



INFERÊNCIA ESTATÍSTICA: TESTE DE HIPÓTESES

2011

Teste de hipóteses

Exemplo. Uma indústria adquire de um certo fabricante pinos cuja resistência média à ruptura é especificada em 60 unid. (**valor nominal** da especificação). Em um determinado dia a indústria recebeu um grande lote de pinos e a equipe técnica da indústria deseja verificar se o lote atende às especificações.

H_0 : O lote atende às especificações (Hipótese **nula**).

H_1 : O lote não atende às especificações (Hipótese **alternativa**).

A v. a. X (resistência à ruptura) é tal que $X \sim N(\mu, 25)$. O problema pode ser resolvido testando as hipóteses

H_0 : $\mu = 60$ (hipótese **simples**: um único valor)
e H_1 : $\mu \neq 60$ (hipótese **composta**: mais de um valor)

Teste de hipóteses

Uma **hipótese estatística** é uma afirmação sobre o(s) parâmetro(s) da distribuição de probabilidade de uma característica (v. a. X) da população.

Um **teste** de uma hipótese estatística é um procedimento ou **regra de decisão** que nos possibilita decidir por H_0 ou H_1 com base na amostra X_1, \dots, X_n .

Exemplo. A equipe técnica da indústria decidiu retirar uma amostra aleatória de tamanho $n = 16$ do lote recebido. A resistência de cada pino foi medida e foi calculada a resistência média \bar{x} (**estimador** de μ), que será utilizada para realizar o teste (**estatística de teste**). Podemos afirmar que

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

Obs. Se X_1, X_2, \dots, X_n é uma amostra de uma distribuição $N(\mu, \sigma^2)$, então a **média amostral** tem distribuição $N(\mu, \sigma^2/n)$.

Para quais valores de \bar{x} a equipe técnica deve rejeitar H_0 e portanto rejeitar o lote?

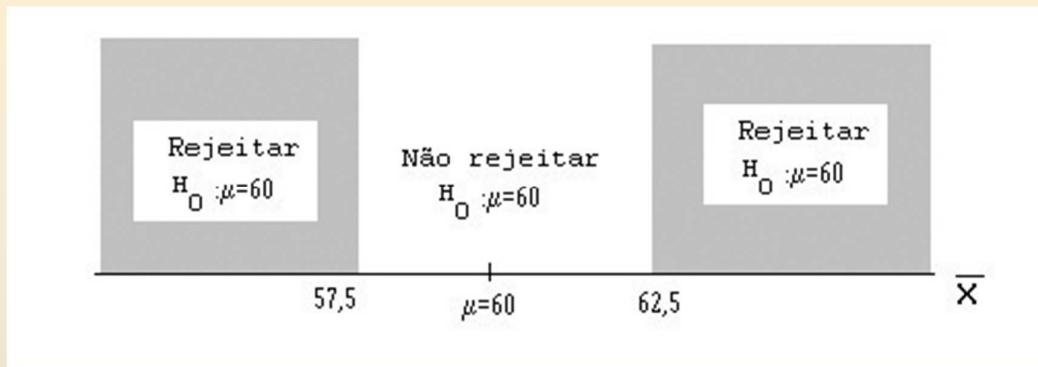
Região crítica (R_c) ou região **de rejeição** é o conjunto de valores assumidos pela estatística de teste para os quais a hipótese nula é **rejeitada**. Seu complementar é a **região de aceitação** (R_a).

Exemplo. Se o lote está fora de especificação, isto é, se $H_1: \mu \neq 60$ for verdadeira, espera-se que a média amostral seja inferior ou superior a 60 unid.

A equipe técnica decidiu adotar a seguinte regra:
rejeitar H_0 se \bar{x} for maior do que 62,5 unid. ou menor do que 57,5 unid.
As duas regiões são

$$R_c = \{\bar{X} > 62,5 \text{ ou } \bar{X} < 57,5\} : \text{região de } \mathbf{rejeição} \text{ de } H_0 \text{ e}$$

$$R_a = \{57,5 \leq \bar{X} \leq 62,5\} : \text{região de } \mathbf{aceitação} \text{ de } H_0.$$



Procedimento (teste):

Se $\bar{x} \in R_c$, rejeita - se H_0 ;

Se $\bar{x} \notin R_c$, não se rejeita (aceita - se) H_0 .

Tipos de erros

Erro tipo I: **rejeitar** H_0 quando H_0 é **verdadeira**.

Erro tipo II: **não rejeitar** (**aceitar**) H_0 quando H_0 é **falsa**.

Exemplo. As hipóteses são

H_0 : O lote atende às especificações;

H_1 : O lote não atende às especificações.

Erro tipo I: **rejeitar** o lote sendo que ele está **de acordo** com as especificações.

Erro tipo II: **não rejeitar** (**aceitar**) o lote sendo que ele **não está de acordo** com as especificações.

Quadro resumo:

Decisão	Situação real e desconhecida	
	Ho verdadeira	Ho falsa
Não rejeitar Ho	Decisão correta	Erro tipo II
Rejeitar Ho	Erro tipo I	Decisão correta

Nível de significância e poder

$P(\text{Erro tipo I}) = \alpha$ (nível de significância).

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

$$\begin{aligned} P(\text{Erro tipo II}) &= \beta = P(\text{Não rejeitar } H_0; H_0 \text{ falsa}) \\ &= P(\text{Não rejeitar } H_0; H_1 \text{ verdadeira}). \end{aligned}$$

$$1 - \beta = P(\text{Rejeitar } H_0; H_0 \text{ é falsa}) : \text{poder do teste.}$$

Obs. Quanto maior o poder, melhor o teste.

Exemplo. As hipóteses são $H_0: \mu = 60$ e $H_1: \mu \neq 60$. Logo,

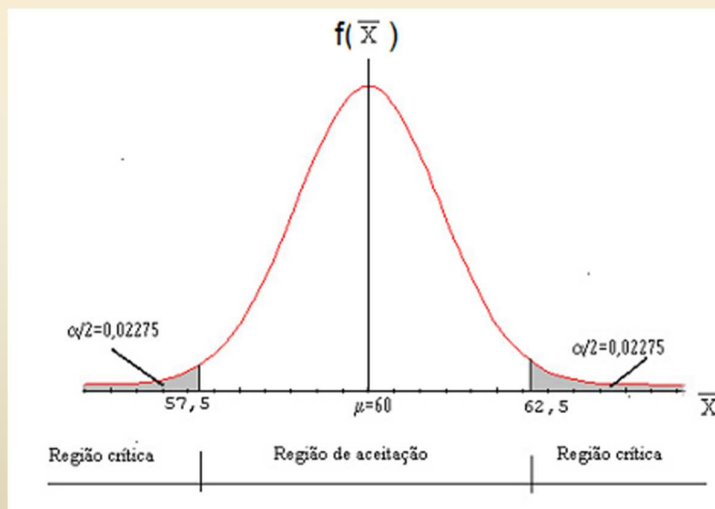
$$\alpha = P(\bar{X} > 62,5 \text{ ou } \bar{X} < 57,5; H_0: \mu = 60).$$

Se H_0 for verdadeira, então $\bar{X} \sim N(60, 25/16)$.

Calculamos o nível de significância:

$$\begin{aligned} \alpha &= P(\bar{X} > 62,5; H_0: \mu = 60) + P(\bar{X} < 57,5; H_0: \mu = 60) \\ &= P\left(\frac{\bar{X} - 60}{\sqrt{25/16}} > \frac{62,5 - 60}{\sqrt{25/16}}\right) + P\left(\frac{\bar{X} - 60}{\sqrt{25/16}} < \frac{57,5 - 60}{\sqrt{25/16}}\right) \\ &= P(Z > 2,00) + P(Z < -2,00) = 0,02275 + 0,02275 = 0,0455. \end{aligned}$$

Cálculo de α :

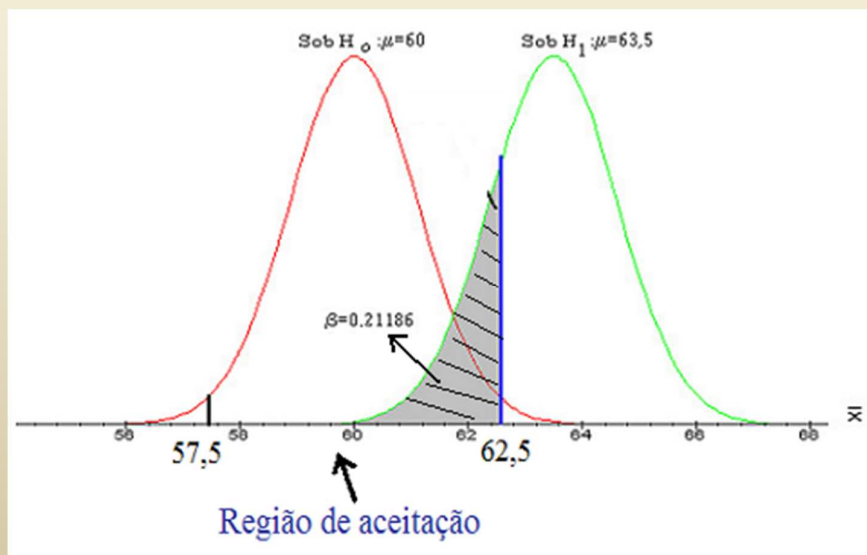


Cálculo de β :

$\beta = P(\text{Não rejeitar } H_0; H_1 \text{ verdadeira}) = P(57,5 \leq \bar{X} \leq 62,5; H_1 : \mu \neq 60)$.

Como exemplo de cálculo de β , selecionamos $H_1: \mu = 63,5$. Logo,

$\bar{X} \sim N\left(63,5; \frac{25}{16}\right)$ e $\beta = P(57,5 \leq \bar{X} \leq 62,5; H_1 : \mu = 63,5)$.



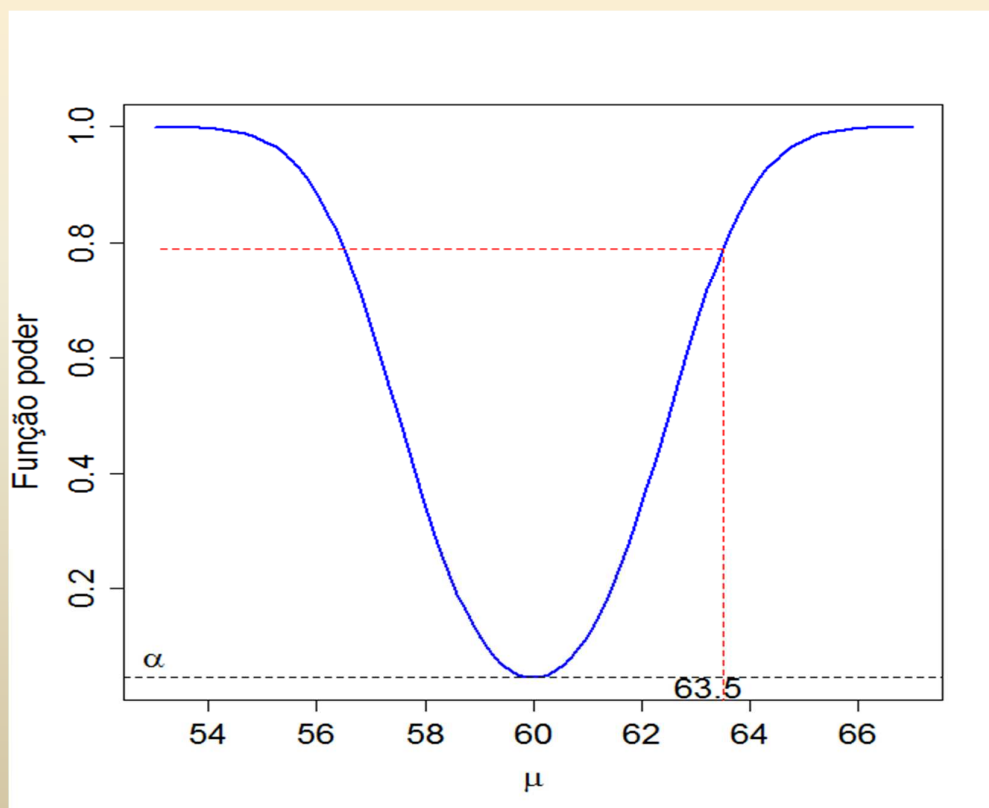
Cálculo de β :

Efetuando o cálculo obtemos

$$\begin{aligned}\beta &= P(57,5 \leq \bar{X} \leq 62,5; H_1 : \mu = 63,5) \\ &= P(\bar{X} \leq 62,5; \mu = 63,5) - P(\bar{X} \leq 57,5; \mu = 63,5) \\ &= P(Z \leq -0,80) - P(Z \leq -4,80) \\ &= 0,2119 - 0,0000 \\ &= 0,2119 .\end{aligned}$$

Logo, se $\mu = 63,5$, o poder do teste é igual a $1 - 0,2119 = 0,7881$.

Função poder



Hipóteses bilateral e unilaterais

Se as hipóteses nula e alternativa são

$$H_0 : \mu = \mu_0;$$

$$H_1 : \mu \neq \mu_0,$$

em que μ_0 é uma constante conhecida (**valor de teste**), o teste é chamado de **bilateral**.

Podemos ter também as hipóteses

$$H_0 : \mu = \mu_0;$$

$$H_1 : \mu < \mu_0, \quad \text{unilateral à esquerda}$$

ou $H_0 : \mu = \mu_0;$

$$H_1 : \mu > \mu_0. \quad \text{unilateral à direita}$$

Sugestão. Expressar H_0 em forma de igualdade.

Exemplo

Um fabricante de um certo componente afirma que o **tempo médio** de vida dos componentes produzidos é de **1000 horas**. Engenheiros de produto têm interesse em verificar se uma **modificação do processo** de fabricação **umenta** a duração dos componentes.

Hipóteses:

$$H_0 : \mu = 1000 \text{ horas};$$

$$H_1 : \mu > 1000 \text{ horas},$$

sendo μ o tempo médio de duração dos componentes.

Procedimento básico de testes de hipóteses

O procedimento de teste de hipóteses relativo ao parâmetro θ de uma população é decomposto em **quatro** passos:

(i) Formulação das **hipóteses**:

$$H_0 : \theta = \theta_0;$$

$$H_1 : \theta < \theta_0 \text{ ou } \theta > \theta_0 \text{ ou } \theta \neq \theta_0.$$

(ii) Identificação da **estatística de teste** e caracterização da sua **distribuição** (por exemplo, método de substituição, lâmina 6).

(iii) **Escolha do nível de significância** do teste ($\alpha = 5\%$, 1% e $0,5\%$ são comuns) e obtenção da **região crítica**.

(iv) Cálculo da estatística de teste e **tomada de decisão** (H_0 deve ser rejeitada ou não?).

Teste de hipóteses para uma média populacional

Considere uma amostra aleatória de tamanho n de uma população normal com média μ (desconhecida) e variância σ^2 (**conhecida**). Iniciamos pelo teste unilateral à esquerda:

(i)

$$H_0 : \mu = \mu_0;$$

$$H_1 : \mu < \mu_0.$$

(ii) A **estatística de teste** é a média amostral \bar{X} (**estimador pontual** de μ). Se a distribuição da população é normal ou se amostra é grande ($n \geq 30$, mesmo que a distribuição da população não seja normal) a distribuição de \bar{X} é $N(\mu, \sigma^2/n)$, aproximadamente. Se H_0 for verdadeira, então

$$Z = \frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma} \sim N(0,1).$$

Teste de hipóteses para uma média populacional

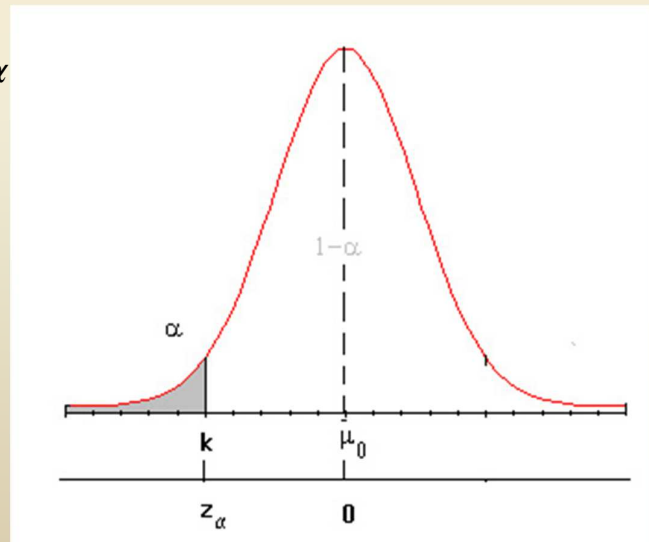
(iii) Rejeitamos H_0 em favor de H_1 se a média amostral \bar{X} é “pequena” em relação μ_0 . A região crítica é obtida selecionando um k tal que $R_c = \{ \bar{X} < k \}$, sendo que $P(\bar{X} < k; H_0: \mu = \mu_0) = \alpha$. Ou seja, sob H_0

$$P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} < \frac{k - \mu_0}{\sigma/\sqrt{n}}\right) = P\left(Z < \frac{k - \mu_0}{\sigma/\sqrt{n}}\right) = \alpha$$

$$\Rightarrow \frac{k - \mu_0}{\sigma/\sqrt{n}} = z_\alpha \Rightarrow k = \mu_0 + z_\alpha \times \frac{\sigma}{\sqrt{n}}$$

$$\Rightarrow R_c = \left\{ \bar{X} < \mu_0 + z_\alpha \times \frac{\sigma}{\sqrt{n}} \right\}.$$

Obs. $z_\alpha < 0$.



(iv) Conclusão: se $\bar{x} \in R_c = \left\{ \bar{X} < \mu_0 + z_\alpha \times \frac{\sigma}{\sqrt{n}} \right\}$, rejeita-se H_0 ; caso contrário não se rejeita H_0 .

IC para a média e testes de hipóteses

O teste da hipótese $H_0: \mu = \mu_0$ contra $H_1: \mu \neq \mu_0$ a um nível de significância α pode ser efetuado utilizando um IC com coeficiente de confiança igual a $1 - \alpha$.

Construímos o IC de $100(1-\alpha)\%$ para μ , dado por

$$[L; U] = [\bar{X} - E; \bar{X} + E], \text{ em que } E = z_{\alpha/2} \frac{\sigma}{\sqrt{n}} : \sigma \text{ conhecido.}$$

$$\text{ou } E = t_{\alpha/2, n-1} \frac{s}{\sqrt{n}} : \sigma \text{ desconhecido.}$$

Se $\mu_0 \notin \text{IC}$, rejeitamos H_0 ; caso contrário, não rejeitamos H_0 .

Exemplo

Um comprador de tijolos suspeita de uma **diminuição** na resistência. De experiências anteriores, sabe-se que a resistência **média** ao desmoronamento de tais tijolos é igual a **200 kg**, com um **desvio padrão** de **10 kg**. Uma amostra de **100 tijolos**, escolhidos ao acaso, forneceu uma **média de 195 kg**. A um nível de significância de 5%, pode-se afirmar que a resistência média ao desmoronamento diminuiu?

(i) As hipóteses de interesse são

$$H_0 : \mu = 200 \text{ kg};$$

$$H_1 : \mu < 200 \text{ kg}.$$

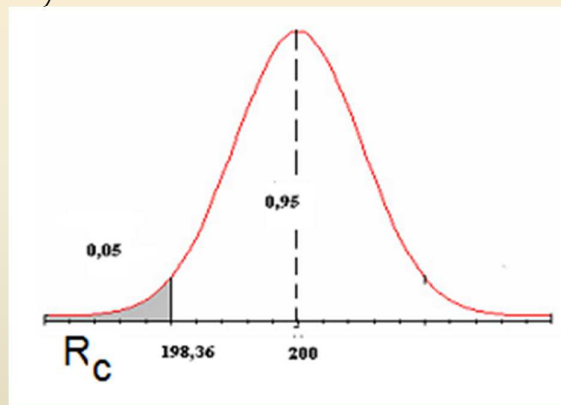
(ii) A estatística de teste é a média amostral \bar{X} . Já que $n = 100 \geq 30$, tem-se que sob H_0 , $\bar{X} \sim N\left(200, \frac{100}{100}\right)$, aproximadamente.

(iii) A região crítica pode ser obtida selecionando k de maneira que $R_c = \{ \bar{X} < k \}$, sendo que $P(\bar{X} < k, H_0: \mu = \mu_0) = \alpha = 0,05$. Ou seja, sob H_0 ,

Exemplo

$$P\left(\frac{\bar{X} - 200}{10/\sqrt{100}} \leq \frac{k - 200}{10/\sqrt{100}}\right) = P\left(Z < \frac{k - 200}{1}\right) = \alpha = 0,05 \Rightarrow k - 200 = -1,64 \Rightarrow k = 198,36$$

$$\Rightarrow R_c = \{ \bar{X} < 198,36 \}.$$



(iv) Do enunciado a média amostral vale 195. Logo, $\bar{x} = 195 \in R_c = \{ \bar{X} < 198,36 \}$. Rejeita-se H_0 a um nível de 5% de significância.

Conclusão. De acordo com os dados coletados e adotando um nível de significância de 5%, concluímos que resistência média ao desmoronamento diminuiu.

Método alternativo

Um método alternativo **prático**: trabalhar diretamente na **escala Z**.

(i) $H_0 : \mu = \mu_0$ contra $H_1 : \mu < \mu_0$.

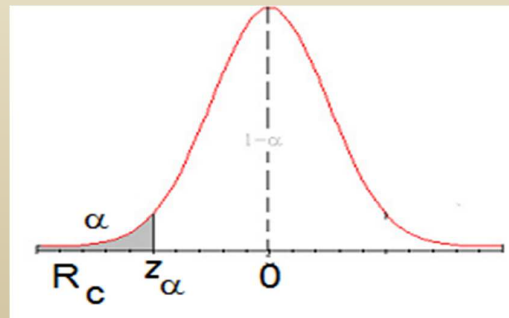
(ii) Estatística de teste:

$$Z = \frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma} \underset{\text{sob } H_0}{\sim} N(0,1), \text{ pelo menos aproximadamente.}$$

(iii) Região crítica para um nível de significância α escolhido:

$$R_c = \{Z < z_\alpha\}.$$

(iv) Se $z \in R_c = \{Z < z_\alpha\}$, rejeita-se H_0 ; caso contrário, não se rejeita H_0 .



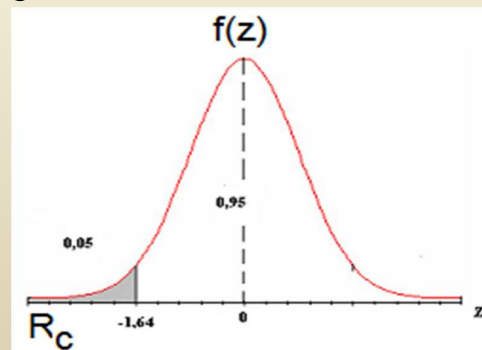
Exemplo

(i) $H_0 : \mu = 200$ contra $H_1 : \mu < 200$.

(ii) Estatística de teste: $Z = \frac{\sqrt{n}(\bar{X} - 200)}{\sigma} \underset{\text{sob } H_0}{\sim} N(0,1).$

(iii) Região crítica para um nível de significância $\alpha = 0,05$:

$$R_c = \{z < -1,64\}.$$



(iv) Calculamos $z = \frac{\sqrt{100}(195 - 200)}{10} = -5 \in R_c$. Rejeita-se H_0 a um nível de significância de 5%.

Procedimento geral

Hipóteses:

(i)

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

$$H_1 : \mu < \mu_0 \quad H_1 : \mu > \mu_0 \quad H_1 : \mu \neq \mu_0$$

À esquerda

À direita

Bilateral

(ii) Estatística de teste:

(a) Variância da população é conhecida:

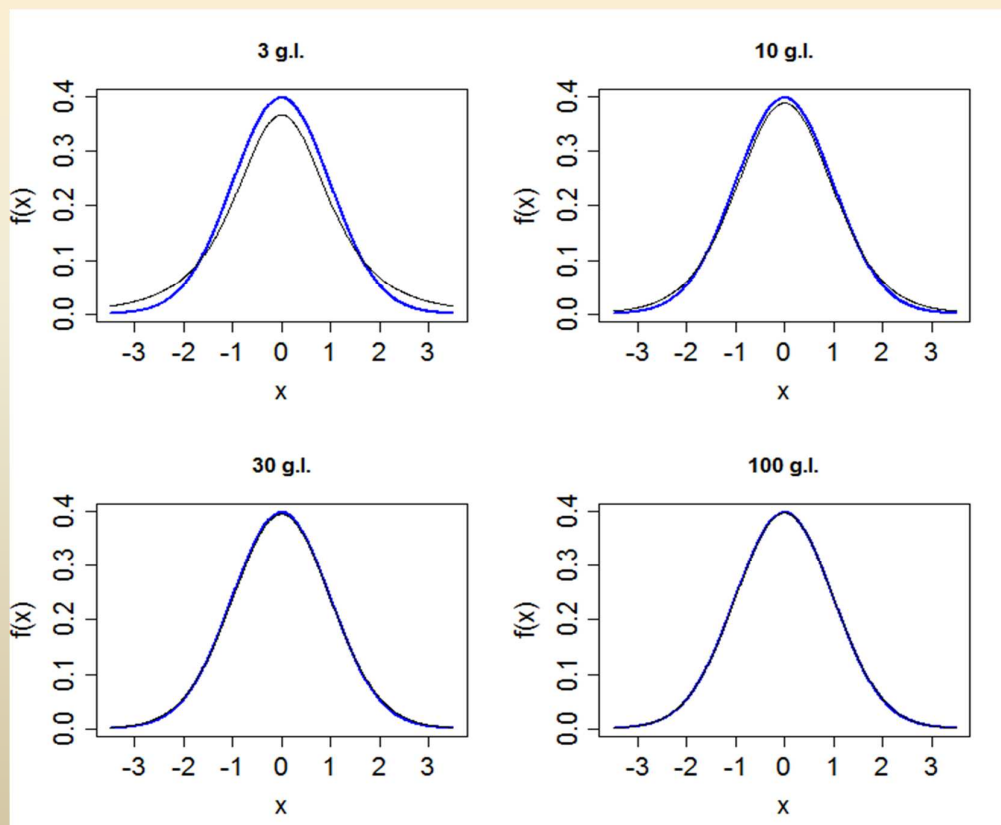
$$Z = \frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma} \underset{\text{sob } H_0}{\sim} N(0,1).$$

(b) Variância da população é desconhecida (s é o desvio padrão amostral):

$$T = \frac{\sqrt{n}(\bar{X} - \mu_0)}{s} \underset{\text{sob } H_0}{\sim} t(n-1).$$

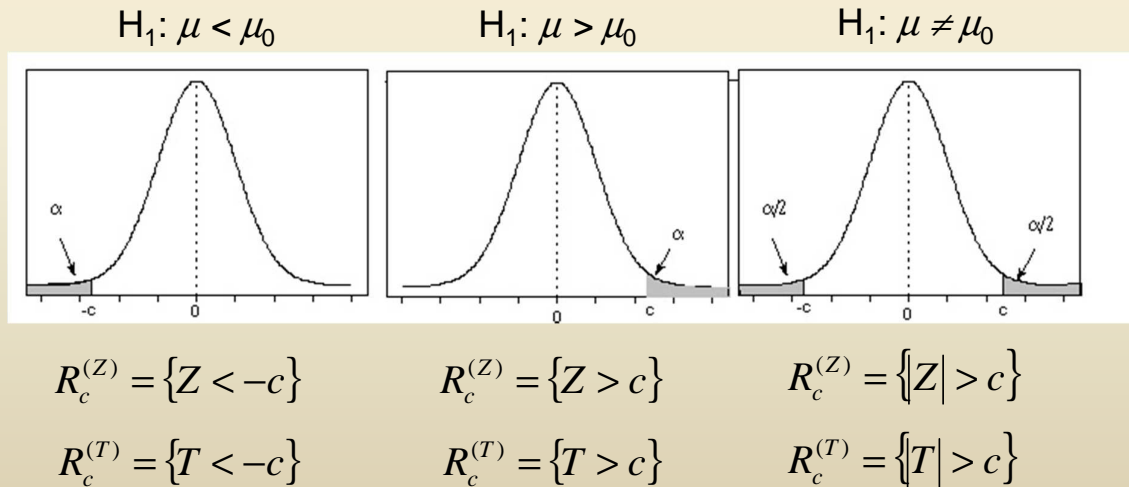
Distribuição t de Student com $n - 1$ graus de liberdade (g.l.).

Distribuições normal e t de Student



Procedimento geral

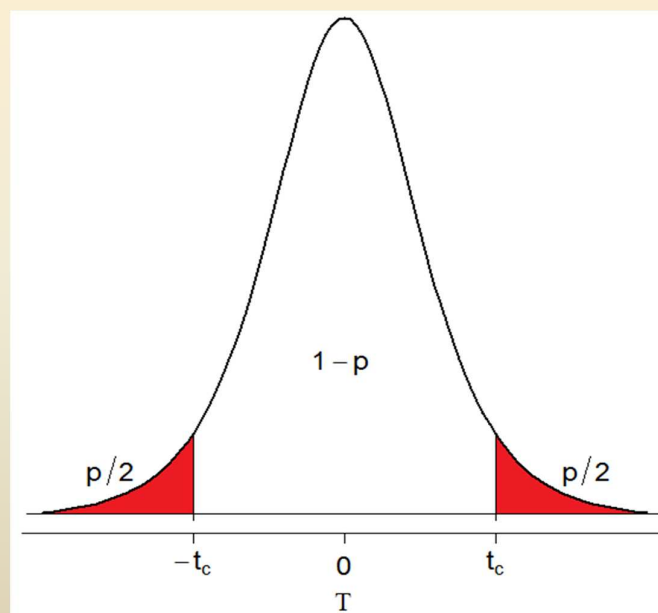
(iii) Região crítica para um nível de significância α escolhido:



(iv) Se $Z \in R_C$ ou $T \in R_C$, rejeita-se H_0 ; caso contrário, não se rejeita H_0 .

Obs. Nas regiões críticas com Z e T o valor de c **não é** o mesmo.

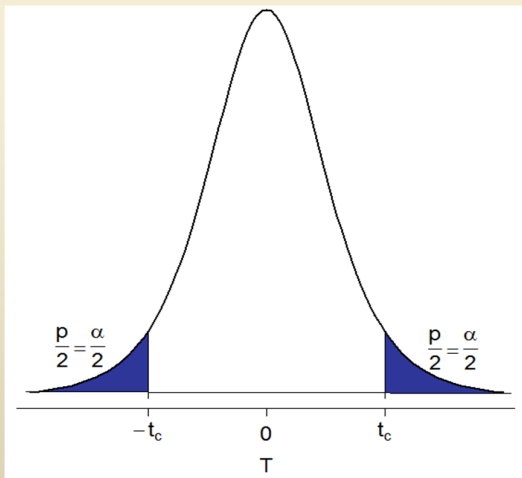
Tabela da distribuição t de Student



A tabela (**Tábua III**) contém os valores de t_c ($t_c > 0$) tais que $P(-t_c \leq T \leq t_c) = 1 - p$ correspondentes a **alguns valores de p** e para **alguns graus de liberdade**.

Tabela da distribuição t de Student

Exemplo. Se $n = 12$, são 11 graus de liberdade. Se tivermos $H_1: \mu \neq \mu_0$, escolhendo $\alpha = 5\%$, temos $p/2 = \alpha/2$, ou seja, $p = 5\%$.



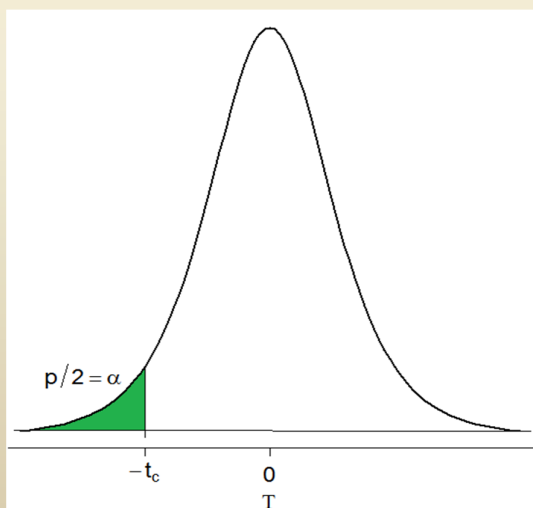
Consultando a tábua III encontramos $t_c = 2,201$ e $R_c = \{|T| > 2,201\}$.

Graus de liberdade	p = 90%	80%	...	5%	...	0,10%
1						
2						
...						
11				2,201		
...						
120						
Infinito				1,960		
	p = 90%	80%	...	5%	...	0,10%

Obs. À medida que aumentam os graus de liberdade, a distribuição t se aproxima da normal (neste exemplo, $t_c \rightarrow 1,960 = z_c$).

Tabela da distribuição t de Student

Exemplo. Se $n = 28$, são 27 graus de liberdade. Se tivermos $H_1: \mu < \mu_0$, escolhendo $\alpha = 1\%$, temos $p/2 = \alpha$, ou seja, $p = 2\alpha = 2\%$.



Consultando a tábua III encontramos $t_c = 2,473$ e $R_c = \{T < -2,473\}$.

Graus de liberdade	p = 90%	80%	...	2%	...	0,10%
1						
2						
...						
27				2,473		
...						
120						
Infinito				2,326		
	p = 90%	80%	...	2%	...	0,10%

Obs. Neste exemplo, se tivéssemos $H_1: \mu > \mu_0$, a região crítica seria $R_c = \{T > 2,473\}$.

Exemplo

Dados **históricos** coletados em uma linha de produção de um certo item indicam **115 kg** como massa **média**. A fim de testar a hipótese de que a **média** de itens recentemente produzidos se **manteve**, retirou-se, ao acaso, uma **amostra** de 20 itens, obtendo-se média igual a 118 kg e desvio padrão 20 kg. Utilize $\alpha = 0,05$.

(i) As hipóteses de interesse são

$$H_0 : \mu = 115 \text{ kg};$$

$$H_1 : \mu \neq 115 \text{ kg}.$$

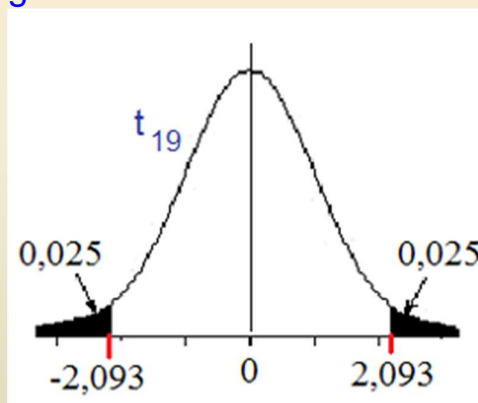
Aproximamos a distribuição da média das 20 notas por uma distribuição normal com média μ e variância σ^2 / n .

(ii) Estatística de teste:

$$T = \frac{\sqrt{n}(\bar{X} - 115)}{S} \underset{\text{sob } H_0}{\sim} t(n-1).$$

Exemplo

(iii) Região crítica para um nível de significância $\alpha = 0,05$ e com $n - 1 = 19$ g.l.:



$$R_c = \{ |T| > 2,093 \}$$

(iv) Calculamos $T = \frac{\sqrt{20}(118 - 115)}{20} = 0,67 \notin R_c$. Não se rejeita H_0 a um nível de significância de 5%. A **diferença não é significativa**.

Conclusão. De acordo com os dados coletados, a um nível de significância de 5% concluímos que a massa média dos itens produzidos se manteve.

Teste de hipóteses para uma proporção populacional

O procedimento para testes de hipóteses sobre a proporção populacional (p) **semelhante** ao utilizado para testes sobre uma média populacional.

Problema. Testar a hipótese que a proporção de sucessos de um ensaio de Bernoulli é igual a um valor especificado p_0 . Isto é, testar um dos seguintes pares de hipóteses:

(i)

$$\begin{array}{ccc} H_0 : p = p_0 & H_0 : p = p_0 & H_0 : p = p_0 \\ \underbrace{H_1 : p < p_0}_{\text{À esquerda}} & \underbrace{H_1 : p > p_0}_{\text{À direita}} & \underbrace{H_1 : p \neq p_0}_{\text{Bilateral}} \end{array}$$

Teste de hipóteses para uma proporção populacional

(ii) Estatística de teste:

$$Z = \frac{\sqrt{n}(\bar{p} - p_0)}{\sqrt{p_0(1 - p_0)}} \underset{\text{sob } H_0}{\sim} N(0,1), \text{ aproximadamente,}$$

sendo que

$$\bar{p} = \frac{\text{Número de sucessos}}{n} = \frac{\sum_{i=1}^n X_i}{n} : \text{estimador pontual de } p.$$

é a **proporção amostral de sucessos** e $X_i = 1$, se o resultado for sucesso; $X_i = 0$, se o resultado for insucesso.

Exemplo

Um estudo é realizado para determinar a presença de pequenas anomalias em chapas metálicas de uma certa dimensão. Segundo o fabricante, a **proporção** de chapas com anomalias é **inferior a 25%**. Foram inspecionadas **50** chapas escolhidas ao acaso e sete delas apresentaram algum tipo de anomalia. Estes dados justificam a afirmação do fabricante? Adote um nível de significância igual a 0,05.

(i) Hipóteses :

$$H_0 : p = 0,25 ;$$

$$H_1 : p < 0,25 .$$

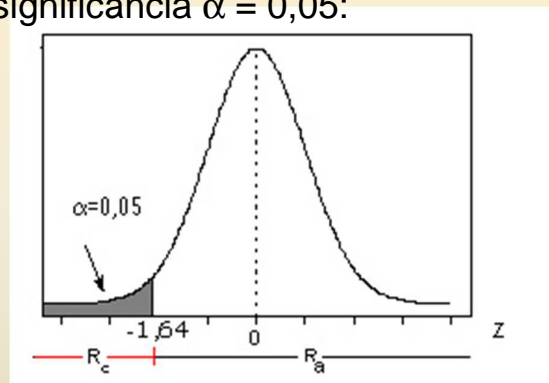
(ii) Estatística de teste:

$$Z = \frac{\sqrt{50}(\bar{p} - 0,25)}{\sqrt{0,25(1 - 0,25)}} \underset{\text{sob } H_0}{\sim} N(0,1), \text{ aproximadamente.}$$

Exemplo

(iii) Região crítica para um nível de significância $\alpha = 0,05$:

$$R_c = \{z < -1,64\}.$$



(iv) Temos $n = 50$. Calculamos $\bar{p} = \frac{7}{50} = 0,14$ e $z = \frac{\sqrt{50}(0,14 - 0,25)}{\sqrt{0,25 \times (1 - 0,25)}} = -1,796 \in R_c$.

Rejeita-se H_0 ao nível de 5% de significância.

Conclusão. Adotando um nível de significância de 5% concluímos a partir dos dados que a proporção de chapas produzidas com anomalias é inferior a 25%.