

# Introdução à Inferência Estatística

Mário Olímpio de Menezes

# Inferência estatística

Na **inferência estatística** os dois principais objetivos são:

- **Estimar** um parâmetro populacional
  - Estimativa pontual
  - Estimativa intervalar
- **Testar** uma hipótese ou afirmativa sobre um parâmetro populacional

# Estimação

Se  $\theta$  é um parâmetro da distribuição de uma v.a.  $X$  e  $X_1, \dots, X_n$  é uma *amostra* desta distribuição, temos as duas abordagens de estimativa:

## 1. Estimação pontual

- Apresentar **um valor** para  $\theta$  que é uma função da amostra  $X_1, \dots, X_n$  (cálculo de  $\theta$ ), chamada de *estimador* de  $\theta$
- Espera-se que o estimador tenha boas propriedades:
  - i. **em média** esteja próximo de  $\theta$
  - ii. o estimador **se aproxima de**  $\theta$  quando **n** aumenta, ...

# Estimação

## 2. Estimação intervalar

- Apresentar **um intervalo de possíveis valores** para  $\theta$ , chamado de *intervalo de confiança*. Os limites do intervalo são funções da amostra  $X_1, \dots, X_n$  (são aleatórios).

$\theta ?$

Limite inferior =  
 $L(X_1, \dots, X_n)$

Limite superior =  
 $U(X_1, \dots, X_n)$

- A probabilidade de que o intervalo contenha  $\theta$  deve ser **alta**
- A *amplitude* do intervalo de ser tão **pequena** quanto possível (intervalo mais *preciso*)

# Testes de hipótese

## Hipótese

É uma afirmativa sobre uma propriedade da população

## Teste de hipótese

- É um procedimento para se testar uma afirmativa sobre uma propriedade da população
- Permite tomar **decisões** sobre a população com base em informações de dados amostrais.

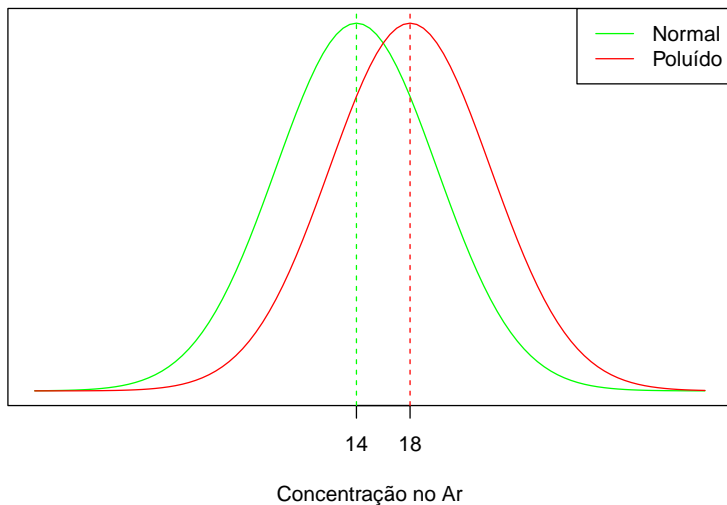
# Exemplo 1

Suponha que, no ar não poluído, a concentração de certa substância se comporta segundo um modelo Normal com média 14 unidades/ml e desvio padrão 6 unidades/ml.

Quando um determinado poluente está presente em excesso, esta substância tem concentração média alterada para 18 unidades/ml.

Admitimos que o modelo Normal com desvio padrão 6 unidades/ml, continua representado de forma adequada a concentração da substância no ar poluído.

# Exemplo 1



# Exemplo 1

Para averiguar se um tratamento é eficaz contra este tipo de poluição, selecionamos uma amostra de 30 volumes de ar **submetidos ao tratamento**.

Assumimos que todos os elementos da amostra  $X_1, \dots, X_{30}$  possuem a mesma distribuição:  $X_i \sim N(\mu, 36)$ , onde:

- $\mu = 14$  se o tratamento for eficiente
- $\mu = 18$  se o tratamento não for eficiente

Se a média da amostra for próxima de 14, temos **evidências** de que o tratamento é eficaz. Se for mais próxima de 18, as **evidências** são contrárias ao tratamento.

Então a pergunta é: o quão próximo é “**próximo**”?



## Exemplo 2

Deseja-se estudar a tolerância de um equipamento eletrônico ao número de impactos termo-elétricos.

Pelas características de fabricação do equipamento, é possível admitir que a probabilidade de falha seja constante, isto é, após cada impacto, existe uma probabilidade  $p$  que ele falhe.

Representando por  $X$  a variável número de impactos anteriores à falha, pretende-se verificar se o modelo Geométrico com  $p = 0.4$  é adequado para caracterizar essa variável.

# Variância conhecida

## Exemplo 1 - continuação

- Interesse geral  $\mu = 14$ ?
- Distribuição da média amostral para  $n = 30$ :  $N(\mu, 36/30)$ .
- Critério para decidir sobre o valor de  $\mu$ .
- **Valor crítico**, digamos  $x_c$  tal que se a média amostral ( $\bar{x}_{obs}$ ) for maior que  $x_c$  concluímos que a amostra pertence ao ar com média 18.
- Como  $\bar{X}$  é uma variável aleatória, devem existir erros associados.

# Tipos de hipóteses

## Hipótese nula $H_0$

É uma afirmativa de que o valor de um parâmetro populacional é **igual** a algum valor especificado. (O termo *nula* é usado para indicar nenhuma mudança, nenhum efeito).

- No ex. 1 temos:  $\mu = 18$  unidades/ml (tratamento não teve efeito)
- No ex. 2 temos:  $p = 0,4$

## Hipótese alternativa $H_1$ ou $H_a$

É uma afirmativa de que o parâmetro tem um valor, que, de alguma forma, difere da hipótese nula. Ex.:

- $p \neq 0,4$        $p < 0,4$        $p > 0,4$

# Tipos de hipóteses

Quando fazemos um teste de hipótese, chegamos a um dos dois possíveis resultados:

- **Rejeitar**  $H_0$ : em favor da hipótese alternativa  $H_1$
- **Não rejeitar**  $H_0$ : e conclui-se que não existem diferenças

## Atenção!

- Mesmo quando **aceitamos** a hipótese nula, não significa dizer que temos 100% de certeza da decisão.
- Sempre existe uma probabilidade de **erro** associado a todo teste de hipótese

# Hipóteses

## Hipótese simples:

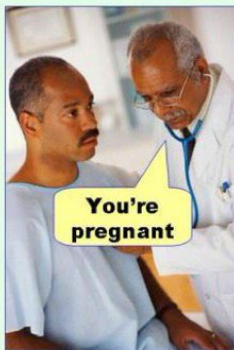
- $H_0$  : O tratamento não é eficaz ( $\mu = 18$ )
- $H_1$  : O tratamento é eficaz ( $\mu = 14$ )

## Hipóteses compostas:

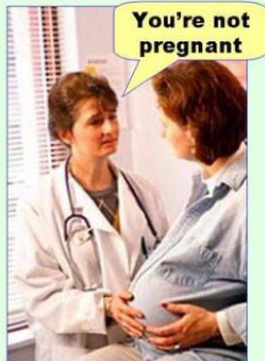
- Hipótese unilateral à esquerda
  - $H_0$  : O tratamento não é eficaz ( $\mu = 18$ );
  - $H_1$  : O tratamento é eficaz ( $\mu < 18$ ).
- Hipótese bilateral:
  - $H_0$  : O tratamento não é eficaz ( $\mu = 18$ );
  - $H_1$  : O tratamento é eficaz ( $\mu \neq 18$ ).

# Erros ao realizar um teste de hipótese

**Type I error**  
(false positive)



**Type II error**  
(false negative)



# Erros ao realizar um teste de hipótese

- **Erro Tipo I:** rejeitar  $H_0$ , quando  $H_0$  é verdadeira.
- **Erro Tipo II:** não rejeitar  $H_0$  quando  $H_0$  é falsa.

|                    | $H_0$ verdadeira | $H_0$ falsa     |
|--------------------|------------------|-----------------|
| Não rejeitar $H_0$ | Decisão correta  | Erro tipo II    |
| Rejeitar $H_0$     | Erro tipo I      | Decisão correta |

# Erros ao realizar um teste de hipótese

Definimos por  $\alpha$  e  $\beta$  as probabilidades de cometer os erros do tipo I e II:

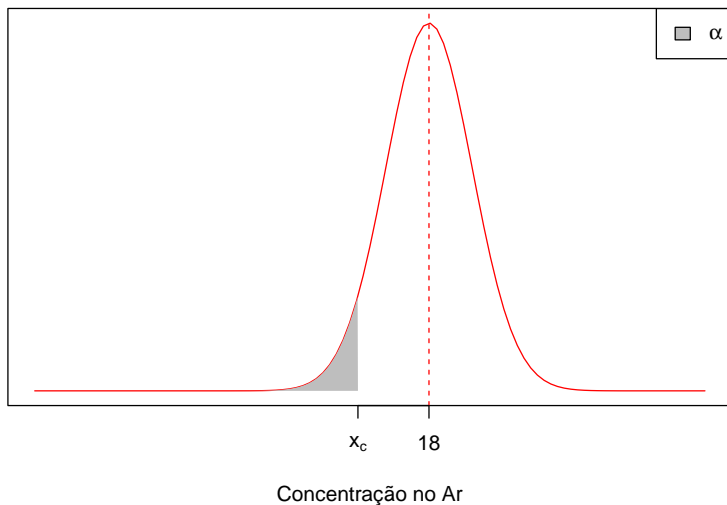
- $\alpha = P(\text{erro tipo I}) = P(\text{rejeitar } H_0 \mid H_0 \text{ verdadeira})$
- $\beta = P(\text{erro tipo II}) = P(\text{não rejeitar } H_0 \mid H_0 \text{ falsa})$

No exemplo 1, se  $H_0 : \mu = 18$  e  $H_a : \mu < 18$ , então:

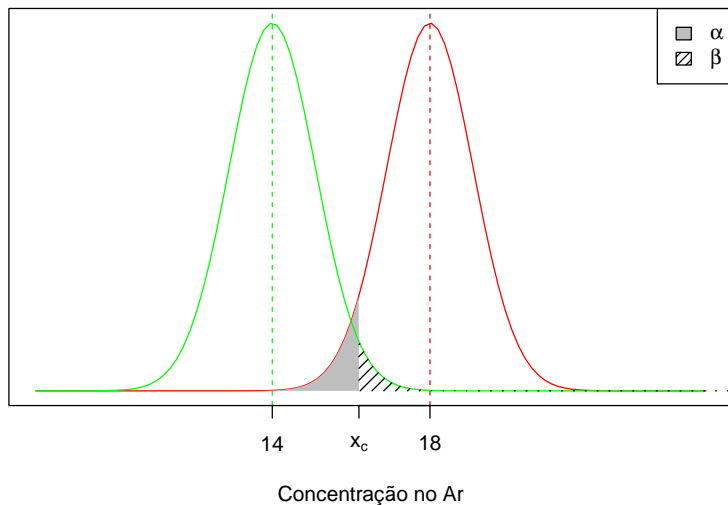
- $\alpha = P(\text{concluir que o tratamento é eficaz quando na verdade não é})$
- $\beta = P(\text{concluir que o tratamento não é eficaz quando na verdade é})$



# Erros ao realizar um teste de hipótese



# Erros ao realizar um teste de hipótese



# Erros ao realizar um teste de hipótese

A situação ideal é aquela em que ambas as probabilidades,  $\alpha$  e  $\beta$ , são próximas de zero.

No entanto, à medida que diminuimos  $\alpha$ , a probabilidade  $\beta$  tende a aumentar.

Levando isso em conta, ao formular as hipóteses, **devemos cuidar para que o erro mais importante a ser evitado seja o erro do tipo I.**

Por isso, a probabilidade  $\alpha$  recebe o nome de **nível de significância** do teste, e é esse erro que devemos controlar.

## Valor crítico

Supondo  $\alpha$  conhecido podemos determinar o valor crítico  $x_c$ .

$$\begin{aligned}\alpha &= P(\text{erro tipo I}) = P(\text{rejeitar } H_0 \mid H_0 \text{ verdadeira}) \\ &= P(\bar{X} < x_c \mid \mu = 18) = P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < \frac{x_c - 18}{6/\sqrt{30}}\right) \\ &= P(Z < z_c)\end{aligned}$$

com  $Z \sim N(0, 1)$ .

## Obtendo o valor crítico

Dado  $\alpha$  encontramos  $z_c$  na tabela normal padrão.

Obtemos  $x_c$

$$z_c = \frac{x_c - 18}{6/\sqrt{30}} \Rightarrow x_c = 18 + z_c \frac{6}{\sqrt{30}}$$

Supondo  $\alpha = 0.05$  temos

$$0.05 = P(Z < z_c) \Rightarrow z_c = -1.64$$

logo

$$x_c = 18 - 1.64 \frac{6}{\sqrt{30}} = 16.2$$

## Região Crítica

Dada uma amostra, se  $\bar{x}_{obs} < 16.2$ , **rejeitamos**  $H_0$ , concluindo que o tratamento é eficaz.

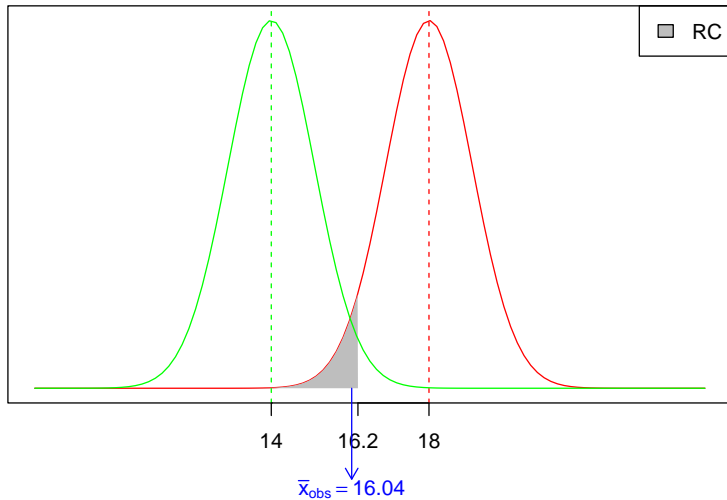
O conjunto dos números reais menores que 16.2 é denominado de **Região de Rejeição** ou **Região Crítica** (RC), isto é:

$$RC = \{x \in \mathbb{R} : x < 16.2\}.$$

**No exemplo 1**, se a média amostral dos 30 volumes foi  $\bar{x}_{obs} = 16.04$ , então **rejeitamos**  $H_0$ , ao nível de significância  $\alpha = 0.05$ .

Nesse caso,  $\bar{x}_{obs} < x_c$  está dentro da RC.

# Região Crítica



# Teste de hipótese bilateral

Definindo as hipóteses

$$H_0 : \mu = \mu_0 \quad \text{e} \quad H_a : \mu \neq \mu_0$$

A Região Crítica será dada por

$$RC = \{x \in \mathbb{R} \mid x < x_{c_1} \quad \text{ou} \quad x > x_{c_2}\}$$

Para um valor de  $\alpha$  fixado, determinamos  $x_{c_1}$  e  $x_{c_2}$  de modo que

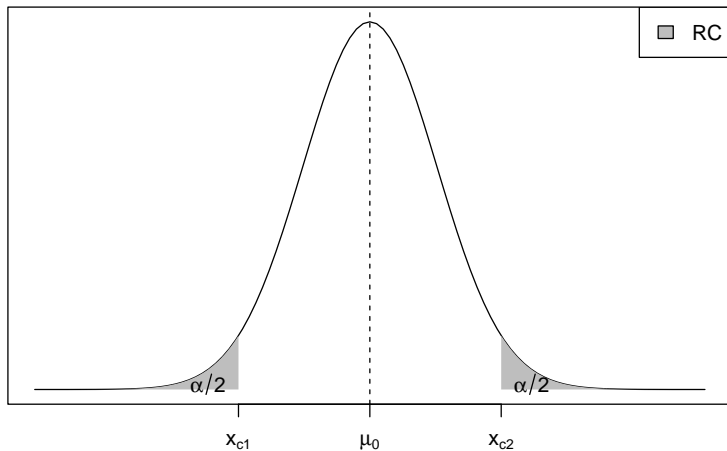
$$P(\bar{X} < x_{c_1} \cup \bar{X} > x_{c_2}) = \alpha$$

Assim, distribuimos a área  $\alpha$  igualmente entre as duas partes da RC

$$P(\bar{X} < x_{c_1}) = \frac{\alpha}{2} \quad \text{e} \quad P(\bar{X} > x_{c_2}) = \frac{\alpha}{2}$$



# Teste de hipótese bilateral



# Etapas de um teste de hipótese

- 1 Estabelecer as hipóteses 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  $\alpha$  e obter a região crítica.
  1. Em algumas situações pode ser o inverso: Fixamos a região crítica (limites) e calculamos o  $\alpha$
- 5 Concluir o teste com base na estimativa e na região crítica.

## 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  $\mu$ ), 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?

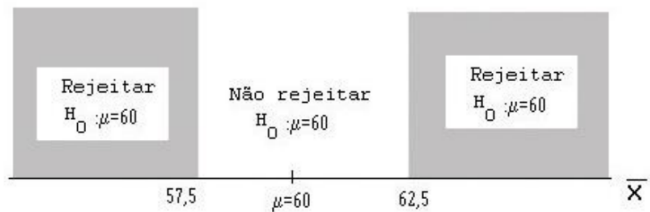
# Teste de Hipóteses

- **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.
  - neste caso, a *região crítica* foi estabelecida (através de limites); depois calculamos o  $\alpha$  correspondente.
- As duas regiões são:

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

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

# Teste de Hipóteses



## 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.

# Nível de Significância e Poder

- $P(\text{Erro tipo I}) = \alpha$  (**nível de significância**)
  - $\alpha = P(\text{Rejeitar } H_0; H_0 \text{ verdadeira})$
- $P(\text{Erro tipo II}) = \beta = P(\text{Não rejeitar } H_0; H_0 \text{ falsa})$ 
  - $= P(\text{Não rejeitar } H_0; H_1 \text{ verdadeira})$
  - $1 - \beta = P(\text{Rejeitar } H_0; H_0 \text{ é falsa})$  : **poder do teste.**

**Obs.:** Quanto maior o poder, melhor o teste.



## Nível de Significância e Poder

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  $\bar{X} \sim N(60, 25/16)$

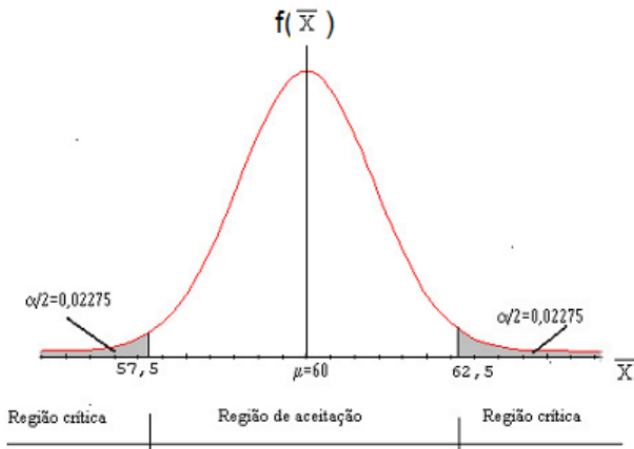
Calculamos o nível de significância:

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

```
alpha = 1 - pnorm(2) + 1 - pnorm(-2)
alpha
[1] 0.04550026
```

# Nível de Significância e Poder

Cálculo de  $\alpha$ :



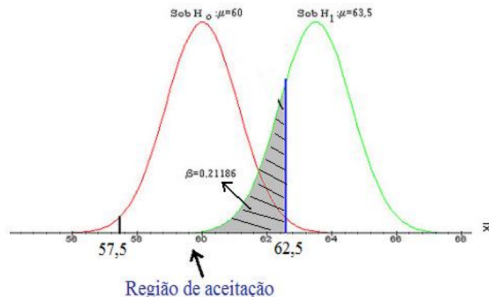
# Nível de Significância e Poder

Cálculo de  $\beta$

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

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

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



# Nível de Significância e Poder

Cálculo de  $\beta$ ; efetuando o cálculo no R:

```
n = 16
sigma = sqrt(25)
sem = sigma/sqrt(n)
# alpha calculado anteriormente
mu0 = 60
I = c(alpha/2, 1 - alpha/2)
q = qnorm(I, mean = mu0, sd = sem)
q
[1] 57.5 62.5
mu = 63.5
p = pnorm(q, mean = mu, sd = sem)
p
[1] 7.933282e-07 2.118554e-01
diff(p)
[1] 0.2118546
poder = 1 - diff(p)
poder
[1] 0.7881454
```

Se a média verdadeira for 63.5, a probabilidade que nós rejeitemos a hipótese nula é de aproximadamente 78.81%

# Hipóteses bilateral e unilaterais

Se as hipóteses nula e alterantiva são:

- $H_0 : \mu = \mu_0;$
- $H_1 : \mu \neq \mu_0;$

em que  $\mu_0$  é uma constante conhecida (**valor de teste**), o teste é chamado de **bilateral**

Podemos ter também as hipóteses:

- $H_0 : \mu = \mu_0;$
- $H_1 : \mu < \mu_0. \leftarrow \textit{unilateral à esquerda}$

ou

- $H_0 : \mu = \mu_0;$
- $H_1 : \mu > \mu_0. \leftarrow \textit{unilateral à direita}.$

É interessante expressar  $H_0$  em forma de igualdade, ou seja, fazer teste bilateral.

# Hipóteses bilateral e unilateral

## 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 *aumenta* a duração dos componentes.

Hipóteses:

- $H_0 : \mu = 1000$  horas
- $H_1 : \mu > 1000$  horas

sendo  $\mu$  o tempo médio de duração dos componentes

# Procedimento básico de testes de hipóteses

O procedimento de teste de hipóteses relativo ao parâmetro  $\theta$  de uma população é decomposto em quatro passos:

## (i) Formulação das hipóteses

- $H_0 : \theta = \theta_0$ ;
- $H_1 : \theta < \theta_0$  ou  $\theta > \theta_0$  ou  $\theta \neq \theta_0$ .

## (ii) Identificação da estatística de teste

e caracterização da sua **distribuição** (por exemplo, método de substituição)

# Procedimento básico de testes de hipóteses

## (iii) Escolha do nível de significância do teste

$\alpha = 0.05, 0.01, 0.005$  são valores 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  $\mu$  (desconhecida) e variância  $\sigma^2$  (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  $\mu$ ). Se a distribuição da população é normal ou se a amostra é grande ( $n \geq 30$ ), mesmo que a distribuição da população não seja normal) a distribuição de  $\bar{X} \sim N(\mu, \sigma^2/n)$ . 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)

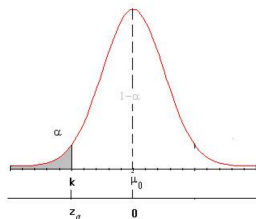
Rejeitamos  $H_0$  em favor de  $H_1$  se a média amostral  $\bar{X}$  é “pequena” em relação a  $\mu_0$ . A região crítica é obtida selecionando-se um  $k$  tal que  $R_c = \{\bar{X} < k\}$ , sendo que  $P(\bar{X} < k; H_0 : \mu = \mu_0) = \alpha$ ,

Ou seja, sob  $H_0$

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

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

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



**Obs.:**  $z_\alpha < 0$ .

# Teste de hipóteses para uma média populacional

## (iv) Conclusão:

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

# Teste de hipóteses para uma média populacional

## 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:

- $H_0 : \mu = 200 \text{ kg}$
- $H_1 : \mu < 200 \text{ kg}$

## Exemplo – cont.

### (ii) Estatística de teste

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

### (iii) Região Crítica

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  $H_0$ :

$$P\left(\frac{\bar{X} - 200}{10/\sqrt{(100)}} \leq \frac{k - 200}{10/\sqrt{(100)}}\right) = P\left(Z < \frac{k - 200}{1}\right) = \alpha = 0.05$$

Precisamos encontrar um  $k$  que satisfaça a relação  $P(Z < (k - 200)) = 0.05$   
O valor de  $Z$  é obtido facilmente:  $Z = \frac{195 - 200}{10/\sqrt{(100)}} = -5$

## Exemplo – cont.

### (iii) Região Crítica – cont.

Nosso teste é unilateral ( $H_1 : \mu < 200$  kg), então precisamos achar o valor de  $P$  que divide a curva de probabilidade em 0.05 e 0.95.

$$Z + 200 < k$$

$$-5 + 200 < k$$

$$195 < k$$

$$P(k) = 0.05$$

```
# qual valor divide a área da gaussiana
k <- qnorm(0.05, mean = 200)
k
[1] 198.3551
```

$$\Rightarrow R_c = \{\bar{X} < 198.3551\}$$

## Exemplo – cont.

### (iv) Conclusão

Do enunciado, a média amostral vale 195. Logo,  
 $\bar{x} = 195 \in R_c = \{\bar{X} < 198.3551\}$ .

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

# Método alternativo

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

## Exemplo

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

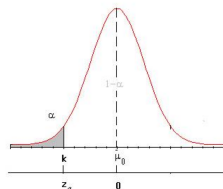
ii. Estatística de teste:

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

iii. Região crítica para um nível de significância  $\alpha$  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

```
xbar = 195
mu0 = 200
sigma = 10
n = 100
z = (xbar - mu0)/(sigma/sqrt(n))
z
[1] -5
alpha = 0.05
# este é o valor que divide a curva em 0.05 e 0.95
rc = qnorm(alpha)
rc
[1] -1.644854
```

A estatística de teste  $z = -5$  está fora do intervalo da região de aceitação (ou seja, está na região crítica),  $-5 \in R_c$ ,  $-5 < 1.645$ . Portanto, a um nível de significância de 5%, nós rejeitamos a hipótese nula que a média seja igual a 200.

# Teste de hipóteses para uma média populacional

## Exercício

Suponha que o peso médio dos Pinguins Reis encontrados em uma colônia na Antártica no último ano foi 15.4 kg. Em uma amostra de 35 pinguins na mesma época neste ano, na mesma colônia, o peso médio foi 14.6kg. Assuma que o desvio padrão da população seja 2.5kg. A um nível de significância de 5%, nós podemos rejeitar a hipótese nula de que o peso médio dos pinguins não difere do último ano?

1. As hipóteses são:

- $H_0 : \mu = 15.4 \text{ kg}$
- $H_1 : \mu \neq 15.4 \text{ kg}$

# Teste de hipóteses para uma média populacional

## Solução

```
xbar = 14.6 # media da amostra
mu0 = 15.4 # valor da hipótese
sigma = 2.5 # desvio padrão da população
n = 35 # tamanho da amostra
z = (xbar - mu0)/(sigma/sqrt(n)) # estatística de teste
z
[1] -1.893146
```

Agora os valores críticos em um nível de significância de 0.05 (bilateral)

```
alpha = 0.05
z.half.alpha = qnorm(1 - alpha/2)
c(-z.half.alpha, z.half.alpha)
[1] -1.959964 1.959964
```

A estatística de teste  $z = -1.8931$  está entre os valores críticos  $-1.96$ ,  $1.96$ , (está dentro da região de aceitação). Portanto, a um nível de significância de 0.05, nós **não** rejeitamos a hipótese nula de que o peso médio dos pinguins não difere do último ano.

# Teste de hipóteses para uma média populacional

## Solução – alternativa

Ao invés de se utilizar o valor crítico  $z_c$ , nós aplicamos a função `pnorm` para calcular o **p-value** da estatística de teste. A comparação é com o nível de significância  $\alpha$  definido para o teste, no nosso caso 5%. Como nosso teste é bilateral, a área (probabilidade) em cada lado é de  $\alpha/2 = 0.025$ .

Temos duas alternativas:

- considerar o valor “dobrado” do `pnorm(z)` calculado, para comparar com o nível de significância  $\alpha = 0.05$ , ou
- considerar o valor obtido no `pnorm(z)` e comparar com  $\alpha/2 = 0.025$

Se o **p-value** obtido for maior do que o nível de significância, nós **não** rejeitamos a hipótese nula de que  $\mu = 15.4$

Considerando o valor “dobrado” do **p-value**:

```
pval = 2 * pnorm(z) # cauda inferior dobrada
pval
[1] 0.05833852
```

Como o valor `pval = 0.0583` é maior do que o nível de significância  $\alpha$ , nós **não** rejeitamos a hipótese nula de que a média seja 15.4

# Procedimento Geral de Teste de Hipótese

Hipóteses:

(i)

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

$$\underbrace{H_1 : \mu < \mu_0}_{\text{À esquerda}} \quad \underbrace{H_1 : \mu > \mu_0}_{\text{À direita}} \quad \underbrace{H_1 : \mu \neq \mu_0}_{\text{Bilateral}}$$

(ii) Estatística de teste:

(a) Variância da população é conhecida:

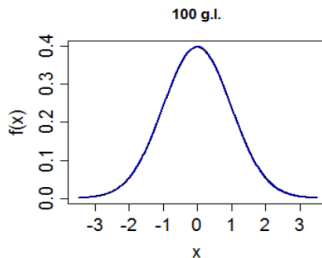
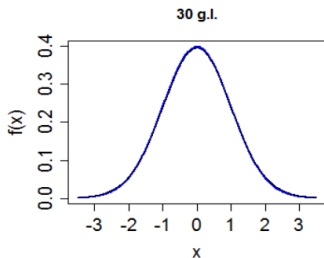
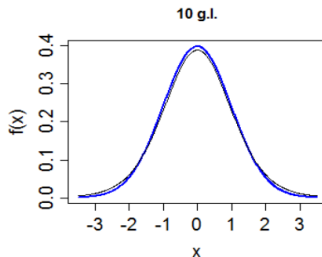
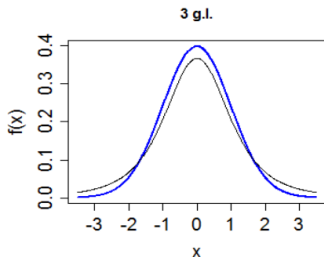
$$Z = \frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma} \underset{sob H_0}{\sim} N(0,1).$$

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

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

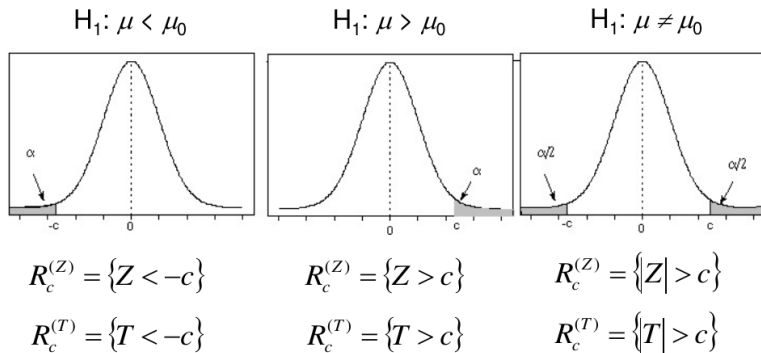
Distribuição t de Student com  
n - 1 graus de liberdade (g.l.).

# Procedimento Geral de Teste de Hipótese



# Procedimento Geral de Teste de Hipótese

(iii) Região crítica para um nível de significância  $\alpha$  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.

# Procedimento Geral de Teste de Hipótese

## Exemplo

Se  $n = 12$ , são 11 graus de liberdade. Se tivermos  $H_1 : \mu \neq \mu_0$  escolhendo  $\alpha = 0.05$ , temos  $p/2 = \alpha/2$ , ou seja,  $p = 0.05$ .

```
alpha = 0.05
df = 11
tc = qt(alpha/2, df = df)
tc
[1] -2.200985
```

## Exemplo

Se  $n = 28$ , são 27 graus de liberdade. Se tivermos  $H_1 : \mu < \mu_0$  escolhendo  $\alpha = 0.01$ , temos  $p/2 = \alpha$ , ou seja,  $p = 2\alpha = 0.02$ .

```
alpha = 0.02
df = 27
tc = qt(alpha/2, df = df)
tc
[1] -2.47266
```



## Exercício

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:

- $H_0 : \mu = 115 \text{ kg};$
- $H_1 : \mu \neq 115 \text{ kg}.$

Utilize uma aproximação da distribuição das médias das 20 notas por uma distribuição normal, com média  $\mu$  e variância  $\sigma^2/n$

### (ii) Estatística de Teste:

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

## Exercício – solução

```
alpha = 0.05
df = 19
n = 20 # tamanho da amostra
S = 20 # desvio padrão da amostra
mu0 = 115
xbar = 118
tc <- qt(alpha/2, df = df)
tc
[1] -2.093024
Tt = (sqrt(n) * (xbar - mu0))/S
Tt
[1] 0.6708204
```

O valor da estatística de teste  $T = 0.6708$  está fora da Região Crítica, ou seja,  $|T| < 2.093$ . Portanto, **não rejeitamos** a  $H_0$  a um nível de significância de 5%. A diferença não é significativa.

# Teste de hipóteses para uma proporção populacional

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

## Problema

Testar a hipótese de 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)

| À esquerda      | À direita       | Bilateral          |
|-----------------|-----------------|--------------------|
| $H_0 : p = p_0$ | $H_0 : p = p_0$ | $H_0 : p = p_0$    |
| $H_1 : p < p_0$ | $H_1 : p > p_0$ | $H_1 : p \neq p_0$ |

# Teste de hipóteses para uma proporção populacional

(ii) Estatística de teste:

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

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

$$H_0 : p = 0.25 \text{ contra } H_1 : p < 0.25$$

### (ii) Estatística de teste:

$$Z = \frac{\sqrt{50}(\bar{p} - 0.25)}{\sqrt{0.25(1 - 0.25)}} \underset{\text{sob } \widetilde{H} - 0}{\approx} N(0, 1), \text{ aproximadamente}$$

# Exemplo

iii Região Crítica para um nível de significância  $\alpha = 0.05$ :

```
alpha = 0.05
rc_z <- qnorm(alpha)
rc_z
[1] -1.644854
```

iv Temos  $n = 50$ . Calculamos:

```
n = 50
ne = 7
pbar = ne/n
pbar
[1] 0.14
p_0 = 0.25
z <- (sqrt(n) * (pbar - p_0))/(sqrt(p_0 * (1 - p_0)))
z
[1] -1.796292
```

Como  $-1.7962925 \in R_c$ , ou seja  $-1.7962925 < -1.6448536$ , rejeita-se  $H_0$  ao nível de significância de 5%, ou seja, concluimos, a partir dos dados, que a proporção de chapas produzidas com anomalias é inferior a 25%.

## Exemplo

- ❶ Os níveis de glicose no sangue para pacientes obesos tem uma média de 100 com um desvio padrão de 15. Um pesquisador acredita que uma dieta rica em amido de milho cru terá um efeito positivo nos níveis de glicose no sangue. Uma amostra de 36 pacientes que fizeram a dieta do amido de milho cru tiveram um nível de glicose médio de 108. Teste a hipótese de que o amido de milho cru foi (ou não) efetivo, a um nível de significância de 5%.

- ❷ Hipótese: a média da população  $\mu = 100$

$$H_0 : \mu = 100 \text{ contra } H_1 : \mu > 100$$

- ❸ Nível de significância: 5%

- ❹ Calcular a estatística de teste, utilizando o *z-score*

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

## Exemplo – cont.

```
mu_0 = 100
xbar = 108
sd = 15
n = 36
z <- (xbar - mu_0)/(sd/sqrt(n))
z
[1] 3.2
pvalue <- pnorm(z)
pvalue <- 1 - pvalue
```

A probabilidade de termos um valor menor do que 108 é 0.9993 e maior ou igual a 108, a probabilidade é 0.000687

- IV Este valor é menor do que o nível de significância de 5%, de modo que rejeitamos a hipótese nula ( $H_0$ ), isto é, **há algum efeito do amido de milho cru.**



## Exercício 1

Dois pesquisadores publicaram um artigo em 2002 onde relatam que a nota média da população na parte quantitativa de uma prova nacional (Graduate Record Examination – GRE) para estudantes que fizeram o exame entre 1994 e 1997 foi  $558 \pm 139$  ( $\mu \pm \sigma$ ). Suponha que selecionemos uma amostra de 100 participantes ( $n = 100$ ). Registramos uma média da amostra igual a 585. Calcule o *p-value* para verificar se a hipótese nula ( $\mu = 558$ ) se mantém a um nível de significância de 0.05 ( $\alpha = 0.05$ ).

## Exercício 2

Um pesquisador deseja estudar o efeito de certa substância no tempo de reação de seres vivos a um certo tipo de estímulo.

Um experimento é desenvolvido com cobaias que são inoculadas com a substância e submetidas a um estímulo elétrico, com seus tempos de reação (em segundos) anotados. Os seguintes valores foram obtidos: 9.1; 9.3; 7.2; 7.5; 13.3; 10.9; 7.2; 9.9; 8.0; 8.6.

Admite-se que o tempo de reação segue o modelo Normal com média 8 e desvio padrão  $\sigma = 2$  segundos. O pesquisador desconfia que o tempo médio sofre alteração por influência da substância.

Efetue um teste de hipótese para verificar se o pesquisador tem razão. Use  $\alpha = 0.06$ , e calcule a probabilidade do erro tipo II supondo que  $\mu = 9$ .

## Exercício 3

Um relatório de uma companhia afirma que 40% de toda a água obtida, através de poços artesianos no nordeste, é salobra.

Há muitas controvérsias sobre essa afirmação, alguns dizem que a proporção é maior, outros que é menor.

Para dirimir as dúvidas, 400 poços foram sorteados e observou-se, em 120 deles água salobra. Qual seria a conclusão, ao nível de 3%?

# Variância desconhecida

## Teste para a média com variância desconhecida

Quando não conhecemos  $\sigma^2$ , usamos como estimativa a variância amostral

$$S^2 = \frac{1}{n-1} \left( \sum_{i=1}^n X_i^2 - n\bar{X}^2 \right),$$

e que a variável padronizada

$$T = \frac{\bar{X} - \mu}{\sqrt{S^2/n}} = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

segue uma distribuição *t-Student* com  $n - 1$  graus de liberdade, ou seja,

$$T \sim t_{(n-1)}$$

## Exemplo 3

Deseja-se investigar se uma certa moléstia que ataca o rim altera o consumo de oxigênio desse órgão. Para indivíduos sadios, admite-se que esse consumo tem distribuição Normal com média  $12 \text{ cm}^3/\text{min}$ .

Os valores medidos em cinco pacientes com a moléstia foram: 14.4; 12.9; 15.1; 13.7; 13.5.

Qual seria a conclusão, ao nível de 1% de significância?

# Testes Qui-Quadrado

Os testes que utilizam o modelo **Qui-Quadrado** como estrutura probabilística são denominados genericamente de **Testes Qui-Quadrado**.

Esse mesmo modelo pode ser utilizado para testes com finalidades diferentes, por exemplo:

- **Teste de aderência:** verifica se uma variável segue uma determinada distribuição
- **Teste de independência:** verifica a independência entre duas variáveis
- **Teste de homogeneidade:** verifica se uma variável se comporta de maneira similar em várias subpopulações

## Testes Qui-Quadrado (cont.)

De maneira geral, qualquer teste Qui-Quadrado envolve a quantidade

$$Q^2 = \sum_{i,j} \frac{(o_{ij} - e_{ij})^2}{e_{ij}}$$

que mede as discrepâncias entre **frequências observadas** ( $o_{ij}$ ) e as **frequências esperadas** ( $e_{ij}$ ). Pode-se mostrar que  $Q^2$  possui distribuição Qui-Quadrado com  $\nu$  **graus de liberdade**, ou seja,

$$Q^2 \sim \chi^2_\nu$$

onde  $\nu$  depende do teste utilizado.

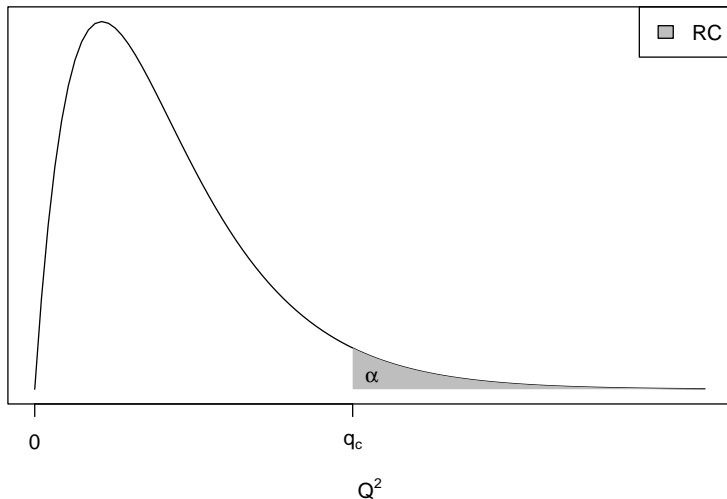
A **Região Crítica** (RC) é constituída de valores grandes de  $Q^2$ ,

$$RC = \{\omega : \omega \geq q_c\}$$

com  $q_c$  sendo determinado pelo nível de significância do teste

$$\alpha = P(Q^2 \geq q_c \mid H_0 \text{ verdadeira})$$

# Distribuição Qui-Quadrado





## Exemplo 4 (teste de aderência)

No exemplo 2, definimos  $X$  como sendo o número de impactos anteriores à falha em um equipamento eletrônico.

Uma amostra de 80 ensaios foi obtida, cada ensaio representando os testes feitos até a interrupção por falha no equipamento, resultado 80 observações da variável de interesse.

Pretende-se verificar se o modelo Geométrico com  $p = 0.4$  é adequado.

## Exemplo 5 (teste de independência)

A tabela abaixo contém os resultados obtidos por estudantes do ensino médio, em um exame com questões nas disciplinas de física e matemática. Para efeito de apresentação na tabela e análise de comportamento, as notas foram classificadas nas categorias alta, média e baixa.

| F     | M    |       |       | Sum |
|-------|------|-------|-------|-----|
|       | Alta | Média | Baixa |     |
| Alta  | 56   | 71    | 12    | 139 |
| Média | 47   | 163   | 38    | 248 |
| Baixa | 14   | 42    | 85    | 141 |
| Sum   | 117  | 276   | 135   | 528 |

Deseja-se testar se existe dependência entre as notas dessas duas disciplinas.

## Exemplo 5 (teste de independência)

Vamos testar as hipóteses:

$H_0$  : As notas de física e matemática são independentes

$H_a$  : As notas não são independentes

Para isso, construímos uma tabela de valores esperados para a casela  $i, j$ :

$$e_{i,j} = \frac{\text{Total da linha } i \times \text{Total da coluna } j}{\text{Total geral}}$$

E assim, obtemos

|       | Alta  | Média  | Baixa |
|-------|-------|--------|-------|
| Alta  | 30.80 | 72.66  | 35.54 |
| Média | 54.95 | 129.64 | 63.41 |
| Baixa | 31.24 | 73.70  | 36.05 |

## Exemplo 5 (teste de independência)

Calculamos

$$q_{obs}^2 = \frac{(56 - 30.8)^2}{30.8} + \dots + \frac{(85 - 36.05)^2}{36.05} = 145.78$$

Nesse caso,  $Q^2$  possui distribuição Qui-Quadrado com  $(r - 1) \times (s - 1)$  graus de liberdade ( $r$  linhas e  $s$  colunas). Para  $\alpha = 0.01$

$$P(Q^2 \geq q_c | H_0) = \alpha \Rightarrow P(Q^2 \geq q_c | H_0) = 0.01$$

obtemos  $q_c = 13.28$  para  $\chi_4^2$ , e portanto,

$$R_C = \{\omega : \omega \geq 13.28\}.$$

Como  $q_{obs}^2 \in R_C$ , concluímos por rejeitar a hipótese nula (de independência entre as notas).

**That's all folks!**