_{obs} = \dfrac{7}{50} = 0.14$.

Passo 5: Como $\hat{p}_{obs} = 0.14 \in RC$, rejeitamos H_0 ao nível de significância de 5%. Então, temos o indicativo de que a proporção verdadeira da anomalia em embriões é inferior a 25%.

Teste para a média populacional (Variância conhecida)

Teste para a média populacional (Variância conhecida)

Considere uma amostra aleatória de tamanho n de uma população normal com média μ (desconhecida) e variância σ^2 (conhecida). Suponha que tem-se interesse em verificar as seguintes hipóteses:

1. As hipóteses são as seguintes:

$$H_0 : \mu = \mu_0 \quad \begin{array}{ll} \times & (i) H_1 : \mu \neq \mu_0. \text{ (bilateral)} \\ \times & (ii) H_1 : \mu > \mu_0. \text{ (Unilateral à direita)} \\ \times & (iii) H_1 : \mu < \mu_0. \text{ (Unilateral à esquerda)} \end{array}$$

2. A estatística escolhida é a média.

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right).$$

Teste para a média populacional (Variância conhecida)

3. Fixado um valor de α, devemos construir a RC, supondo H_0 verdadeira. Ou seja, podemos escrever

$$\bar{X} \sim N\left(\mu_0, \frac{\sigma^2}{n}\right).$$

A RC dependerá da hipótese alternativa considerada:

- Para a alternativa (i) : $RC = \{\bar{X} \in \mathbb{R} : \bar{X} < \mu_1 \ ou \ \bar{X} > \mu_1\}$
- Para a alternativa (ii) : $RC = \{\bar{X} \in \mathbb{R} : \bar{X} > \mu_1\}$
- Para a alternativa (iii) : $RC = \{\bar{X} \in \mathbb{R} : \bar{X} < \mu_1\}$

Os passos 4 e 5 dependerão dos valores da amostra, vamos descrevê-los num exemplo.

Exemplo

Uma máquina automática para encher pacotes de café enche-os segundo uma distribuição normal, com média μ e variância igual a $400g^2$. A máquina foi regulada para $\mu = 500g$. Desejamos, periodicamente, colher uma amostra de 16 pacotes e verificar se a produção está sob controle, isto é, se $\mu = 500g$ ou não. Se uma dessas amostras apresentasse uma média $\bar{X} = 492g$, você pararia ou não a produção para regular a máquina?

> Solução:

Passo 1: As hipóteses são:

$$H_0 : \mu = 500\text{g} \times H_1 : \mu \neq 500g.$$

Passo 2: A estatística de teste:

$$\bar{X} \sim N\left(\mu, \frac{400}{16}\right).$$

Exemplo

Passo 3: Fixaremos $\alpha = 0,01$ e sob a suposição de H_0 ser verdadeira, temos

$$\bar{X} \sim N\left(500, \frac{400}{16}\right), \rightarrow \bar{X} \sim N\left(500, 25\right).$$

Dessa forma, temos que

$$RC = \{\bar{X} \in \mathbb{R} : \bar{X} < x_1 \text{ ou } \bar{X} > x_2\}.$$

Com,

$$0.01 = P\left(\bar{X} \in RC \mid \bar{X} \sim N\left(500, 25\right)\right) \Leftrightarrow 0.01 = P\left(\bar{X} < x_1 \text{ ou } \bar{X} > x_2 \mid \bar{X} \sim N\left(500\right.\right.$$

Para resolver a expressão acima, tomaremos x_1 e x_2 tais que

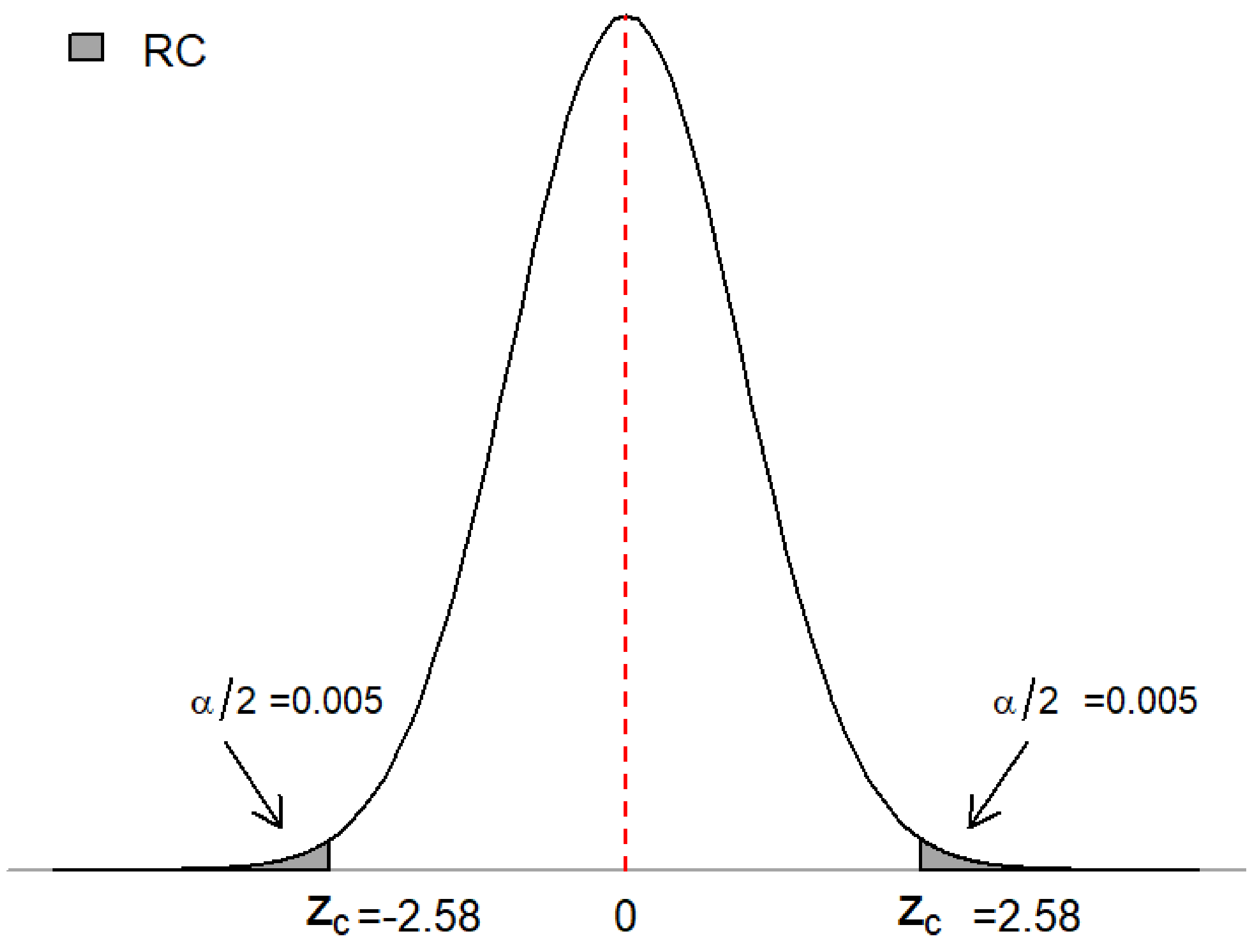
$$0.005 = P\left(\bar{X} < x_1 \mid \bar{X} \sim N\left(500, 25\right)\right) \text{ e } 0.005 = P\left(\bar{X} > x_2 \mid \bar{X} \sim N\left(500, 25\right)\right)$$

Ou simplesmente,

$$0.005 = P\left(Z < \frac{x_1 - 500}{\sqrt{25}}\right) \text{ e } 0.005 = P\left(Z > \frac{x_2 - 500}{\sqrt{25}}\right)$$

Exemplo

Graficamente,



Exemplo

Assim temos que,

$$-2.58 = \frac{x_1 - 500}{\sqrt{25}} \text{ e } 2.58 = \frac{x_2 - 500}{\sqrt{25}}.$$

Isolando temos que $x_1 = 487.1$ e $x_2 = 512.9$. Portanto a regra de decisão será,

$$\text{Rejeitar } H_0 \text{ se } \bar{X} < 487.1 \text{ ou } \bar{X} > 512.9.$$

Passo 4: Com os valores amostrais temos que $\bar{X}_{obs} = 492g$.

Passo 5: Como $\bar{X}_{obs} = 492g$ não pertence à RC, ou seja, não rejeitamos H_0 ao nível de significância de 1%. Dessa forma, temos indicativo de que a máquina está regulada e não há necessidade de parar a produção.

Teste para a média populacional (Variância desconhecida)

Teste para a média populacional (Variância desconhecida)

Considere uma amostra aleatória de tamanho n de uma população normal com média μ (desconhecida) e variância σ^2 (desconhecida). Suponha que tem-se interesse em verificar as seguintes hipóteses:

1. As hipóteses são as seguintes:

$$H_0 : \mu = \mu_0 \quad \begin{array}{ll} \times & (i) H_1 : \mu \neq \mu_0. \text{ (bilateral)} \\ \times & (ii) H_1 : \mu > \mu_0. \text{ (Unilateral à direita)} \\ \times & (iii) H_1 : \mu < \mu_0. \text{ (Unilateral à esquerda)} \end{array}$$

2. A estatística escolhida é a proporção amostral.

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} = \frac{\sqrt{n}(\bar{X} - \mu)}{S} \sim t_{n-1}.$$

Teste para a média populacional (Variância desconhecida)

3. Fixado um valor de α, devemos construir a RC, supondo H_0 verdadeira. Ou seja, podemos escrever

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{\sqrt{n}(\bar{X} - \mu_0)}{S} \sim t_{n-1}.$$

- A RC dependerá da hipótese alternativa considerada:
- Para a alternativa (i) : $RC = \{T \in \mathbb{R} : T < t_1 \; ou \; T > t_2\}$
- Para a alternativa (ii) : $RC = \{T \in \mathbb{R} : T > t_2\}$
- Para a alternativa (iii) : $RC = \{T \in \mathbb{R} : T < t_1\}$

Os passos 4 e 5 dependerão dos valores da amostra, vamos descrevê-los num exemplo.

Exemplo

Os registros dos últimos anos de um colégio atestam para calouros admitidos uma nota média 115 (teste vocacional). Para testar a hipótese de que a média de uma nova turma é a mesma das turmas anteriores, retirou-se, ao acaso, uma amostra de 20 notas, obtendo-se média 118 e desvio padrão 20. Use $\alpha = 0.05$.

> Solução:

Passo 1: As hipóteses são:

$$H_0 : \mu = 115 \times H_1 : \mu \neq 115.$$

Passo 2: A estatística de teste:

$$T = \frac{\sqrt{n}(\bar{X} - \mu)}{S} \sim t_{20-1=19}.$$

Exemplo

Passo 3: Fixaremos $\alpha = 0,05$ e sob a suposição de H_0 ser verdadeira, temos

$$T = \frac{\sqrt{n}(\bar{X} - 115)}{S} \sim t_{19}.$$

Dessa forma, temos que

$$RC = \left\{T \in \mathbb{R} : T < t_1 \text{ ou } T > t_2\right\}.$$

Com,

$$0.05 = P\left(T \in RC \mid T \sim t_{19}\right) \Leftrightarrow 0.05 = P\left(T < t_1 \text{ ou } T > t_2 \mid T \sim t_{19}\right).$$

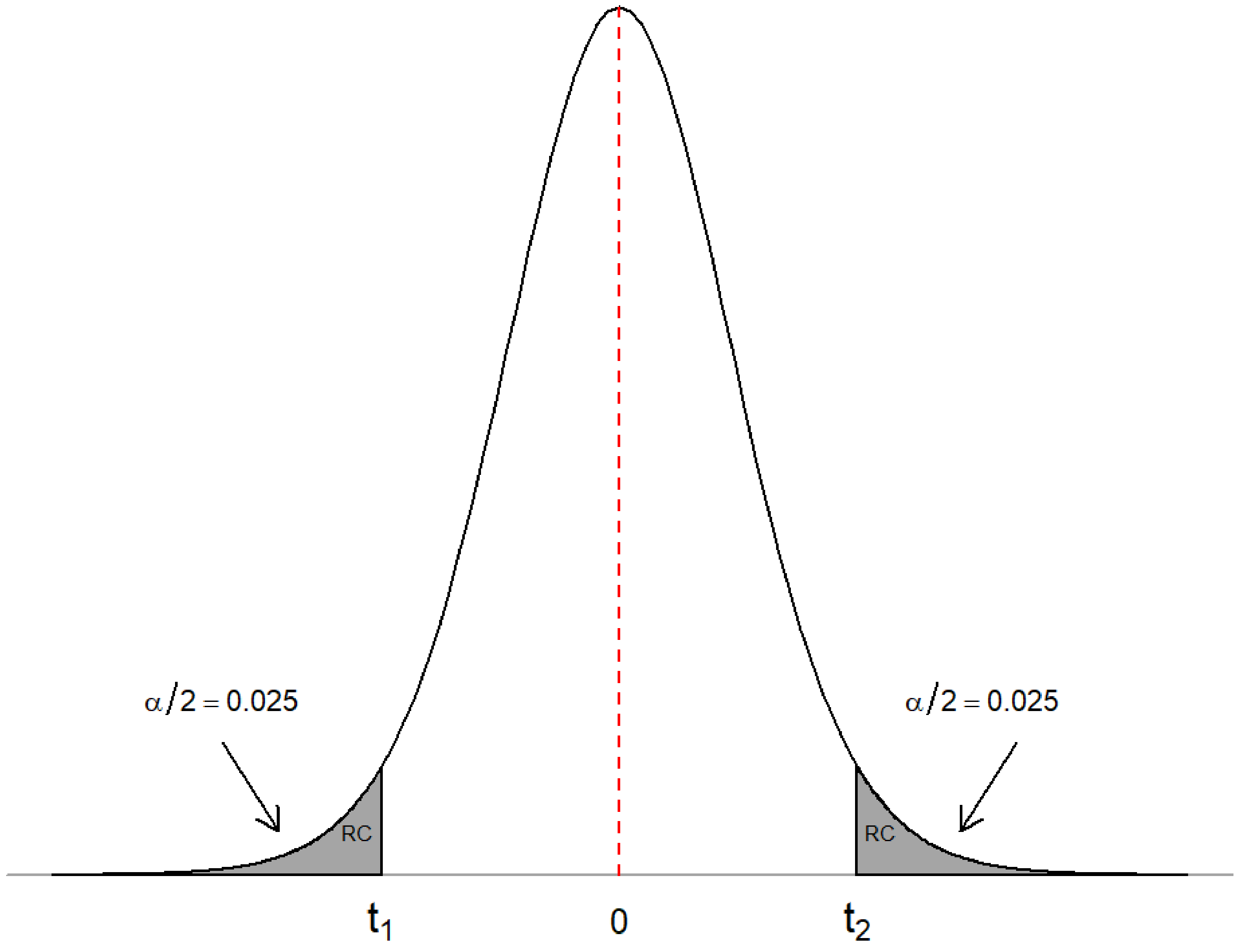
Para resolver a expressão acima, tomaremos t_1 e t_2 tais que

$$0.025 = P\left(T < t_1 \mid T \sim t_{19}\right) \text{ e } 0.025 = P\left(T > t_2 \mid T \sim t_{19}\right)$$

Note que os valores de t_1 e t_2 podem ser obtidos através da tabela da distribuição t-Student.

Exemplo

Graficamente,



Exemplo

Com os valores de $t_1 = -2.093$ e $t_2 = 2.093$, temos que a regra de decisão será:

$$\text{Rejeitar } H_0 \text{ se } T < -2.093 \text{ ou } T > 2.093.$$

Passo 4: Com os valores amostrais, temos que:

$$T_{obs} = \frac{\sqrt{20}(118 - 115)}{20} = 0.6708.$$

Passo 5: Como $T_{obs} = 0.6708$ não pertence à região crítica (RC), ou seja, não rejeitamos H_0 ao nível de significância de 5%. Dessa forma, temos indicativo de que a média de uma nova turma seria a mesma das turmas anteriores.

Teste para a Variância de uma Normal

- $X_1, X_2, \cdots, X_n$ uma a.a. de $X \sim N(\mu, \sigma^2)$
- A estatística a ser usada para o teste é:

$$Q = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

Exemplo: Uma das maneiras de manter sob controle a qualidade de um produto é controlar sua variabilidade. Uma máquina de encher pacotes de café está regulada para enchê-los com média de 500 g e desvio padrão de 10 g. O peso de cada pacote segue uma distribuição $N(\mu, \sigma^2)$. Colheu-se uma amostra de 16 pacotes e observou-se uma variância $S^2 = 169g^2$. Com esse resultado, você diria que a máquina está desregulada com relação à variância?

Passo 1: $H_0 : \sigma^2 = 100$ contra $H_1 : \sigma^2 \neq 100$

Exemplo

Passo 2: $Q = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{15}$.

Passo 3: Sob H_0, temos $Q = \frac{(n-1)S^2}{100} \sim \chi^2_{15}$. Fixando $\alpha = 0,05$, temos que a região de rejeição deve ser tal que

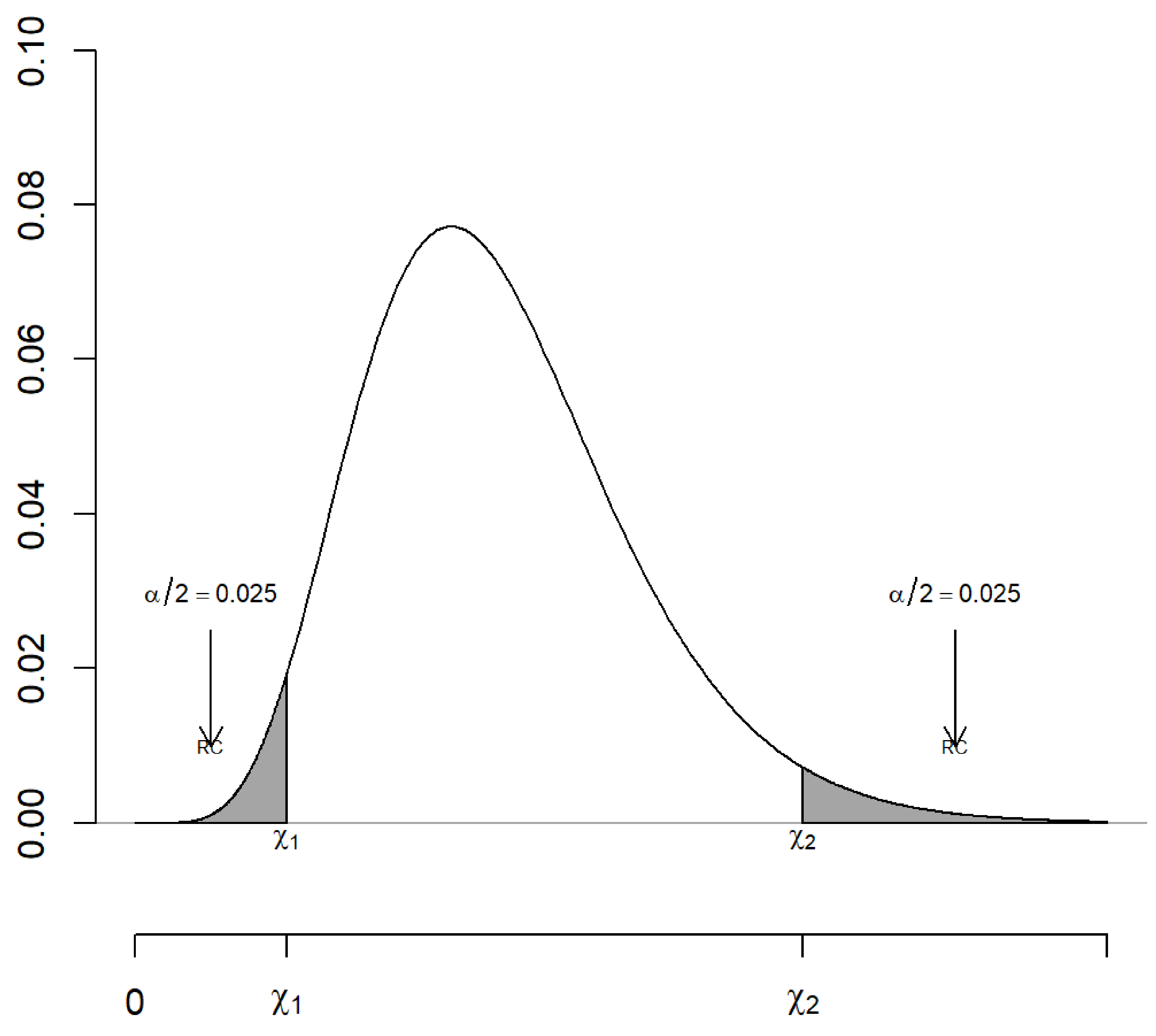
$$RC = \{Q \in \mathbb{R}_+ : \; Q < \chi_1 \text{ ou } Q > \chi_2\},$$

com

$$0,025 = P(Q < \chi_1 | Q \sim \chi^2_{15}) \quad \text{e} \quad 0,025 = P(Q > \chi_2 | Q \sim \chi^2_{15})$$

Exemplo

Graficamente,



Exemplo

Passo 2: $Q = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{15}$.

Passo 3: Sob H_0, temos $Q = \frac{(n-1)S^2}{100} \sim \chi^2_{15}$. Fixando $\alpha = 0,05$, temos que a região de rejeição deve ser tal que

$$RC = \{Q \in \mathbb{R}_+ : \ Q < \chi_1 \text{ ou } Q > \chi_2\},$$

com

$$0,025 = P(Q < \chi_1 | Q \sim \chi^2_{15}) \quad \text{e} \quad 0,025 = P(Q > \chi_2 | Q \sim \chi^2_{15})$$

$$\iff \quad \chi_1 = 6,262 \quad \text{e} \quad \chi_2 = 27,488$$

Portanto, a regra de decisão será: Rejeitar H_0 se $Q < 6,262$ ou $Q > 27,488$.

Passo 4: $q_o = \frac{15 \times 169}{100} = 25,35$

Passo 5: Como o valor observado $q_o = 25,35$ não pertence a RC, então não rejeitamos H_0, ou seja, a máquina está sob controle quanto à variância.

Probabilidade de significância

Probabilidade de significância

- Na construção de um teste de hipóteses, o nível de significância α é fixado.
- Pode-se argumentar que esse procedimento pode levar à rejeição da hipótese nula para um valor de α e à não-rejeição para um valor menor.
- Outra maneira de proceder consiste em apresentar a `probabilidade de significância`, `nível descritivo` ou `p-valor` do teste.
- A principal diferença está em não construir a região crítica.
- O que se faz é indicar a probabilidade de ocorrer valores da estatística mais extremos do que o observado, sob a hipótese de H_0 ser verdadeira.
- Chamamos a estatística de teste de significante quando rejeitamos H_0.
- O p-valor é o menor nível α em que a estatística é significante.
- Pode determinar quão significante os dados são:
  - p-valor pequeno (p-valor $< \alpha$) $\rightarrow$ muito provável que H_0 é falsa.
  - p-valor grande (p-valor $> \alpha$) $\rightarrow$ muito provável que H_0 é verdadeira.

Probabilidade de significância - Teste para a média

Para hipóteses unilaterais, em que $H_0 : \mu = \mu_0$, temos:

$$\text{p-valor} = P(\bar{X} < \bar{X}_{obs} \mid H_0 \text{ verdade}), \quad (H_1 : \mu < \mu_0)$$
$$\text{p-valor} = P(\bar{X} > \bar{X}_{obs} \mid H_0 \text{ verdade}), \quad (H_1 : \mu > \mu_0)$$

Para hipótese bilateral, em que $H_0 : \mu = \mu_0$ e $H_1 : \mu \neq \mu_0$, temos:

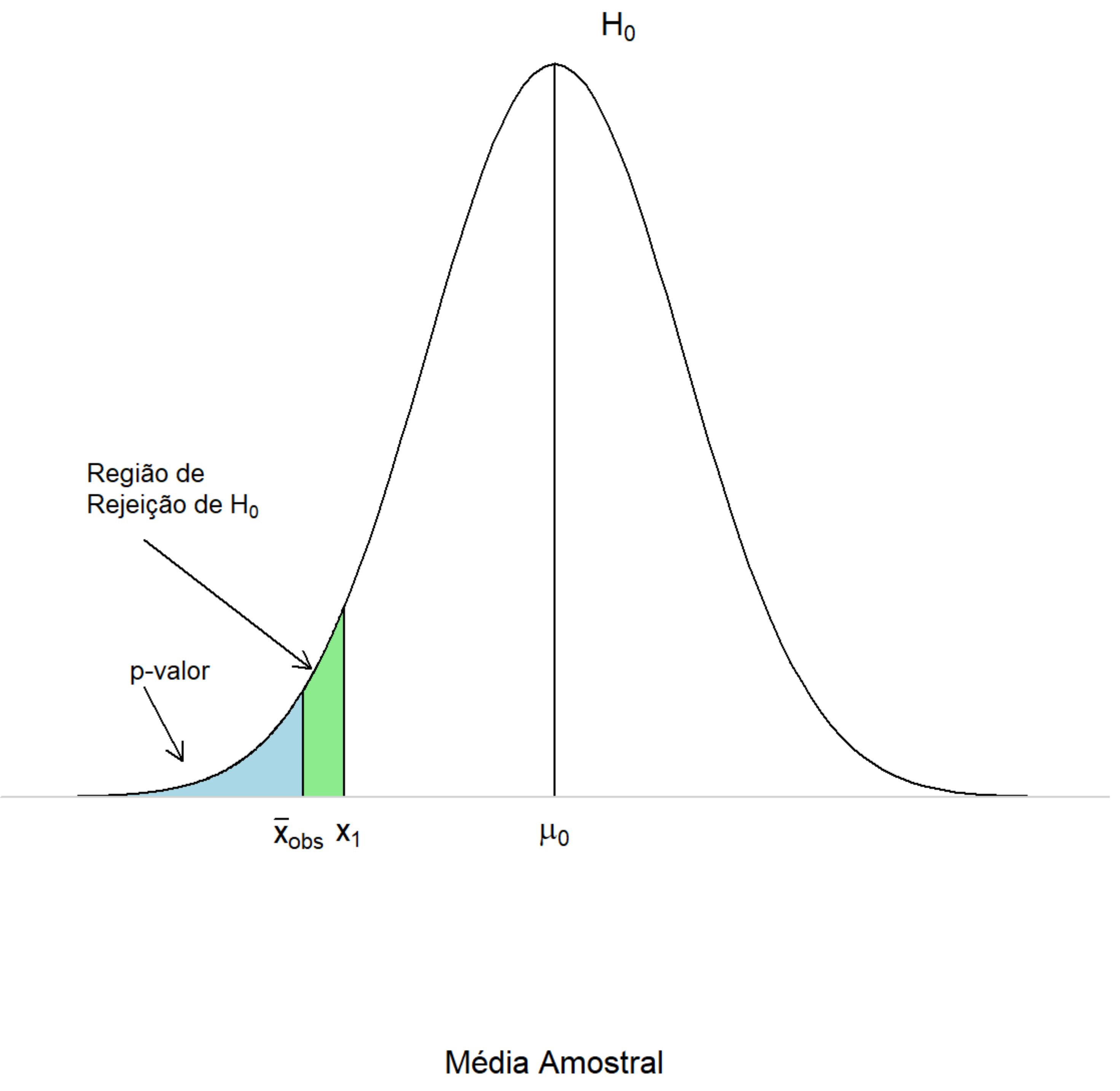
$$\text{p-valor} = 2 * P(\bar{X} < \bar{X}_{obs} \mid H_0 \text{ verdade}),$$

ou

$$\text{p-valor} = 2 * P(\bar{X} > \bar{X}_{obs} \mid H_0 \text{ verdade}).$$

Probabilidade de significância - Teste para a média

Graficamente,



Teste para a proporção - Exemplo

Voltando ao exemplo 1, das anomalias em embriões, vamos obter a probabilidade da proporção amostral ser inferior à observada, supondo $H_0 : p = 0.25$ verdadeira:

A proporção amostral observada $\hat{p}_{obs}$ foi igual a $7/50 = 0.14$, logo,

$$\text{p-valor} = P(\hat{p} < 0.14 \mid \hat{p} \sim N(0.25, 0.0037)) = P\left(Z < \frac{0.14 - 0.25}{\sqrt{0.0037}}\right) = 0.0368.$$

Interpretação: Como o $\text{p-valor} = 0.0368$ é menor que $\alpha = 0.05$, rejeitamos H_0. E dessa forma, temos indicativo de que a proporção de anomalias é menor que 0.25.

Teste para a média com variância conhecida - Exemplo

Voltando ao exemplo 2, o peso de pacotes de café, vamos obter a probabilidade da média amostral ser diferente da observada, supondo $H_0 : \mu = 500g$ verdadeira:

A média amostral observada $\bar{X}_{obs}$ foi igual a $492g$, logo,

$$\text{p-valor} = 2 * P(\bar{X} < 492 \mid \bar{X} \sim N(500, 25)) = 2 * P\left(Z < \frac{492 - 500}{\sqrt{25}}\right) = 0.1151.$$

ou

$$\text{p-valor} = 2 * P(\bar{X} > 492 \mid \bar{X} \sim N(500, 25)) = 2 * P\left(Z > \frac{492 - 500}{\sqrt{25}}\right) = 0.1151.$$

Interpretação: Como o $\text{p-valor} = 0.1151$ é maior que $\alpha = 0.05$, não rejeitamos H_0. E dessa forma, temos indicativo de que a máquina está bem calibrada.

Testes de Hipótese vs Intervalo de Confiança

- Estamos testando,

$$H_0 : \theta = \theta_0 \ vs \ H_1 : \theta \neq \theta_0.$$

- Seja $[l, u]$ um intervalo com confiança $100(1-\alpha)$% para θ.
- Se $\theta_0 \in [l, u]$ não rejeitamos H_0.
- Se $\theta_0 \notin [l, u]$ rejeitamos H_0.

Exemplo:

- No exemplo, da média das notas de novos alunos, o intervalo com 95% de confiança para a média é dada por

$$IC(\mu, 95\%) = \bar{X} \pm t_{19}\frac{S}{\sqrt{n}} = 118 \pm 2.093\frac{20}{\sqrt{20}} = [108.63; 127.37].$$

- A hipótese nula $H_0 : \mu = 115$, com $\mu_0 = 115$, observamos que μ_0 pertence ao intervalo de confiança de 95%.
- logo, não rejeitamos H_0, e dessa forma, temos o indicativo de que a média das provas de uma nova turma seria igual às das anteriores.

Poder de um Teste

Poder de um Teste

$$X \sim f(x|\theta)\ , \qquad \theta \in \Theta$$

$$H_0 : \theta \in \Theta_0 \quad \text{contra} \quad H_1 : \theta \in \Theta_1$$

- Quando $\Theta_0 = \{\theta_0\}$, ou seja, a hipótese H_0 especifica um único valor para o parâmetro, dizemos que H_0 é uma **hipótese simples**, caso contrário, quando a hipótese especifica vários valores para θ, dizemos que a **hipótese** é **composta**.

Definição 3 O **Poder do Teste** com região crítica RC para testar $H_0 : \theta = \theta_0$ contra $H_1 : \theta = \theta_1$ é dado por

$$\pi(\theta_1) = P(\text{Rejeitar } H_0 | H_0 \text{ é falsa}) = P_{H_1}(\boldsymbol{X} \in RC) = P(\boldsymbol{X} \in RC | \theta = \theta_1)$$

Observações:

i. $\pi(\theta_1) = 1 - \beta$, onde $\beta =$ P(Erro II) = P(aceitar $H_0 | H_0$ é falsa$)$;

ii. Quando a hipótese alternativa $\hat{e}$ composta, não podemos calcular o valor de β, consequentemente, não obtemos o poder do teste. Neste caso, podemos construir uma **função poder** para o teste.

Exemplo

Exemplo: Voltando ao Exemplo da máquina de café, a v.a. X representava o peso dos pacotes de café, e assumimos que $X \sim N(\mu, 400)$. Para testar se a máquina estava regulada ($H_0 : \mu = 500$) ou não ($H_1 : \mu \neq 500$), construímos a seguinte $RC = \{\overline{x} < 487,1 \text{ ou } \overline{x} > 512,9\}$. A probabilidade β do erro tipo II não pode ser calculada, pois

$$\beta = P(\text{aceitar } H_0 | H_0 \text{ é falsa})$$

$$= P(487,1 < \overline{X} < 512,9 | \overline{X} \sim N(\mu, 400/16),\ \mu \neq 500)$$

- Observe que o valor de β depende do particular valor que μ pode assumir na hipótese alternativa, ou seja, $\beta = \beta(\mu)$. Por exemplo, se a máquina se desregular para $\mu = 505$, teremos

$$\beta(505) = P(487,1 < \overline{X} < 512,9 | \overline{X} \sim N(505, 400/16))$$

$$= P(-3,58 < Z < 1,58) = 0,9428.$$

Exemplo

- A função $\pi(\mu) = 1 - \beta(\mu) = P$(Rejeitar $H_0 | H_0$ é falsa) é chamada a função poder do teste.
- A Tabela 1 apresenta os valores de $\pi(\mu)$ e $\beta(\mu)$ para diversos valores de μ.

Tabela 1: Valores de $\pi(\mu)$ e $\beta(\mu)$

Valores de μ			
à esquerda de 500	à direita de 500	$\pi(\mu)$ (em %)	$\beta(\mu)$ (em %)
500	500	1,0	99,0
498	502	1,7	98,3
495	505	5,7	94,3
492	508	16,4	83,6
490	510	28,1	71,9
487	513	49,0	51,0
485	515	66,3	34,7
480	520	92,1	7,9

Exemplo

Observe que quanto maior for a distância entre o valor fixado em $H_0 (\mu = 500)$ e o valor atribuído para a hipótese alternativa, maior será a probabilidade de tomar a decisão correta. As seguintes propriedades de $\pi(\mu)$ são facilmente verificadas:

i. $\pi(-\infty) = \pi(+\infty) = 1$;

ii. $\pi(500) = \alpha$

iii. $\pi(\mu)$ decresce para $\mu < 500$ e cresce para $\mu > 500$.

Testes para duas amostra

Testes para duas amostra

Existem uma série de testes de hipóteses, cada um com sua finalidade.

Para duas amostra, temos alguns testes:

- Teste para a comparação de médias com variância conhecida (Teste Z);
- Teste para a comparação de médias com variância desconhecida (Teste t de **Student** comparação de grupos);
- Teste para a comparação de variâncias;
- Teste para comparação de amostras dependentes (Teste t de **Student** comparação de grupos pareado)

Teste para Comparação das médias

Teste para Comparação das médias

- Queremos comparar:

$$H_0 : \mu_1 = \mu_2 \quad \begin{array}{ll} \times & (i) H_1 : \mu_1 \neq \mu_2.\ \text{(bilateral)} \\ \times & (ii) H_1 : \mu_1 > \mu_2.\ \text{(Unilateral à direita)} \\ \times & (iii) H_1 : \mu_1 < \mu_2.\ \text{(Unilateral à esquerda)} \end{array}$$

CASO 1: Mesma variância, conhecida. $(\sigma_1^2 = \sigma_2^2 = \sigma^2)$.

- Estatística para o teste:

$$Z = \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sigma\sqrt{\frac{1}{n} + \frac{1}{m}}} \sim N(0,1).$$

Exemplo

Exemplo: Duas técnicas de venda são aplicadas por dois grupos de vendedores: a técnica A, por 12 vendedores, e a técnica B, por 15 vendedores. Espera-se que a técnica B produza melhores resultados. Suponha que as vendas para as duas técnicas seguem uma distribuição Normal, com média desconhecida e variância comum igual a 50. Com base nos dados apresentados na Tabela 1, teste ao nível de 5% se há diferenças significativas entre as vendas resultantes das duas técnicas.

Tabela 1: Dados para duas técnicas de vendas

Medidas	**Técnica A**	**Técnica B**
Média	68	76
Variância	50	75
Vendedores	12	15

Exemplo

Passo 1: $H_0 : \mu_A = \mu_B$ contra $H_1 : \mu_A < \mu_B$

Passo 2:

$$Z = \frac{\overline{X}_A - \overline{X}_B - (\mu_A - \mu_B)}{\sigma\sqrt{\frac{1}{n} + \frac{1}{m}}} \sim N(0,1)$$

Passo 3: Sob H_0, temos $Z = \frac{\overline{X}_A - \overline{X}_B}{\sqrt{50}\sqrt{\frac{1}{n}+\frac{1}{m}}} \sim N(0,1)$

$$RC = \{Z \in \mathbb{R} : Z < z_c\}, \text{ com } 0,05 = P(Z < z_c | Z \sim N(0,1))$$

Assim, temos $z_c = -1,64$. E a regra de decisão é dada por

$$\text{Rejeitar } H_0 \text{ se } Z < -1,64.$$

Exemplo

Passo 4: O valor observado da estatística é

$$z_0 = \frac{68 - 76}{\sqrt{50}\sqrt{\frac{1}{12} + \frac{1}{15}}} = -2,92$$

Passo 5: Como o valor observado de Z pertence a RC, rejeitamos H_0, e concluímos que há evidências de que a técnica B é melhor que a técnica A.

Importante

- Poderíamos construir um I.C. para a diferença $\theta = \mu_A - \mu_B$!!!
- Em que:

$$\text{IC}(\theta, \gamma) = \left(\overline{X}_A - \overline{X}_B \pm z_{\alpha/2}\sigma\sqrt{\frac{1}{n} + \frac{1}{m}}\right)$$

Teste para Comparação das médias

CASO 2: Mesma variância, desconhecida.

- Neste caso, a estatística para o teste:

$$T = \frac{\overline{X}_A - \overline{X}_B - (\mu_A - \mu_B)}{S_p \sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t_{n+m-2},$$

onde

$$S_p^2 = \frac{(n-1)S_A^2 + (m-1)S_B^2}{n+m-2}.$$

- Este teste é conhecido como `teste t para duas amostras!`

Exemplo

Exemplo: Suponha que no exemplo anterior as variâncias populacionais fossem iguais, mas seu valor comum σ^2 desconhecido. Repita o teste anterior.

Passo 1: $H_0 : \mu_A = \mu_B$ contra $H_1 : \mu_A < \mu_B$

Passo 2:

$$T = \frac{\overline{X}_A - \overline{X}_B - (\mu_A - \mu_B)}{S_p\sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t_{25}$$

Passo 3: Sob H_0, temos $T = \frac{\overline{X}_A - \overline{X}_B}{S_p\sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t_{25}$

$$RC = \{T \in \mathbb{R} : T < t_c\}, \text{ com } 0,05 = P(T < t_c | T \sim t_{25})$$

Assim, temos $t_c = -1,708$. E a regra de decisão é dada por: rejeitar H_0 se $T < -1,708$.

Exemplo

Passo 4:

$$S_p^2 = \frac{11 \times 50 + 14 \times 75}{25} = 64$$

O valor observado da estatística é

$$t_0 = \frac{68 - 76}{\sqrt{64}\sqrt{\frac{1}{12} + \frac{1}{15}}} = -2,56$$

Passo 5: Como o valor observado de T pertence a RC, rejeitamos H_0, e concluímos que há evidências de que a técnica B é melhor que a técnica A.

Teste para Comparação das médias

CASO 3: Variâncias desiguais e desconhecidas.

- Pode-se provar que a estatística

$$T = \frac{\overline{X}_A - \overline{X}_B - (\mu_A - \mu_B)}{\sqrt{\frac{S_A^2}{n} + \frac{S_B^2}{m}}},$$

sob H_0, tem uma distribuição aproximadamente t-Student com graus de liberdade, dados aproximadamente por

$$v = \frac{(x+y)^2}{\frac{x^2}{n-1} + \frac{y^2}{m-1}},$$

com $x = \frac{S_A^2}{n}$ e $y = \frac{S_B^2}{m}$.

- Este teste é conhecido como `Problema de Behrens-Fisher!`.

Exemplo

Exemplo: Queremos testar as resistências de dois tipos de vigas de aço, A e B. Tomando-se $n = 15$ vigas do tipo A e $m = 20$ vigas do tipo B, obtemos os valores na Tabela 2. Admita que as variâncias da resistência para os dois tipos de viga não podem ser consideradas iguais. Compare as resistências médias dos dois tipos de viga ao nível de 5%.

Tabela 2: Dados para os dois tipos de vigas de aço

Tipo	Média	Variância
A	70,5	81,6
B	84,3	161,5

Passo 1: $H_0 : \mu_A = \mu_B$ contra $H_1 : \mu_A \neq \mu_B$

Passo 2:

$$T = \frac{\overline{X}_A - \overline{X}_B - (\mu_A - \mu_B)}{\sqrt{\frac{S_A^2}{n} + \frac{S_B^2}{m}}}$$

Exemplo

Passo 3: Sob H_0, temos $T = \frac{\overline{X}_A - \overline{X}_B}{\sqrt{\frac{S_A^2}{n} + \frac{S_B^2}{m}}} \sim t_v$

$$RC = \{T \in \mathbb{R} : T < t_1 \text{ ou } T > t_2\},$$

com

$$0,025 = P(T < t_1 | T \sim t_v) \quad \text{e} \quad 0,025 = P(T > t_2 | T \sim t_v).$$

Agora,

$$v = \frac{((81,6/15) + (161,5/20))^2}{(81,6/15)^2/14 + (161,5/20)^2/19} = 32,9 \simeq 33.$$

Portanto,

$$t_1 = -2,0348 \quad \text{e} \quad t_2 = 2,0348.$$

E a regra de decisão é dada por: Rejeitar H_0 se $T < -2,0348$ ou $T > 2,0348$.

Exemplo

Passo 4: O valor observado da estatística é

$$t_0 = \frac{70,5 - 84,3}{\sqrt{\frac{81{,}6}{15} + \frac{161{,}5}{20}}} = -3,75.$$

Passo 5: Como o valor observado de T pertence a RC, rejeitamos H_0, e concluímos que há evidências de que os dois tipos de vigas têm resistências médias diferentes.