

CE008

Introdução à Bioestatística

Silvia Shimakura
silvia.shimakura@ufpr.br



Laboratório de Estatística e Geoinformação



Objetivo da disciplina

Conhecer metodologias estatísticas para produção, descrição e análise de dados em contextos relacionados às ciências biológicas.

Programa estatístico

- ▣ Ambiente de análise estatística de dados: R
- ▣ Livre - Gratuito e de código aberto
- ▣ Utilizado como ferramenta didática
- ▣ <http://www.r-project.org>



Conteúdo

Introdução

Estatística Descritiva

Estatística Inferencial

Distribuição t de Student e Teste de Hipóteses

Testes Não Paramétricos

Tabelas de Contingência e Teste Qui-quadrado

Quadros de Síntese

Aspectos históricos

- ▣ A palavra **Estatística** provém do latim status, que significa estado.
- ▣ A utilização primitiva envolvia compilações de dados e gráficos que descreviam aspectos de um estado ou país.
- ▣ Com o desenvolvimento das ciências, da Teoria da Probabilidade e da Informática, a Estatística adquiriu status de Ciência com aplicabilidade em praticamente todas as áreas do saber.

Bioestatística

- ▣ Fornece métodos para se tomar decisões na presença de **incerteza**
- ▣ Estabelece **faixas de confiança** para eficácia dos tratamentos
- ▣ Verifica a influência de **fatores de risco** no aparecimento de doenças

[Soares e Siqueira, 2002]

Estatística / Bioestatística

▣ **Estatística Descritiva**

- ▣ Objetivo: Descrever dados amostrais
- ▣ Ferramentas: Tabelas, gráficos, medidas de posição, medidas de tendência central, medidas de dispersão

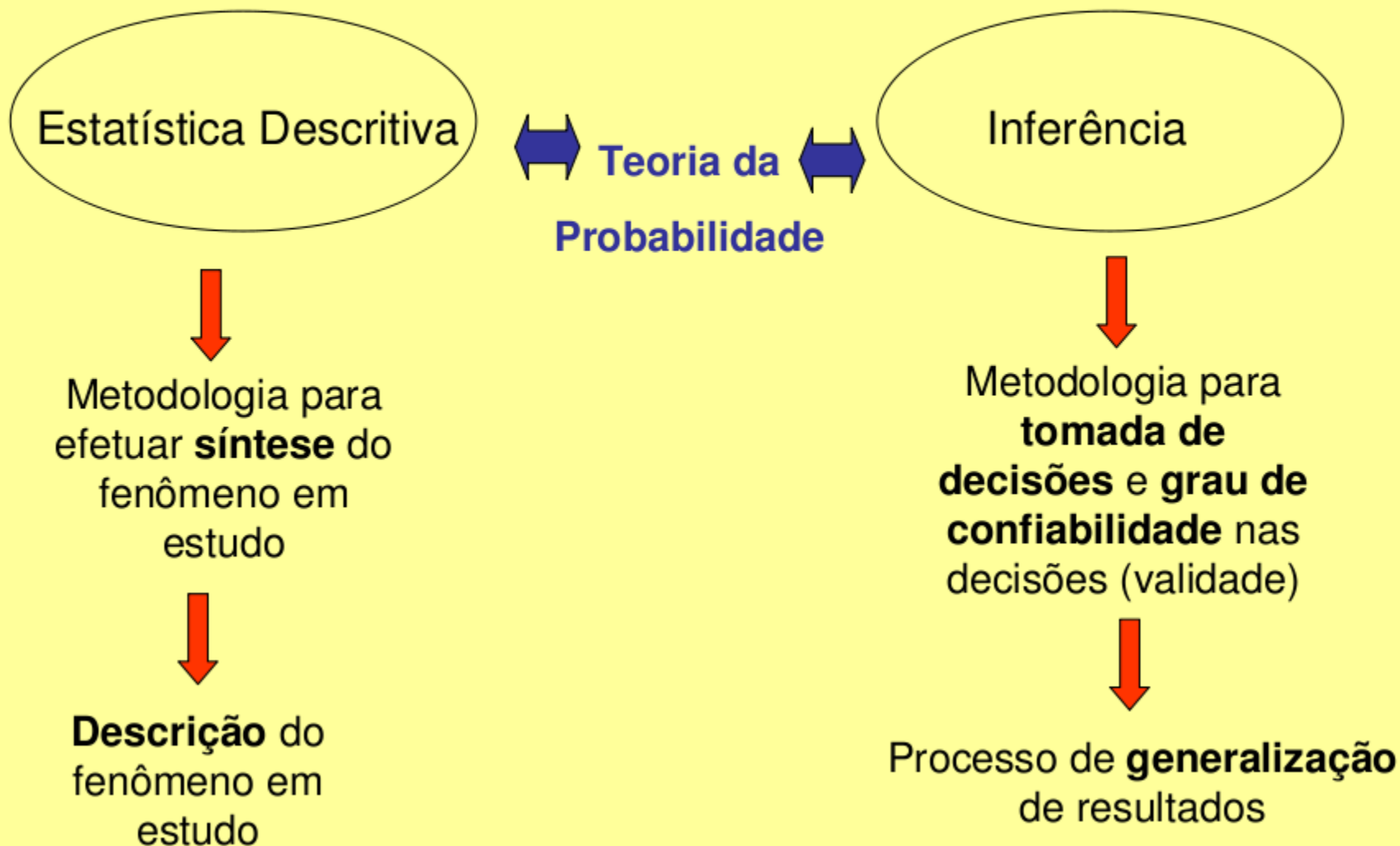
▣ **Estatística Inferencial**

- ▣ Objetivo: Retirar informação útil sobre a população partindo de dados amostrais
- ▣ Ferramentas: Estimativas pontuais e de intervalo de parâmetros populacionais, testes de hipóteses

- ▣ A ligação entre as duas se dá através da **teoria de probabilidades**



Campos ou funções da Estatística



Conceitos

- ▣ **População:** conjunto de elementos que apresentam uma ou mais características em comum, cujo comportamento interessa analisar (inferir)

- ▣ **Fatores limitantes:**
 - Populações infinitas
 - Custo
 - Tempo
 - Processos destrutivos

Conceitos

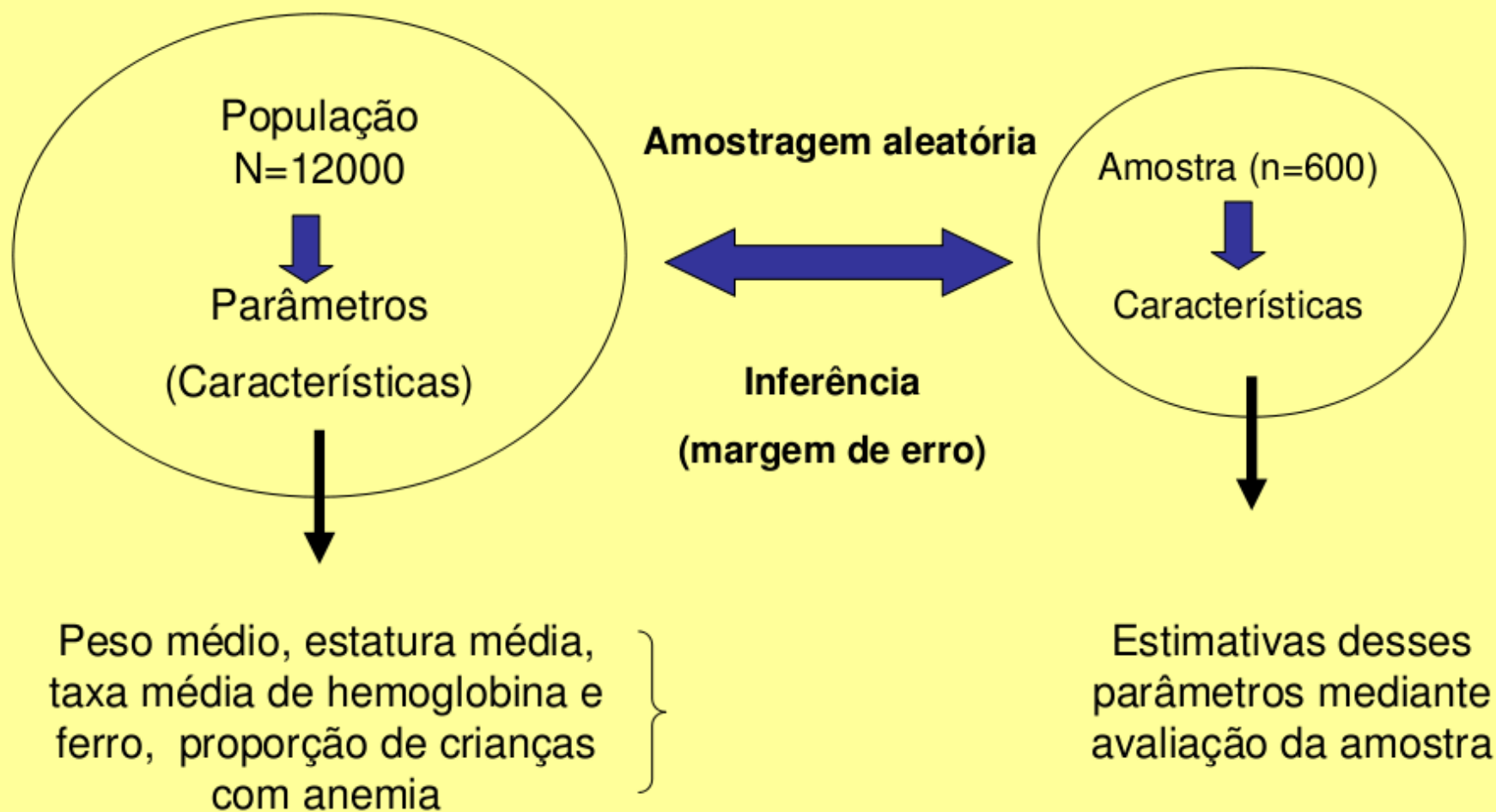
- ▮ **Amostra:** é um subconjunto de os elementos (sujeitos, medidas, valores, etc.) extraídos da população em estudo.
- ▮ Amostragem é um conjunto de técnicas para se obter amostras.

Conceitos relacionados a população e amostra

- **Parâmetro** é um valor ou uma medida numérica que descreve uma característica *populacional*.
(São valores estabelecidos para a população)
- **Estimativa** é um valor ou uma medida que descreve uma característica de uma *amostra*
(são medidas ou valores estabelecidos para uma amostra)

Um exemplo

Estudo da anemia em crianças com idade entre 5 e 7 anos, numa região do município com uma população de 12000 crianças nessa faixa etária.



Estatística Descritiva

Tipos de variáveis, medidas de tendência central, medidas de dispersão, gráficos e tabelas



Tipos de Variáveis

- ▣ Quantitativas
 - ▣ Discretas
 - ▣ Contínuas
- ▣ Qualitativas (Categóricas)
 - ▣ Ordinais
 - ▣ Nominais



Medidas de Tendência Central

- ▮ Moda
- ▮ Média
- ▮ Mediana



Quantis

- ▮ Posição das observações
- ▮ Quantis
- ▮ Mediana
- ▮ Quartis
- ▮ Percentis



Medidas de Dispersão

- ▮ Amplitude
- ▮ Amplitude interquartis
- ▮ Variância
- ▮ Desvio padrão



Tabelas e Gráficos

- ▣ Tabela de frequências
 - ▣ Frequência absoluta
 - ▣ Frequência relativa
 - ▣ Frequência cumulativa
- ▣ Tabelas de contingência (2×2 ; $l \times c$)
- ▣ Gráfico de setores
- ▣ Gráfico de barras
- ▣ Histograma
- ▣ Polígono de frequências
- ▣ Diagrama de dispersão
- ▣ Box plot (mediana, amplitude inter-quartis)
- ▣ Error bar (média, IC 95%)



Probabilidade

- ▣ Qualidade de testes diagnósticos
- ▣ Distribuição Binomial
- ▣ Distribuição Normal



Testes diagnósticos

- ▣ Testes diagnósticos: baseados em observações, questionários ou exames de laboratório utilizados para classificar indivíduos em categorias
 - Ex: taxa de glicose no sangue para diagnóstico de diabetes
- ▣ Os testes podem ser imperfeitos e resultar em classificações incorretas.
- ▣ Antes de ser adotado deve ser avaliado para verificar a capacidade de acerto.
- ▣ Avaliação feita aplicando-se o teste a dois grupos de pessoas: um grupo doente e um grupo não doente.
- ▣ O diagnóstico é feito por um teste chamado **padrão ouro**.

Organização dos resultados

True status	Screening Test Result		Total
	Positive	Negative	
Diseased	a	b	$a + b$
Not diseased	c	d	$c + d$
Total	$a + c$	$b + d$	N

Sensibilidade e Especificidade

- **Sensibilidade:** Probabilidade de teste positivo num paciente doente → capacidade de reação do teste num paciente doente
- **Especificidade:** Probabilidade de teste negativo num paciente não doente → capacidade de não reação do teste num paciente não doente

Organização dos resultados

True status	Screening Test Result		Total
	Positive	Negative	
Diseased	a	b	$a + b$
Not diseased	c	d	$c + d$
Total	$a + c$	$b + d$	N

$$\text{sensitivity} = \frac{a}{a + b}$$

$$\text{specificity} = \frac{d}{c + d}$$

Exemplo: Câncer de colo do útero

- Doença com alta chance de refreamento se detectada no início
- Procedimento de triagem: **Papanicolau**
- 16,25% dos testes realizados em mulheres com câncer resultaram em falsos negativos

$$P(T-|D+)=0,1625$$

$$\text{sensibilidade}=P(T+|D+)=1-P(T-|D+)=0,8375$$

- 83,75% das mulheres que tinham câncer de colo do útero apresentaram resultados positivos

Exemplo: Câncer de colo do útero (cont.)

- Nem todas as mulheres testadas sofriam de câncer de colo do útero.
- 18,64% dos testes resultaram falsos positivos

$$P(T+|D-)=0,1864$$

$$\text{especificidade}=P(T-|D-)=1-P(T+|D-)=0,8136$$

- 81,36% das mulheres que não tinham câncer de colo do útero apresentaram resultados negativos

VPP e VPN

Os índices acima são bons sintetizadores das qualidades gerais de um teste mas: **Não ajudam a decisão do médico que precisa concluir se um paciente com resultado positivo, tem a doença.**

- Probabilidade de uma pessoa ter a doença sabendo-se que tem teste positivo: $P(D+|T+)$

Valor preditivo positivo (VPP)

- Probabilidade de uma pessoa não ter a doença sabendo-se que tem teste negativo: $P(D-|T-)$

Valor preditivo negativo (VPN)

Organização dos resultados

True status	Screening Test Result		Total
	Positive	Negative	
Diseased	a	b	$a + b$
Not diseased	c	d	$c + d$
Total	$a + c$	$b + d$	N

$$\text{sensitivity} = \frac{a}{a + b}$$

$$\text{specificity} = \frac{d}{c + d}$$

$$\text{positive predictive value} = \frac{a}{a + c}$$

$$\text{negative predictive value} = \frac{d}{b + d}$$

VPP e VPN

VPP e VPN só podem ser calculados diretamente da tabela se a prevalência estimada for próxima à prevalência populacional

	T+	T-	Total
D+	10	10	20
D-	30	70	100
Total	40	80	120

VPP=10/40=0,25

	T+	T-	Total
D+	20	20	40
D-	24	56	80
Total	44	76	120

VPP=20/44=0,45

Aplicação do Teorema de Bayes

- Queremos obter $P(D_+|T_+)$

$$P(D_+|T_+) = \frac{P(D_+ \cap T_+)}{P(T_+)} = \frac{P(T_+|D_+)P(D_+)}{P(T_+|D_+)P(D_+) + P(T_+|D_-)P(D_-)}$$

- Temos: $P(T_+|D_+) = 0,8375$ e $P(T_+|D_-) = 0,1864$
- Precisamos de $P(D_+)$ e $P(D_-)$
 - $P(D_+) = 0,000083$ (prevalência: 83 por 1.000.000)
 - $P(D_-) = 1 - P(D_+) = 1 - 0,000083 = 0,999917$

Aplicação do Teorema de Bayes (cont.)

$$P(D_+|T_+) = \frac{0,000083 \times 0,8375}{(0,000083 \times 0,8375) + (0,999917 \times 0,1864)} = 0,000373$$

Para cada 1.000.000 de mulheres com Papanicolau positivos, 373 casos de câncer de colo do útero →
VPP

Aplicação do Teorema de Bayes (cont.)

$$P(D_-|T_-) = \frac{0,999917 \times 0,8136}{(0,999917 \times 0,8136) + (0,000083 \times 0,1625)} = 0,999983$$

Para cada 1.000.000 de mulheres com Papanicolau negativos, 999.983 não sofrem de câncer de colo do útero → **VPN**

Cálculo de VPP e VPN

$$VPP = \frac{sp}{sp + (1 - e)(1 - p)}$$

$$VPN = \frac{e(1 - p)}{(1 - s)p + e(1 - p)}$$

Acurácia

- ▣ Valores preditivos variam de acordo com a prevalência da doença na população
- ▣ Sensibilidade e especificidade não variam com a prevalência da doença pois consideram doentes e não doentes separadamente
- ▣ Para um teste baseado em uma medida contínua, a escolha do ponto de corte é importante pois altera a sensibilidade e a especificidade do teste

Exemplo

Example 1.1: Enzyme tests and myocardial infarction (MI): use of creatinine kinase (CK) assay in a coronary care unit. The data obtained were as follows:

CK activity	MI	non-MI
0-49	2	32
50-99	4	10
100-149	6	5
150-399	14	2
400+	21	0
Total no. patients	47	49

	CK		Total
	< 50 (-ve)	≥ 50 (+ve)	
MI	2	45	47
Non-MI	32	17	49
Total	34	62	96

sensitivity $45/47 = 0.96$

specificity $32/49 = 0.65$

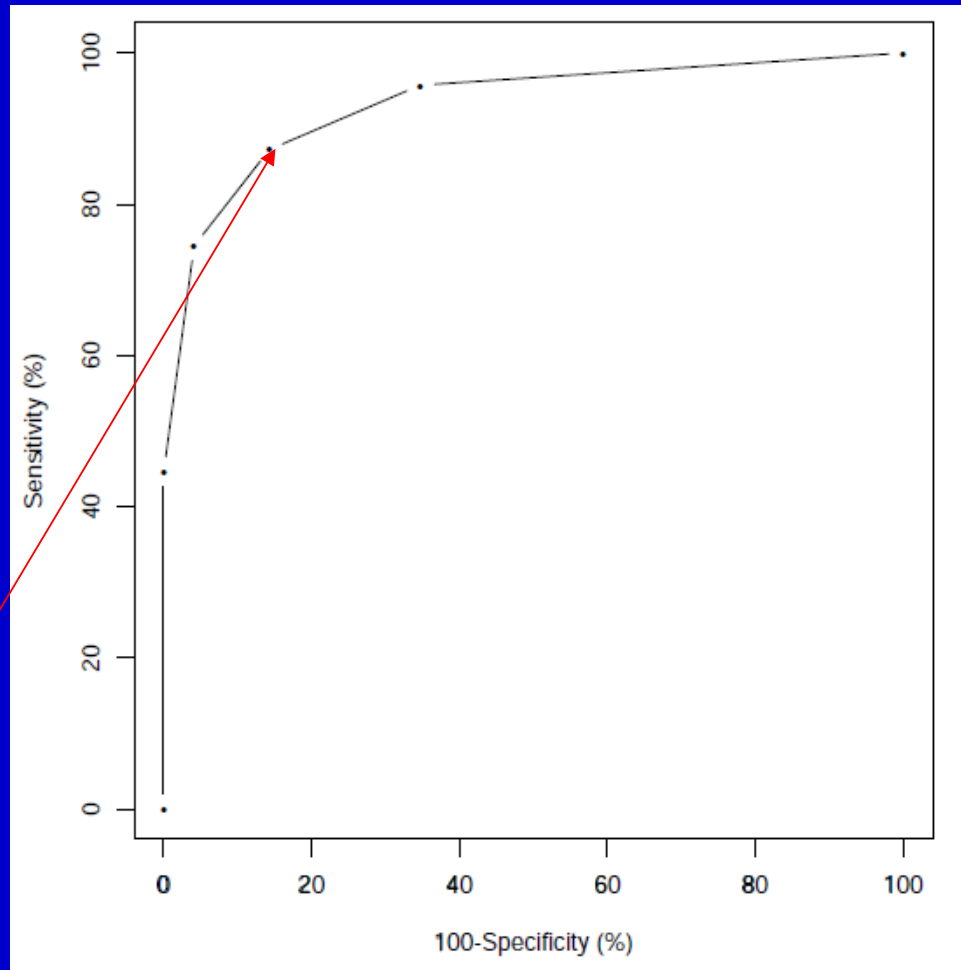
Exemplo (cont.)

Possible cutoff	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Negative predictive value (%)
50	96	65	73	94
100	87	86	85	88
150	74	96	95	80
400	45	100	100	65

Curva ROC

(Receiver Operating Characteristic)

- Não havendo preferência por um teste mais sensível ou mais específico
- Escolhe-se o ponto de corte no canto extremo esquerdo no topo do gráfico



Distribuições de Probabilidade

Exemplo: Eficácia de medicamento

- ▣ Uma indústria farmacêutica afirma que um certo medicamento alivia os sintomas de angina pectoris em 80% dos pacientes.
- ▣ Você prescreve este medicamento a 5 dos seus pacientes com angina mas somente 2 (40%) relatam alívio dos sintomas.
- ▣ Se a afirmação do fabricante for verdadeira, é possível obter resultados tão ruins ou ainda piores do que os que você observou?

Exemplo: Eficácia de medicamento (cont.)

- Assumindo

$$P(\text{alívio dos sintomas})=0,8$$

- X : #pacientes sentem alívio dos sintomas dentre 5 pacientes

- $P(X \leq 2) = P(X=2) + P(X=1) + P(X=0)$

Exemplo: Eficácia de medicamento (cont.)

x	Sequência	P(Sequência)
2	AANNN	$0,8 \times 0,8 \times 0,2 \times 0,2 \times 0,2 = 0,00514$
2	ANANN	$0,8 \times 0,2 \times 0,8 \times 0,2 \times 0,2 = 0,00514$
2	ANNAN	$0,8 \times 0,2 \times 0,2 \times 0,8 \times 0,2 = 0,00514$
2	ANNNA	$0,8 \times 0,2 \times 0,2 \times 0,2 \times 0,8 = 0,00514$
2	NAANN	$0,2 \times 0,8 \times 0,8 \times 0,2 \times 0,2 = 0,00514$
2	NANAN	$0,2 \times 0,8 \times 0,2 \times 0,8 \times 0,2 = 0,00514$
2	NANNA	$0,2 \times 0,8 \times 0,2 \times 0,2 \times 0,8 = 0,00514$
2	NNAAN	$0,2 \times 0,2 \times 0,8 \times 0,8 \times 0,2 = 0,00514$
2	NNANA	$0,2 \times 0,2 \times 0,8 \times 0,2 \times 0,8 = 0,00514$
2	NNNAA	$0,2 \times 0,2 \times 0,2 \times 0,8 \times 0,8 = 0,00514$

$$\binom{5}{2} = 10$$

Sequências possíveis

Exemplo: Eficácia de medicamento (cont.)

X	Sequência	P(Sequência)
2	AANNN	$0,8 \times 0,8 \times 0,2 \times 0,2 \times 0,2 = 0,00514$
2	ANANN	$0,8 \times 0,2 \times 0,8 \times 0,2 \times 0,2 = 0,00514$
2	ANNAN	$0,8 \times 0,2 \times 0,2 \times 0,8 \times 0,2 = 0,00514$
2	ANNNA	$0,8 \times 0,2 \times 0,2 \times 0,2 \times 0,8 = 0,00514$
2	NAANN	$0,2 \times 0,8 \times 0,8 \times 0,2 \times 0,2 = 0,00514$
2	NANAN	$0,2 \times 0,8 \times 0,2 \times 0,8 \times 0,2 = 0,00514$
2	NANNA	$0,2 \times 0,8 \times 0,2 \times 0,2 \times 0,8 = 0,00514$
2	NNAAN	$0,2 \times 0,2 \times 0,8 \times 0,8 \times 0,2 = 0,00514$
2	NNANA	$0,2 \times 0,2 \times 0,8 \times 0,2 \times 0,8 = 0,00514$
2	NNNAA	$0,2 \times 0,2 \times 0,2 \times 0,8 \times 0,8 = 0,00514$
P(X=2)		$10 \times 0,8^2 \times 0,2^3 = 0,0514$

$$\binom{5}{2} = 10$$

Sequências possíveis

Exemplo: Eficácia de medicamento (cont.)

X	Sequência	P(Sequência)
1	ANNNN	$0,8 \times 0,2 \times 0,2 \times 0,2 \times 0,2 = 0,00128$
1	NANNN	$0,2 \times 0,8 \times 0,2 \times 0,2 \times 0,2 = 0,00128$
1	NNANN	$0,2 \times 0,2 \times 0,8 \times 0,2 \times 0,2 = 0,00128$
1	NNNAN	$0,2 \times 0,2 \times 0,2 \times 0,8 \times 0,2 = 0,00128$
1	NNNNA	$0,2 \times 0,2 \times 0,2 \times 0,2 \times 0,8 = 0,00128$

$$\binom{5}{1} = 5$$

Sequências possíveis

Exemplo: Eficácia de medicamento (cont.)

X	Sequência	P(Sequência)
1	ANNNN	$0,8 \times 0,2 \times 0,2 \times 0,2 \times 0,2 = 0,00128$
1	NANNN	$0,2 \times 0,8 \times 0,2 \times 0,2 \times 0,2 = 0,00128$
1	NNANN	$0,2 \times 0,2 \times 0,8 \times 0,2 \times 0,2 = 0,00128$
1	NNNAN	$0,2 \times 0,2 \times 0,2 \times 0,8 \times 0,2 = 0,00128$
1	NNNNA	$0,2 \times 0,2 \times 0,2 \times 0,2 \times 0,8 = 0,00128$
P(X=1)		$5 \times 0,8^1 \times 0,2^4 = 0,0064$

$$\binom{5}{1} = 5$$

Sequências possíveis

Exemplo: Eficácia de medicamento (cont.)

X	Sequência	P(Sequência)
1	ANNNN	$0,8 \times 0,2 \times 0,2 \times 0,2 \times 0,2 = 0,00128$
1	NANNN	$0,2 \times 0,8 \times 0,2 \times 0,2 \times 0,2 = 0,00128$
1	NNANN	$0,2 \times 0,2 \times 0,8 \times 0,2 \times 0,2 = 0,00128$
1	NNNAN	$0,2 \times 0,2 \times 0,2 \times 0,8 \times 0,2 = 0,00128$
1	NNNNA	$0,2 \times 0,2 \times 0,2 \times 0,2 \times 0,8 = 0,00128$
P(X=1)		$5 \times 0,8^1 \times 0,2^4 = 0,0064$
0	NNNNN	$0,2 \times 0,2 \times 0,2 \times 0,2 \times 0,2 = 0,00032$
P(X=0)		$1 \times 0,8^0 \times 0,2^5 = 0,00032$

$$\binom{5}{1} = 5$$

Sequências possíveis

$$\binom{5}{0} = 1$$

Exemplo: Eficácia de medicamento (cont.)

X	Sequência	P(Sequência)
1	ANNNN	$0,8 \times 0,2 \times 0,2 \times 0,2 \times 0,2 = 0,00128$
1	NANNN	$0,2 \times 0,8 \times 0,2 \times 0,2 \times 0,2 = 0,00128$
1	NNANN	$0,2 \times 0,2 \times 0,8 \times 0,2 \times 0,2 = 0,00128$
1	NNNAN	$0,2 \times 0,2 \times 0,2 \times 0,8 \times 0,2 = 0,00128$
1	NNNNA	$0,2 \times 0,2 \times 0,2 \times 0,2 \times 0,8 = 0,00128$
P(X=1)		$5 \times 0,8^1 \times 0,2^4 = 0,0064$
0	NNNNN	$0,2 \times 0,2 \times 0,2 \times 0,2 \times 0,2 = 0,00032$
P(X=0)		$1 \times 0,8^0 \times 0,2^5 = 0,00032$

$$\binom{5}{1} = 5$$

Sequências possíveis

$$\binom{5}{0} = 1$$

$$P(X \leq 2) = 0,0514 + 0,0064 + 0,00032 = 0,05812$$

Exemplo: Eficácia de medicamento (cont.)

Se a afirmação do fabricante for verdadeira, a chance de se obter resultados tão ruins ou ainda piores do que os observados é de 5,8%.

CONCLUSÃO?

Distribuição Binomial

- n : no. ensaios (independentes)
- X : no. sucessos nos n ensaios
- p : prob. sucesso num ensaio

$$P(X=x) = \binom{n}{x} p^x (1-p)^{n-x}$$

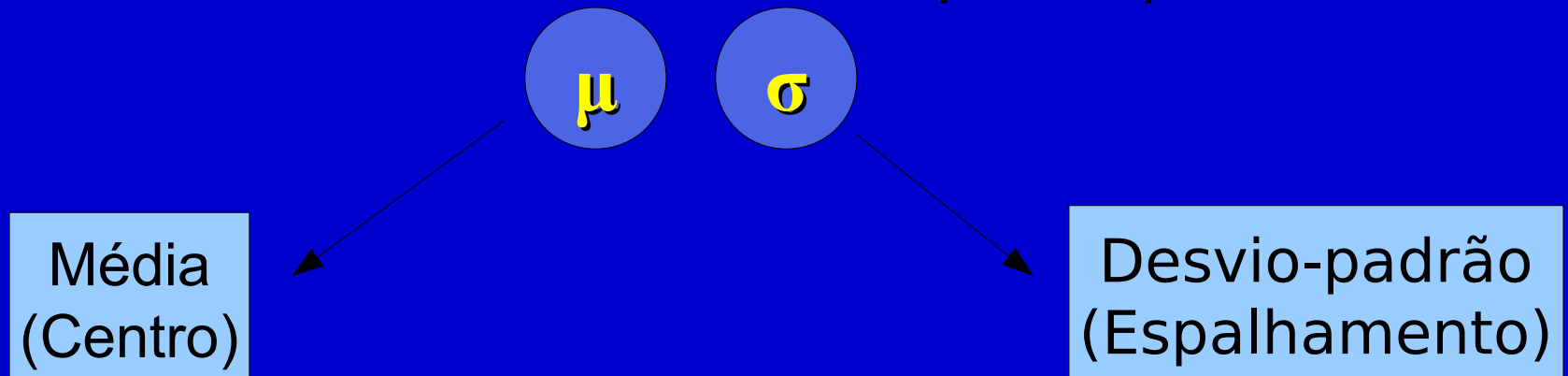
$$P(X=0) + P(X=1) + \dots + P(X=n) = 1$$

Calculadora

<http://onlinestatbook.com/2/java/binomialProb.html>

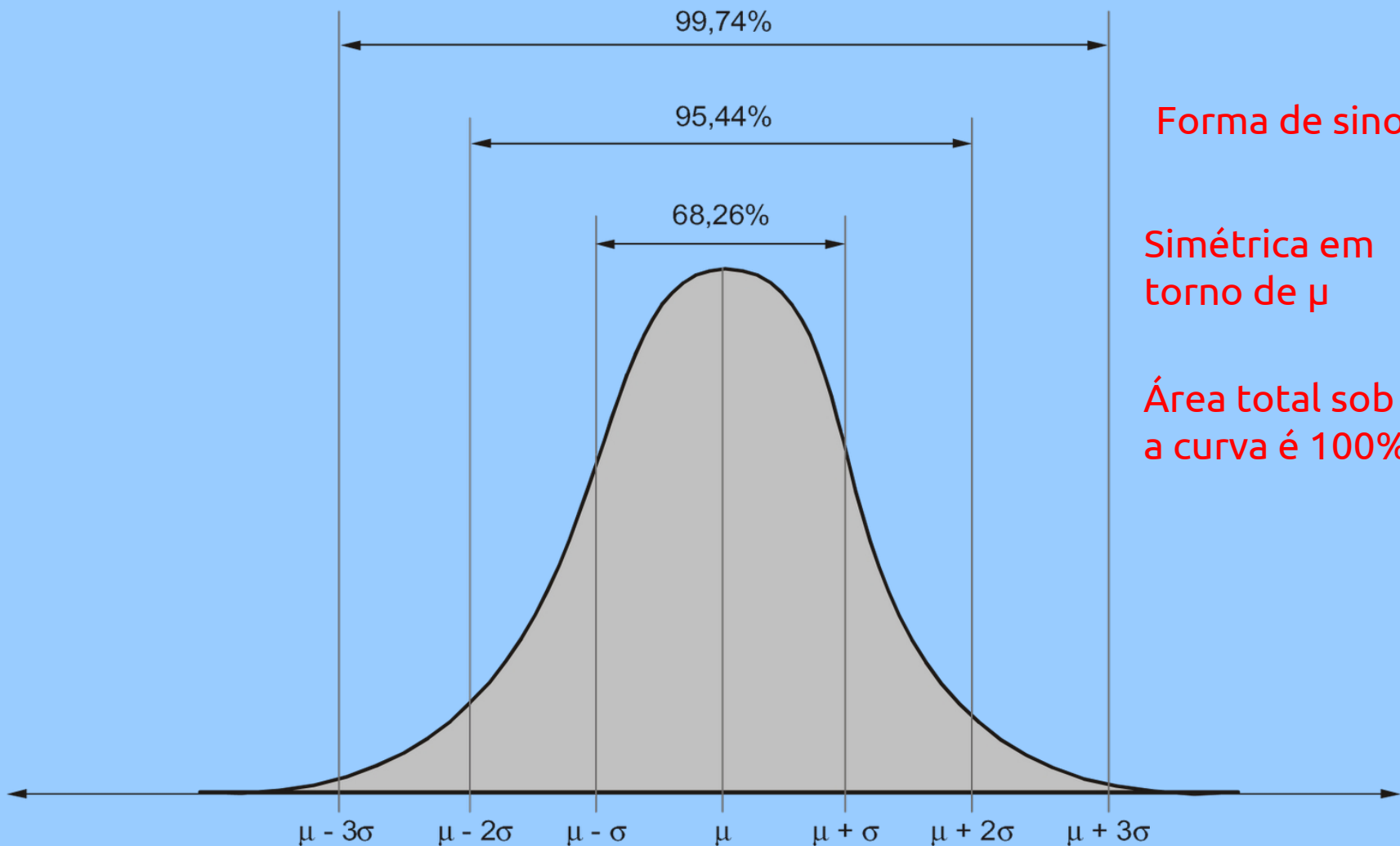
Distribuição Normal

- ▣ Diversas variáveis contínuas tais como, altura, peso, níveis de colesterol, pressão sistólica e diastólica, podem ser descritas pela distribuição normal
- ▣ Formato da curva definido por 2 parâmetros:



Equação:

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\}$$

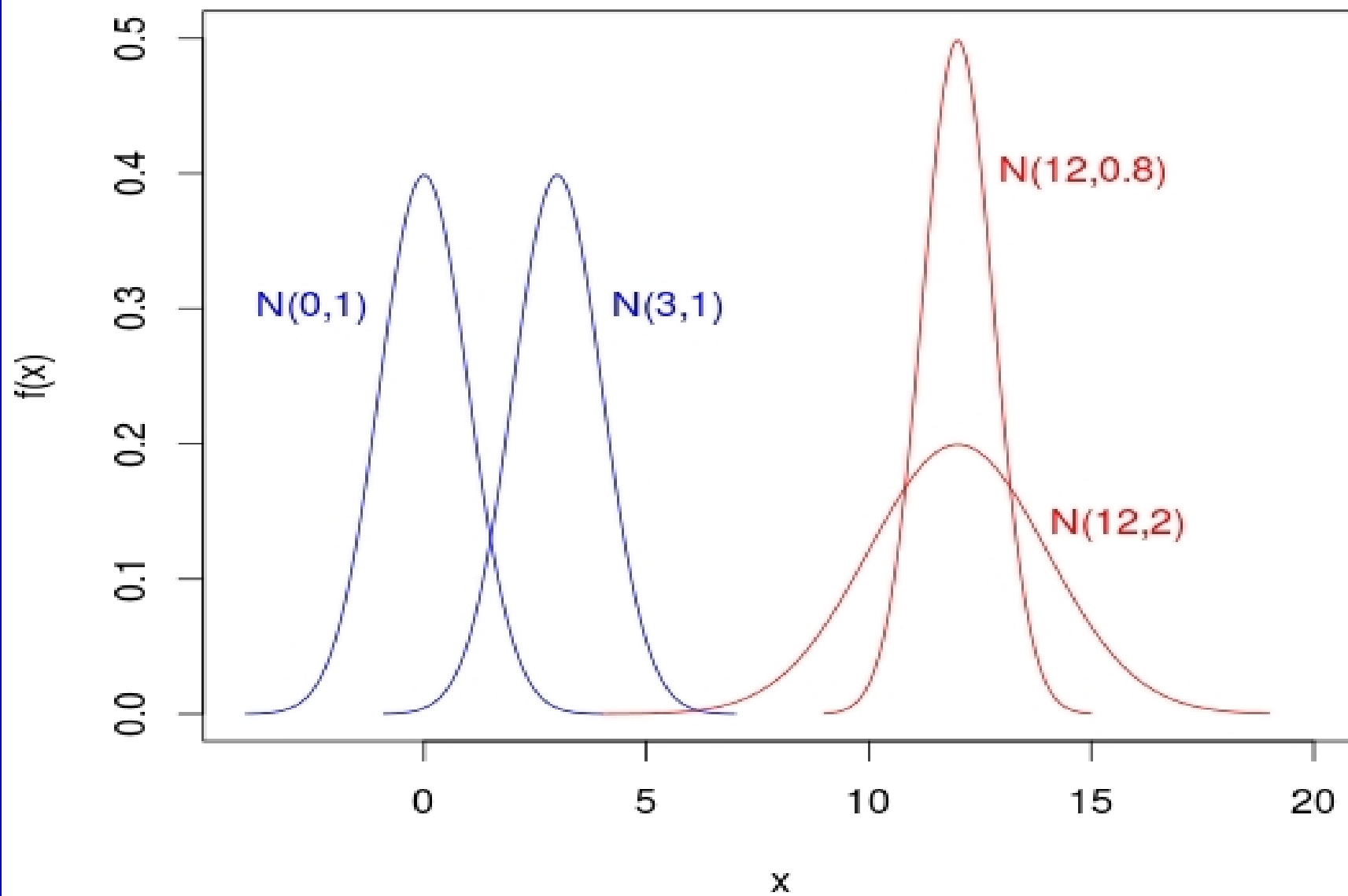


Forma de sino

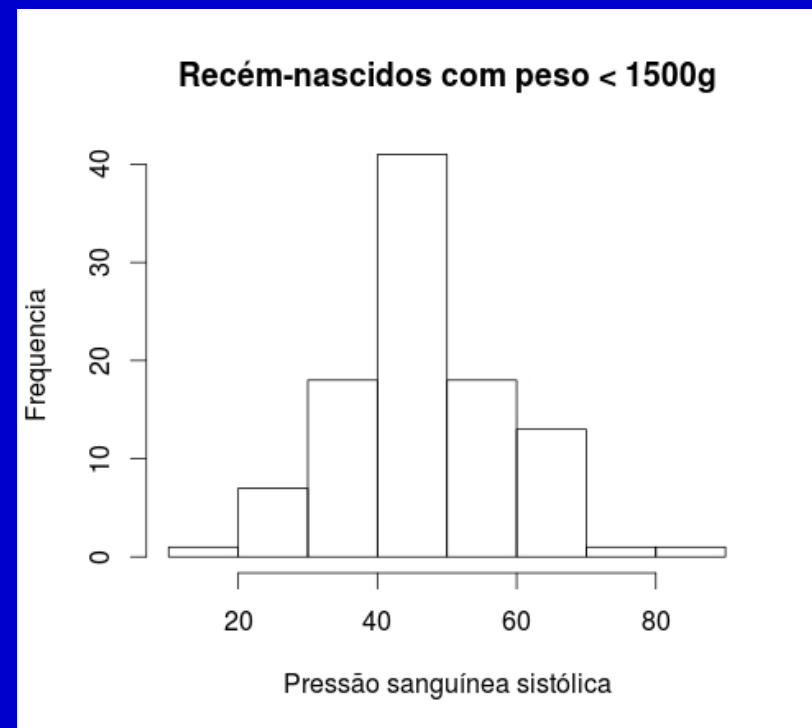
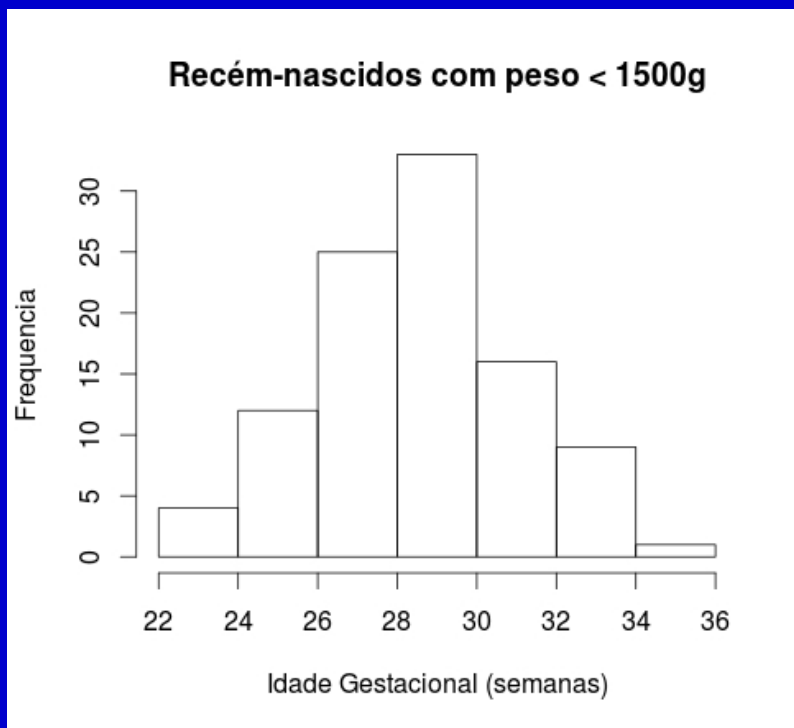
Simétrica em torno de μ

Área total sob a curva é 100%

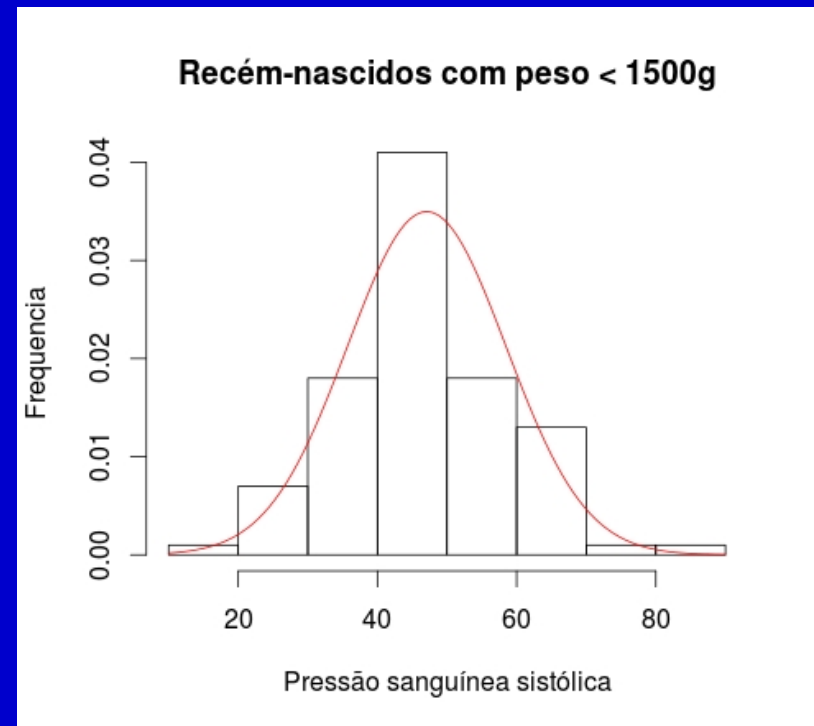
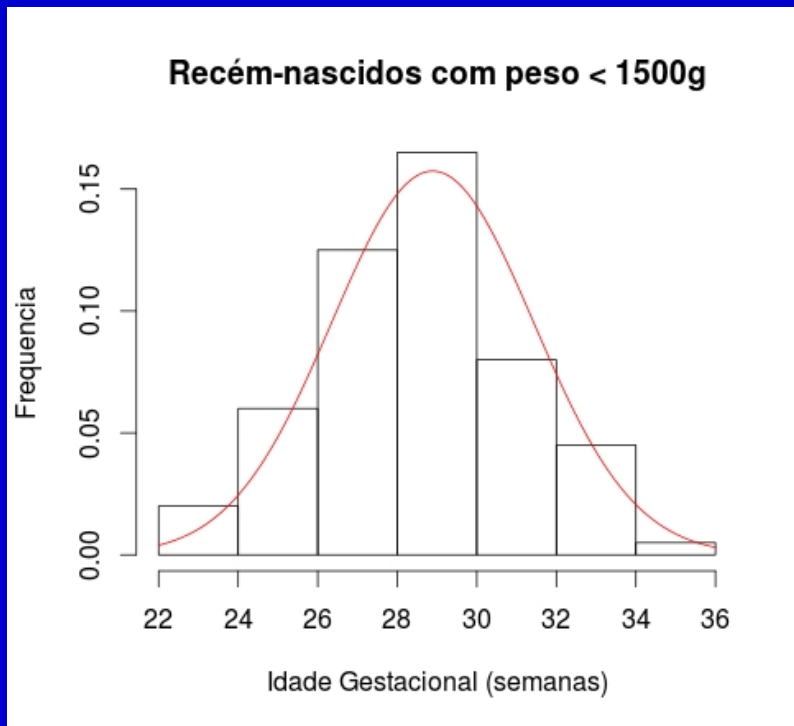
Notação: $N(\mu, \sigma)$



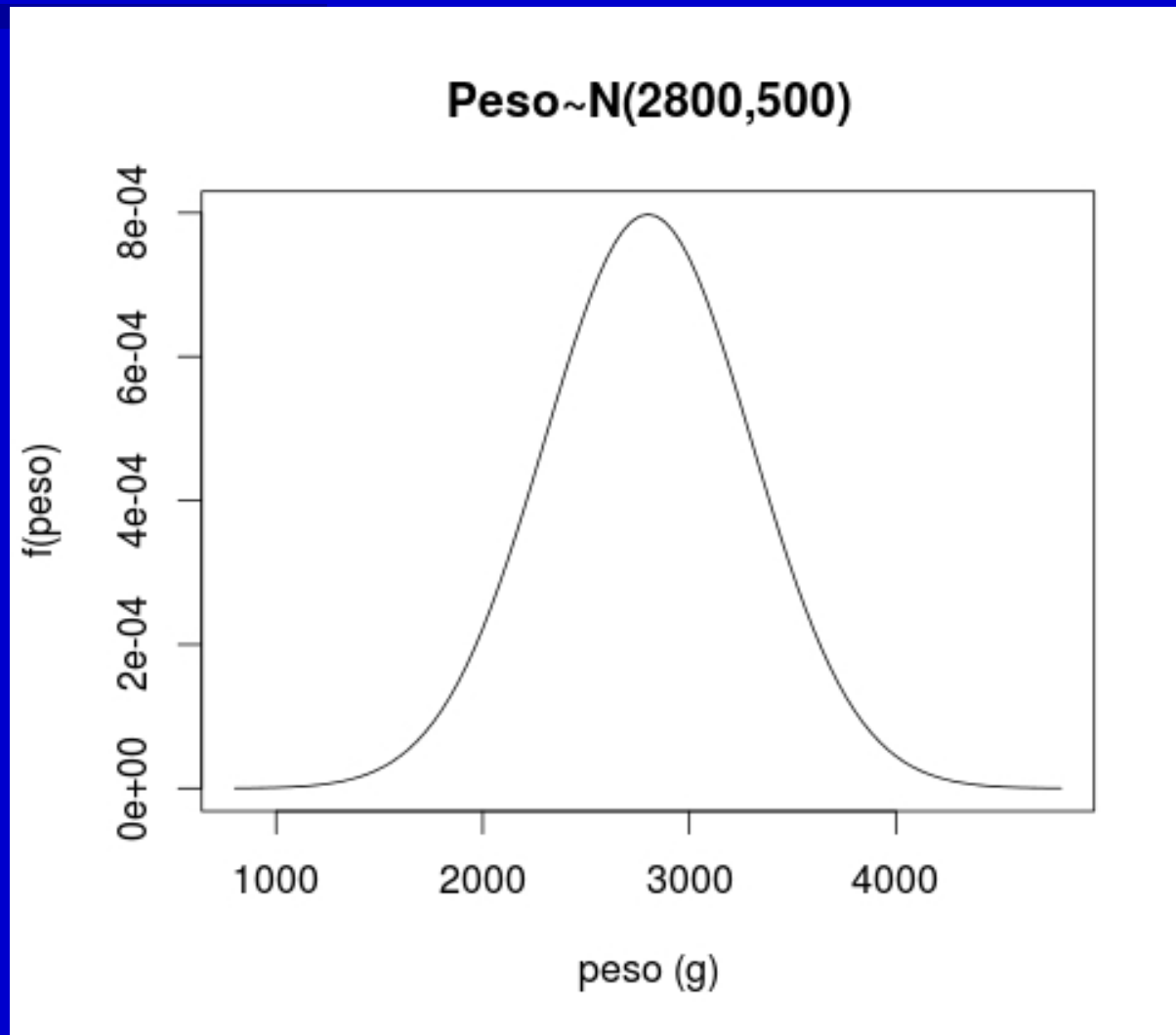
Amostra de 100 recém-nascidos com peso < 1500g em Boston, Massachusetts



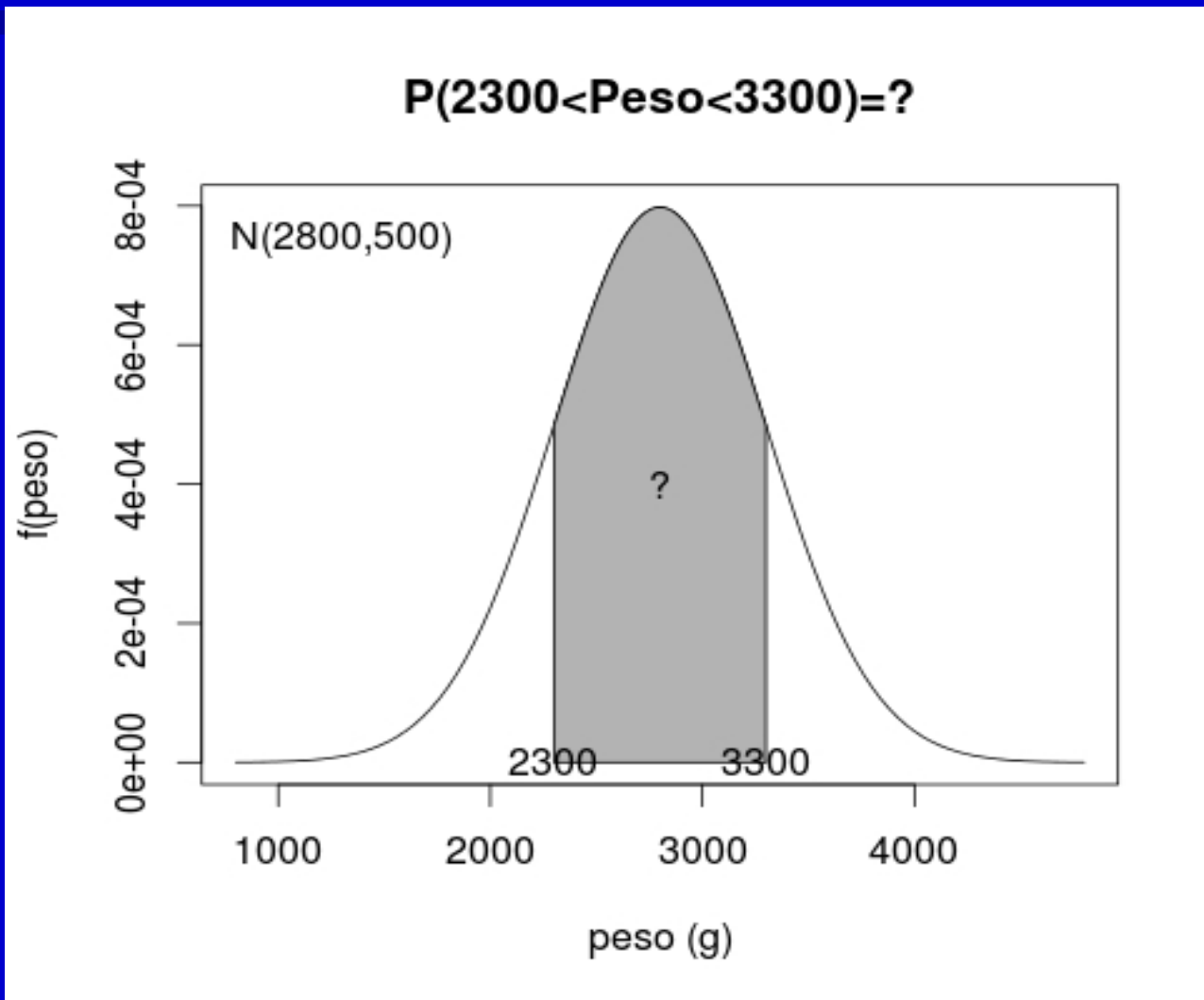
Amostra de 100 recém-nascidos com peso < 1500g em Boston, Massachusetts



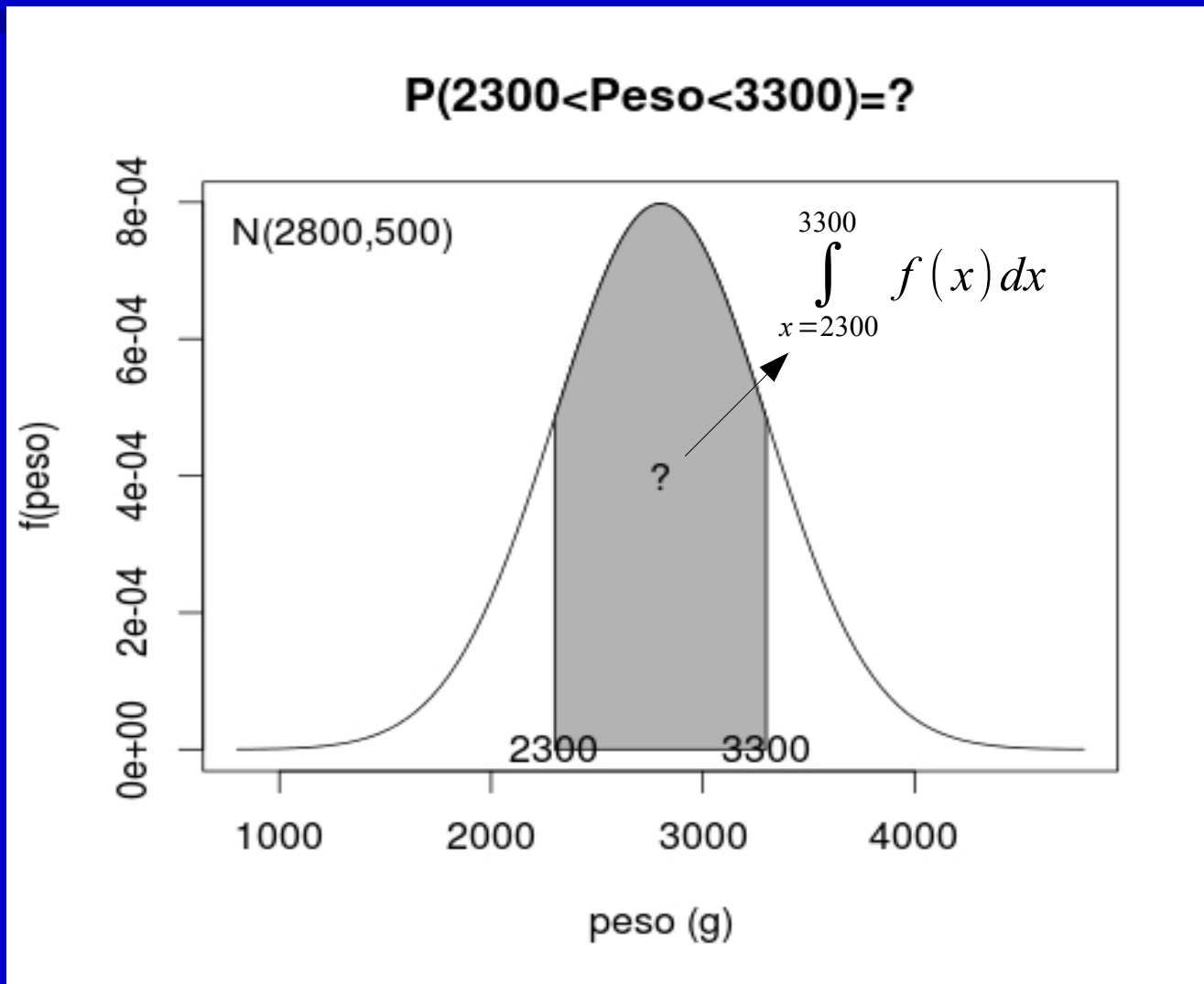
Exemplo: peso de recém-nascidos



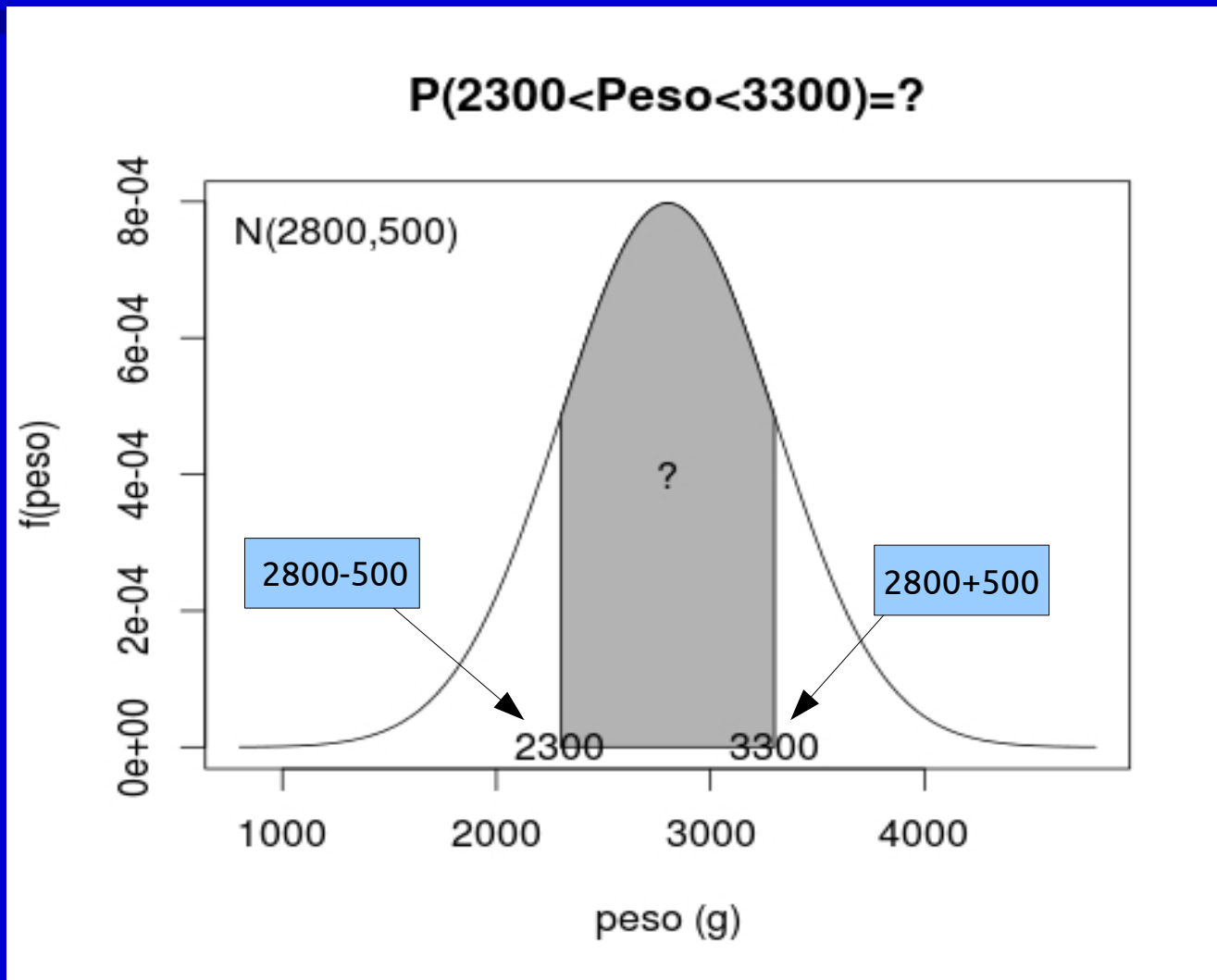
Exemplo: peso de recém-nascidos



Exemplo: peso de recém-nascidos

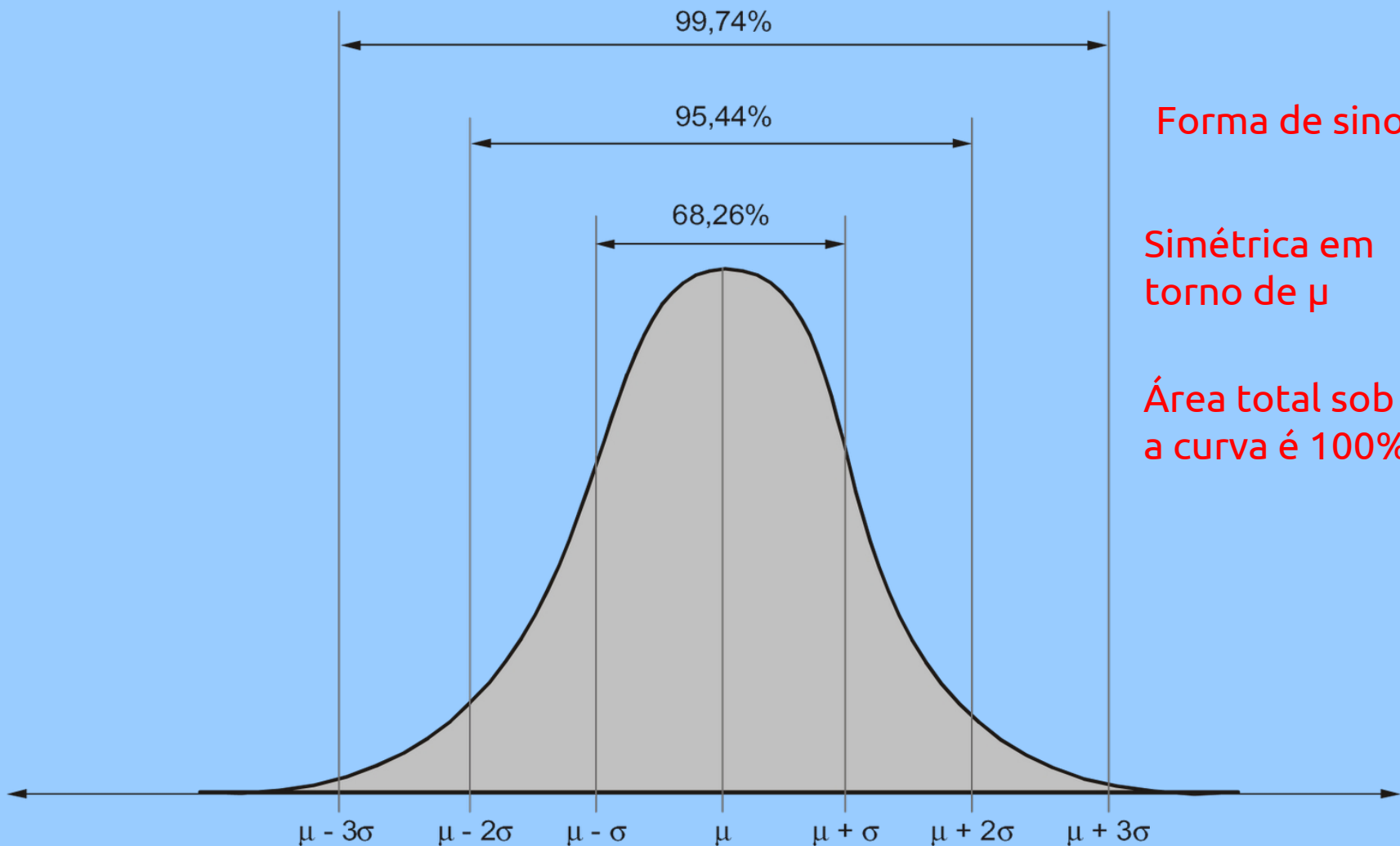


Exemplo: peso de recém-nascidos



Equação:

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\}$$

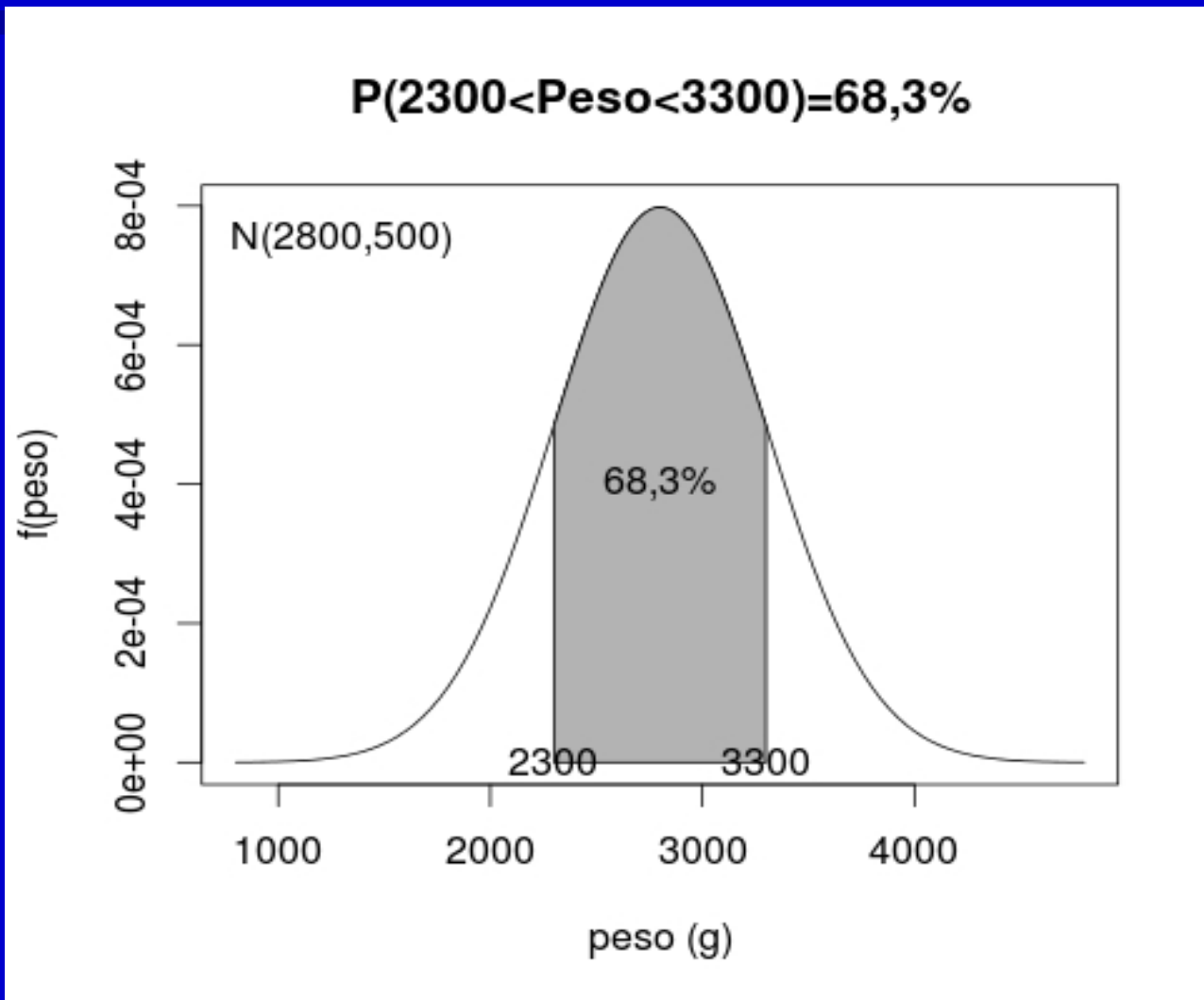


Forma de sino

Simétrica em torno de μ

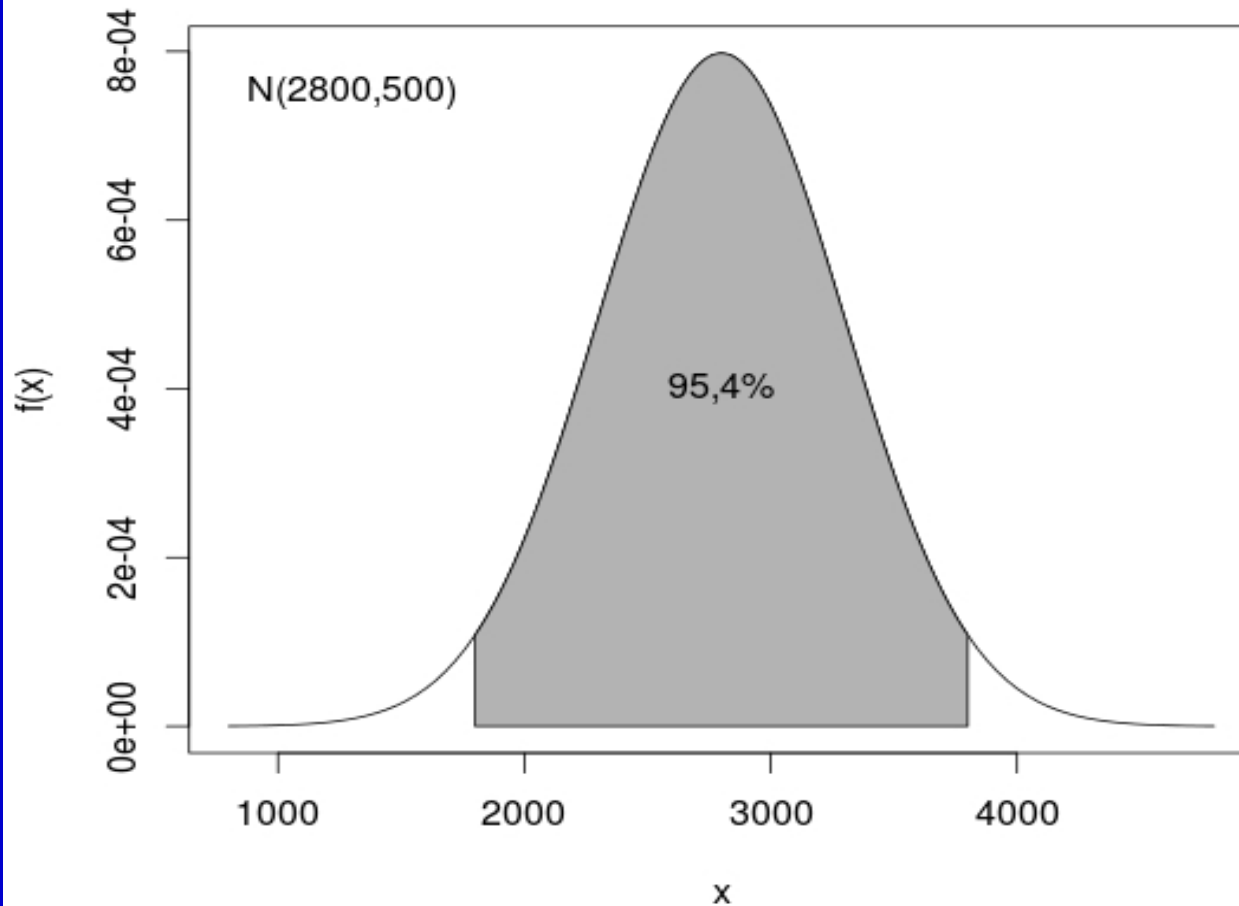
Área total sob a curva é 100%

Exemplo: peso de recém-nascidos

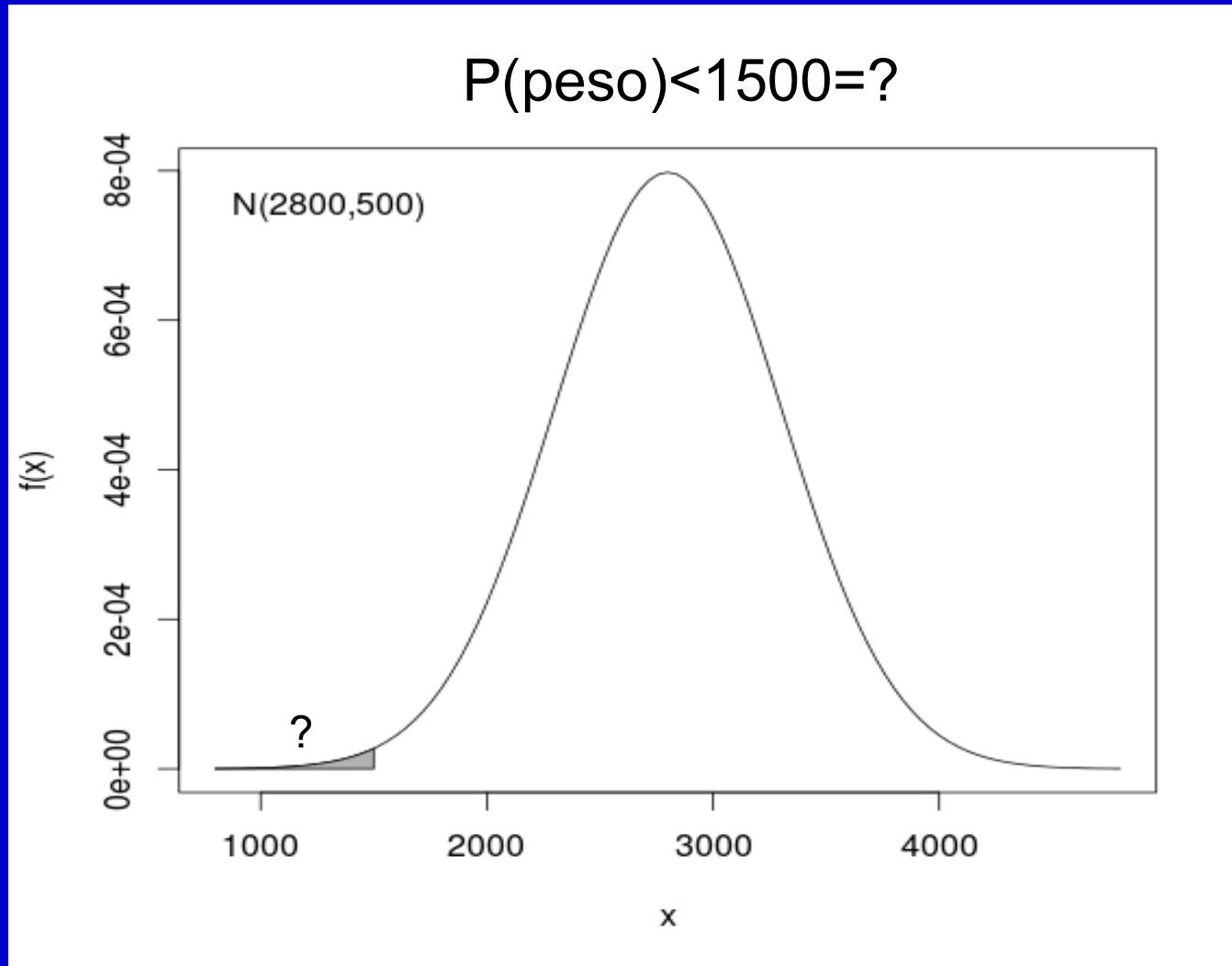


Exemplo: peso de recém-nascidos

$$P(1800 < \text{Peso} < 3800) = 95,4\%$$

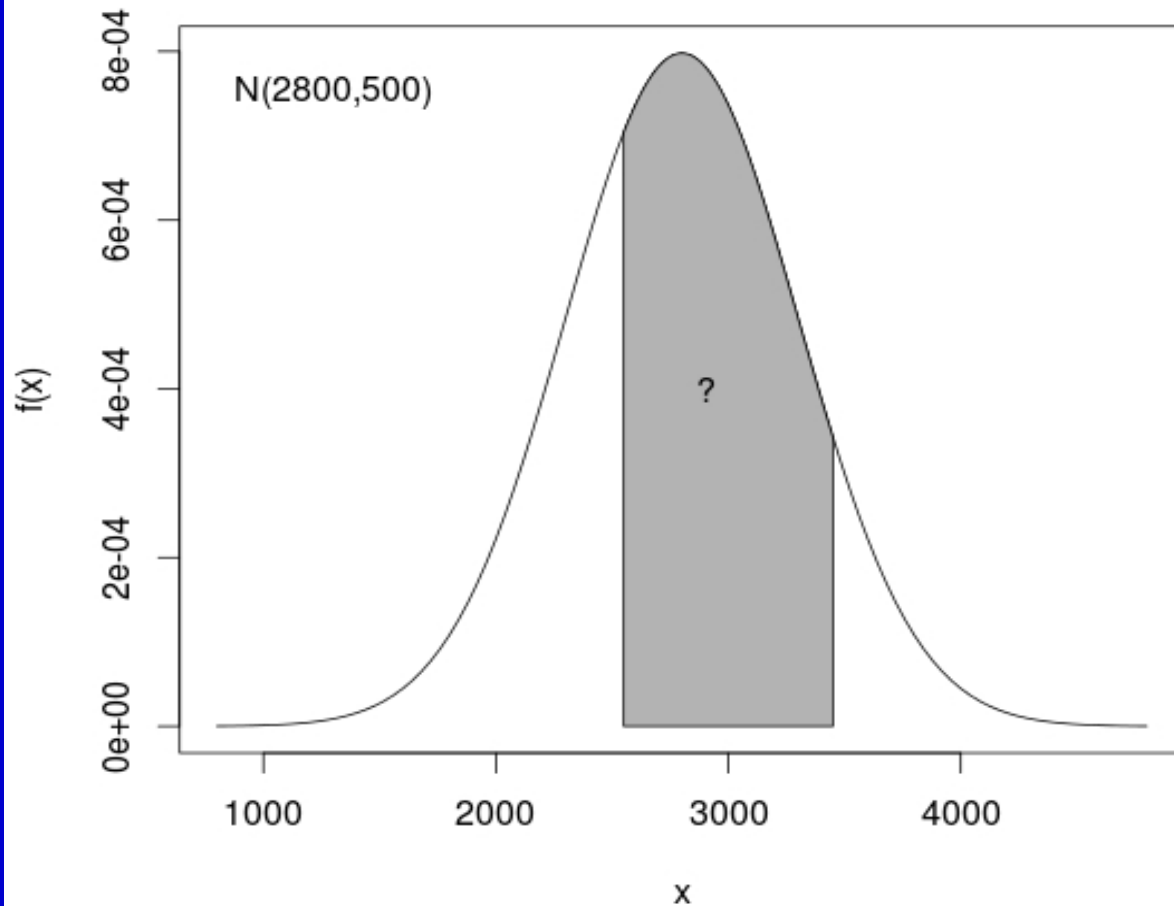


Exemplo: peso de recém-nascidos

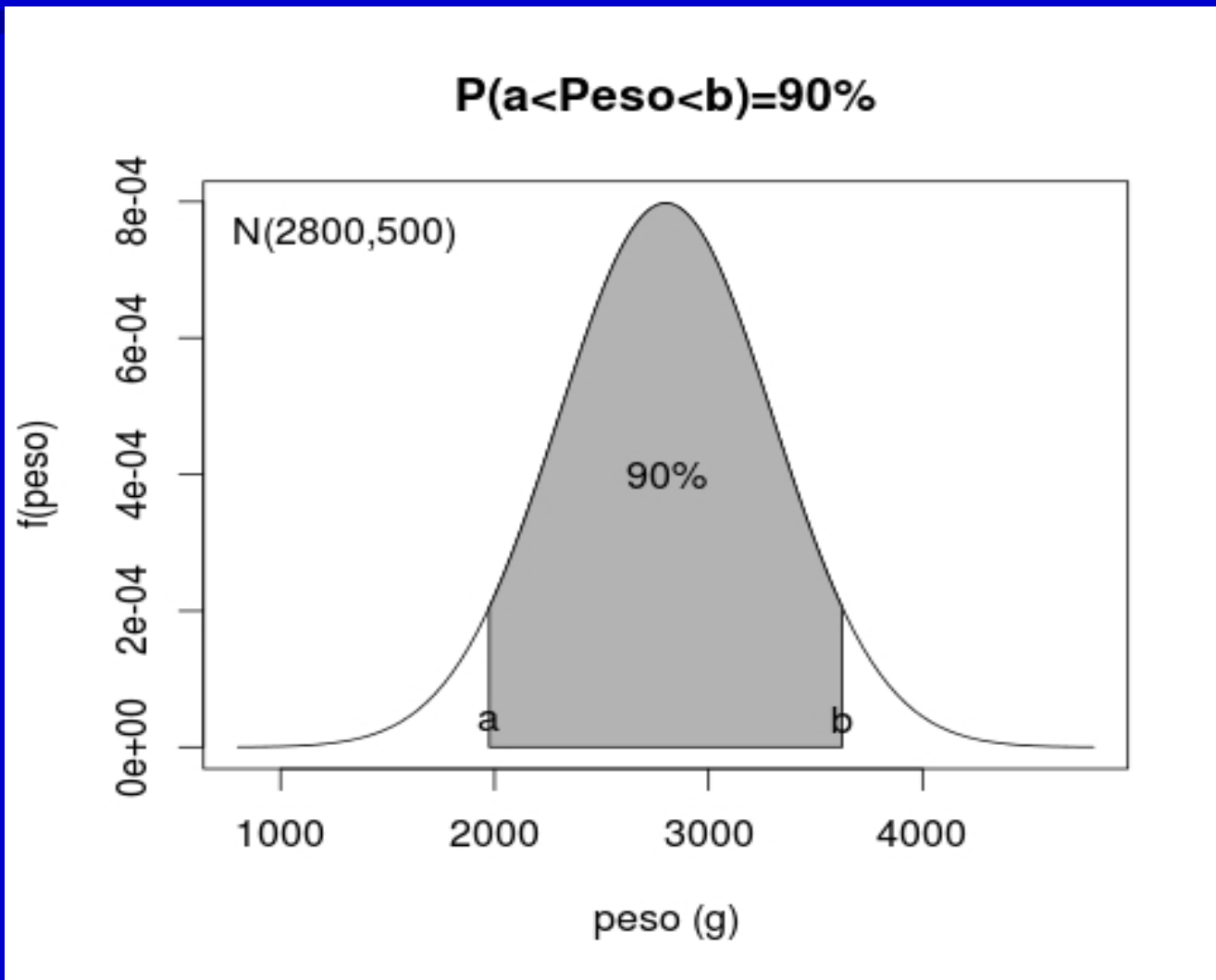


Exemplo: peso de recém-nascidos

$$P(2550 < \text{Peso} < 3450) = ?$$



Exemplo: peso de recém-nascidos



Padronização

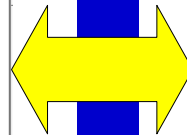
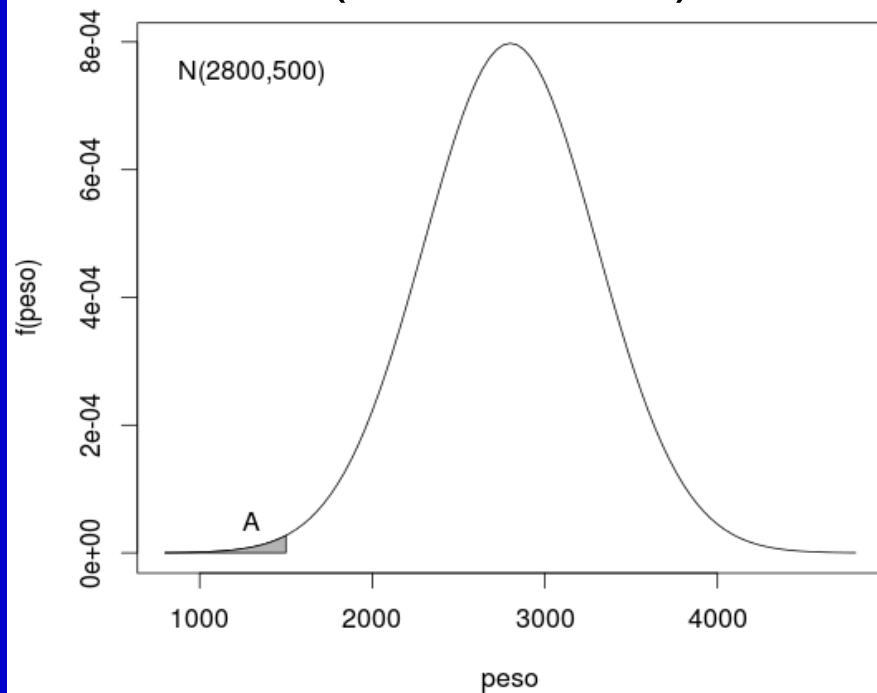
$X \sim N(\mu, \sigma)$ é transformada numa forma padronizada $Z \sim N(0, 1)$

$$Z = \frac{X - \mu}{\sigma}$$

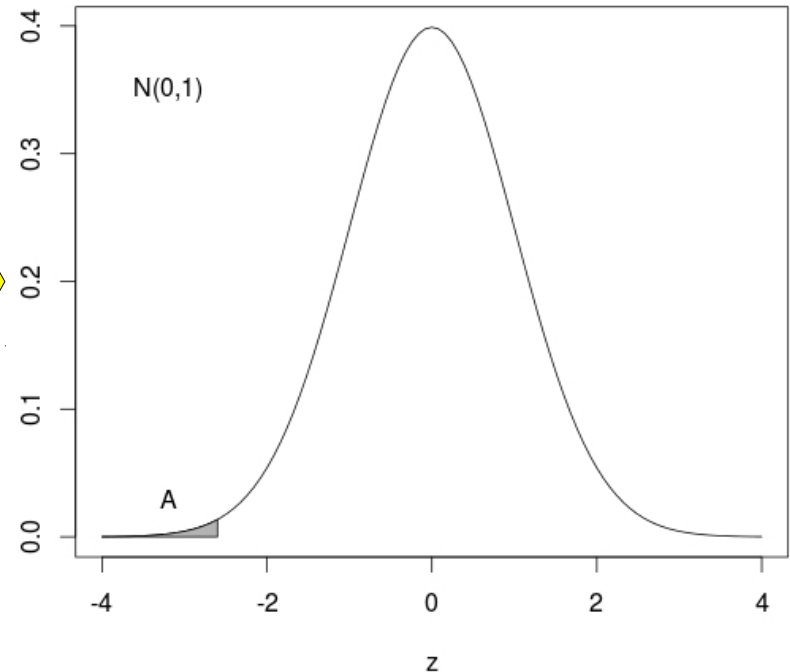
Padronização

Peso $\sim N(2800, 500)$ é transformado em $Z \sim N(0, 1)$

$$P(\text{Peso} < 1500) = A$$



$$P(Z < -2,6) = A$$

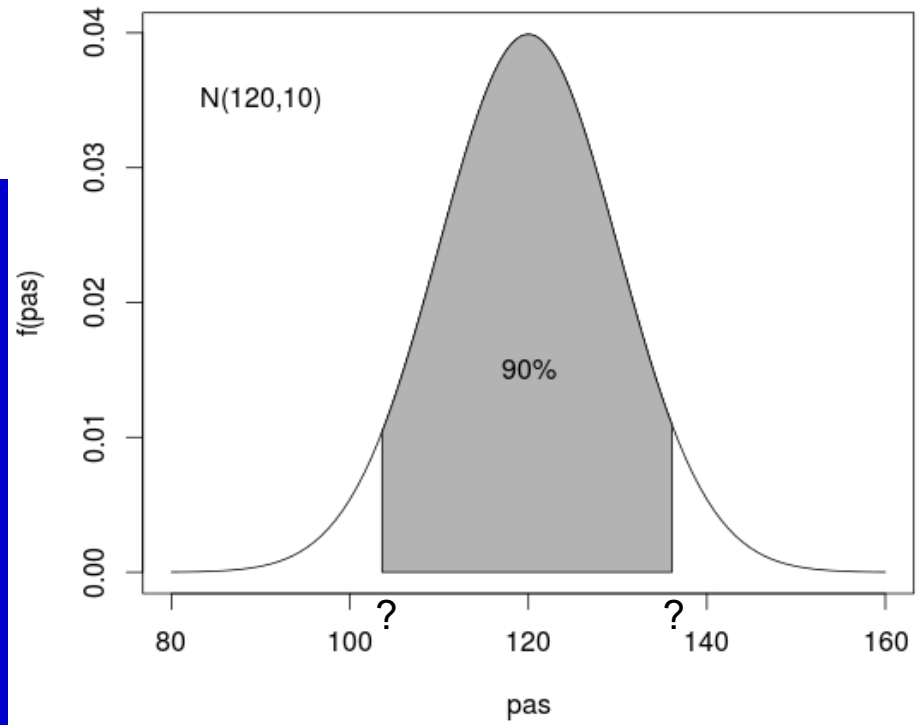
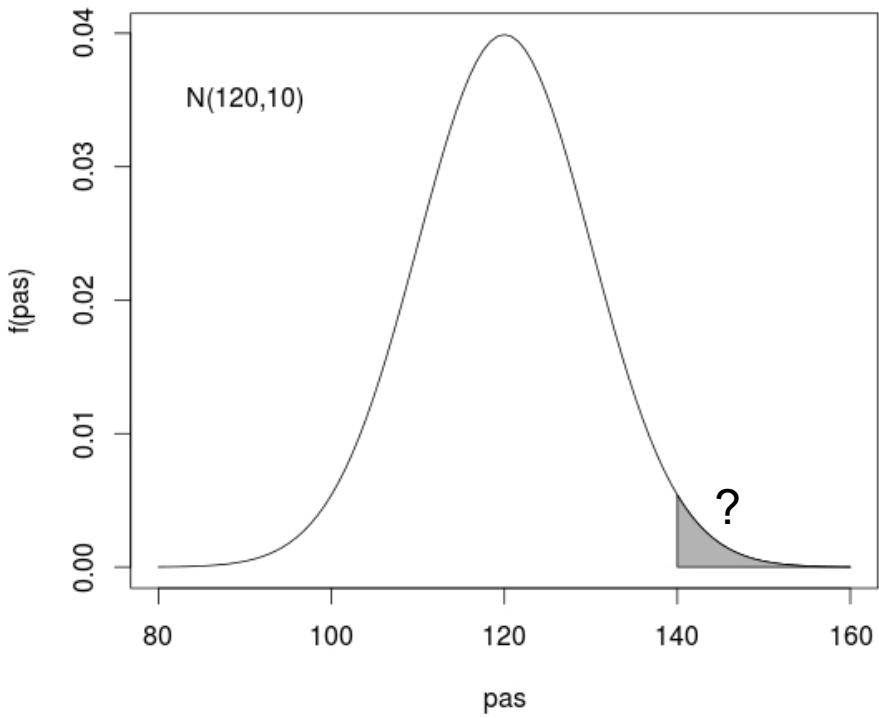


Exemplo: PAS

Suponha que a pressão arterial sistólica de pessoas jovens saudáveis seja $N(120,10)$

Qual é o percentual dessas pessoas com pressão sistólica acima de 140mmHg?

Qual é o intervalo simétrico em torno da média que engloba 90% dos valores das pressões sistólicas de pessoas jovens e saudáveis?



Calculadora

<http://onlinestatbook.com/2/calculators/normal.html>

Estadística Inferencial

Estimación, Intervalos de Confianza,
Testes de hipóteses



Estatística Inferencial

- Populações X Amostras
- Parâmetros X Estimativas
- Estimativas: Pontuais ou Intervalares
- Testes de Hipóteses



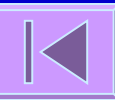
Teoria Elementar da Amostragem

- Teoria da amostragem
 - Retira informação sobre a **população** a partir de **amostras**
 - **Estimativas pontuais** ou **intervalares**
 - **Testes de Hipóteses**
- Números e amostras aleatórias
 - As **conclusões** da teoria de amostragem e da inferência estatística serão **válidas** se as amostras forem **representativas** da população
 - Um método para obter amostras representativas é a **amostragem aleatória simples**

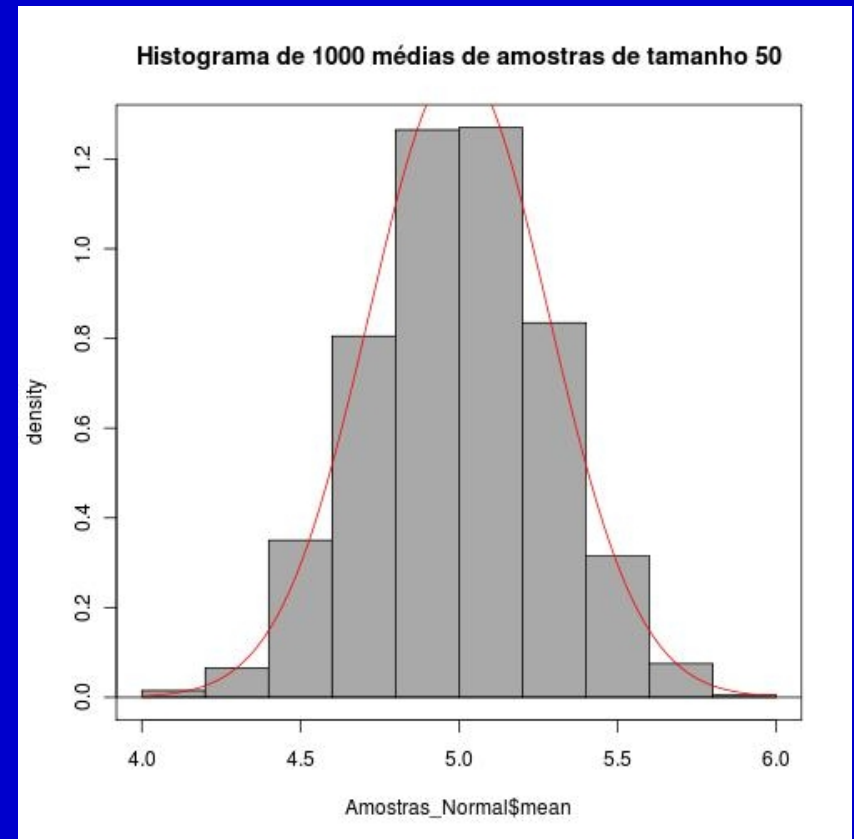
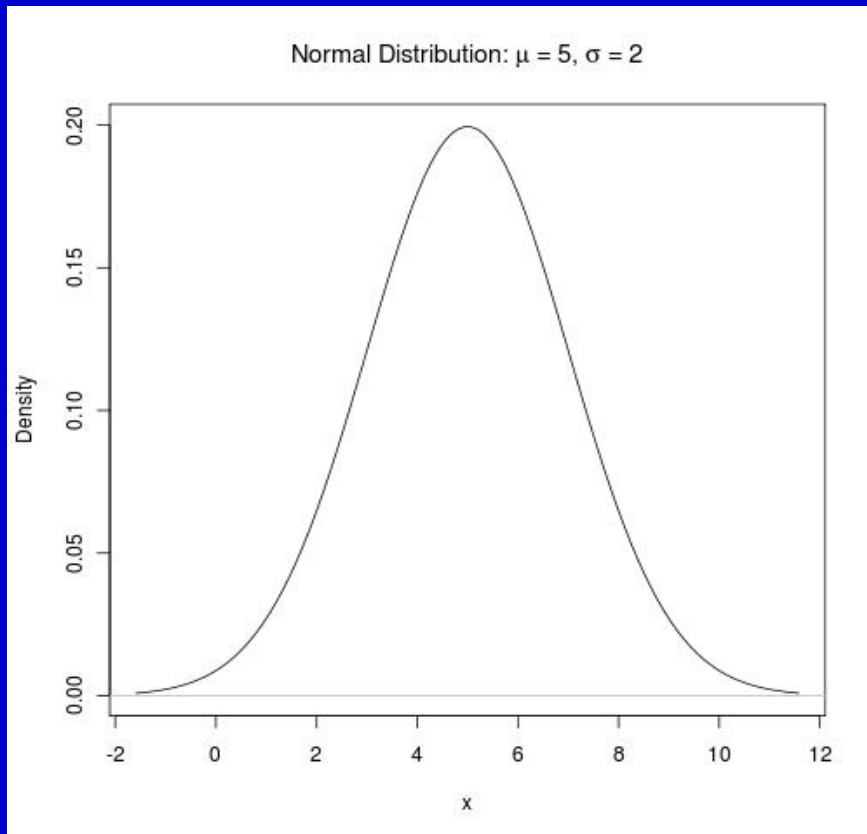


Teorema Central do Limite

- Valores estatísticos amostrais
 - Valores estatísticos obtidos de amostras são eles próprios variáveis
 - Assim, podem ser definidas distribuições a valores estatísticos amostrais
- Teorema central do limite
 - As **médias de amostras** de tamanho n retiradas de uma população normal **têm sempre uma distribuição normal**
 - As médias de amostras de tamanho n retiradas de uma população não normal têm uma distribuição que **tende para a normal à medida que n aumenta** (geralmente, a partir de $n \geq 30$ é já uma boa aproximação da normal)



Exemplo: TCL



Teorema Central do Limite (cont.)

- ▣ A distribuição das médias amostrais tende para uma distribuição **$N(\mu, \sigma/\sqrt{n})$**
- ▣ Erro Padrão
 - ▣ **Erro Padrão** é o desvio padrão das estatísticas amostrais
 - ▣ Assim, o **Erro Padrão da Média** $= \sigma/\sqrt{n}$ uma vez que é o desvio padrão das médias amostrais



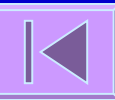
Teoria da Estimaco Paramtrica

▣ Estimaco Paramtrica

- ▣ Um dos problemas da estatística inferencial é a estimaco de parâmetros populacionais, também designada por **Estimaco Paramtrica**

▣ Estimaco

- ▣ **Pontual**
- ▣ **Intervalar**



Teoria da Estimação Paramétrica

- ▣ Intervalos de Confiança para parâmetros populacionais
- ▣ Intervalos de Confiança (IC) para a Média

$$\left(\bar{X} \pm z \frac{\sigma}{\sqrt{n}} \right)$$

- ▣ z é um valor da distribuição normal padrão
- ▣ No caso do IC 95% → $z = 1,96$
- ▣ No caso do IC 99% → $z = 2,58$



Intervalos de Confiança para a Média

▮ Interpretação

O intervalo $\mu \pm 1,96 (\sigma/\sqrt{n})$ contém 95% das possíveis médias amostrais, então, há uma probabilidade de 95% da média da nossa amostra estar dentro deste intervalo

Assim sendo, pode-se afirmar analogamente que 95% dos intervalos definidos por **Média amostral $\pm 1,96 (\sigma/\sqrt{n})$** cobrem a média da população (μ)

O intervalo **Média amostral $\pm 1,96 (\sigma/\sqrt{n})$** é chamado de **Intervalo de Confiança a 95% para a Média**



Distribuição t de Student e Teste de Hipóteses

Distribuição t de Student, Teste de Hipóteses, Teste t para uma média, teste t para a diferença entre duas médias e teste t para dados pareados



Distribuição t de Student

- ▣ Tendo em conta o Teorema Central do Limite, temos que:

$$\left(\frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \right) \sim N(0,1)$$

- ▣ Este resultado assume que σ é conhecido mas na prática não é.



Distribuição t de Student

- Para resolver este problema Gossett (1908), com o pseudônimo de Student, propõe uma distribuição que utiliza o desvio padrão da amostra ao invés do desvio padrão da população

$$t = \left(\frac{\bar{X} - \mu}{s / \sqrt{n}} \right)$$

- Se a variável em estudo segue uma distribuição normal, então t segue uma distribuição t de Student com n-1 graus de liberdade



Distribuição t de Student

- É semelhante à distribuição normal, mas com uma maior dispersão em torno do valor central
- Esta distribuição tem uma forma diferente em função do tamanho da amostra (n)
- À medida que n aumenta a distribuição tende para uma distribuição normal



Distribuição t de Student

- Assim, se não conhecermos o desvio padrão da população o **Intervalo de Confiança de 95% para a Média** poderá ser calculado do seguinte modo:

$$\left(\bar{X} \pm t_{(n-1; 0,05)} \frac{s}{\sqrt{n}} \right)$$

Distribuição t de Student

Intervalo de Confiança a 95% para a Média:

$$IC\ 95\% = \text{Média da amostra} \pm t_{(n-1)} (s/\sqrt{n})$$

Erro
Padrão

Valor apropriado da distribuição t com (n-1) graus de liberdade

Exemplo:

Estadística descritiva (n=462)

			Estadística	Erro Padrão
Peso da criança ao nascer	Média		3263,23	25,752
	Intervalo de confiança a 95% para a média	Limite inferior	3212,62	
		Limite superior	3313,83	

$$IC\ 95\% = 3263,23 \pm t_{(462-1)} (25,752)$$

$$IC\ 95\% = 3263,23 \pm 1,965 (25,752) = [3212,62; 3313,83]$$



Testes de Hipóteses

- Utilizando a mesma estrutura teórica que nos permite calcular Intervalos de Confiança podemos **testar hipóteses** sobre um parâmetro populacional

Exemplo: Queremos testar a hipótese de que a altura média de uma certa população é 160 cm. Numa amostra aleatória de 9 pessoas a altura média amostral foi 170 cm com desvio padrão amostral de 10 cm.

Qual é a probabilidade de se obter uma média amostral tão distante, ou ainda mais distante, da hipótese inicial de 160 cm?



Testes de Hipóteses

- Utilizando a mesma estrutura teórica que nos permite calcular Intervalos de Confiança podemos **testar hipóteses** sobre um parâmetro populacional

Exemplo: Queremos testar a hipótese de que a altura média de uma certa população é 160 cm. Numa amostra aleatória de 9 pessoas a altura média amostral foi 170 cm com desvio padrão amostral de 10 cm.

Qual é a probabilidade de se obter uma média amostral tão distante, ou ainda mais distante, da hipótese inicial de 160 cm?

Se essa probabilidade for muito baixa, podemos rejeitar a hipótese inicial.



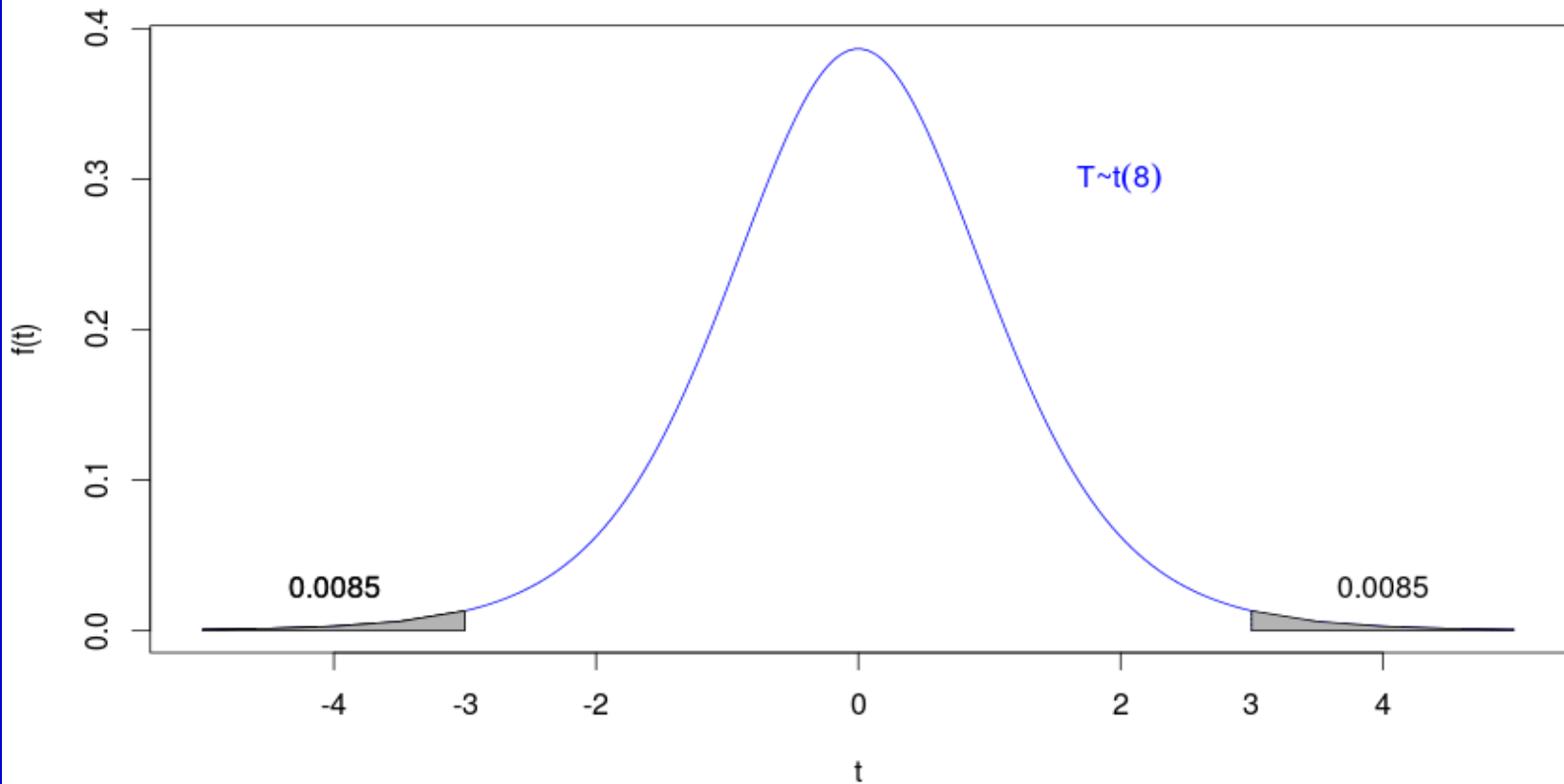
Teste t para uma média

$$H_0: \mu = 160 \text{ cm} \quad H_A: \mu \neq 160 \text{ cm}$$

$$n = 9 \quad \bar{X} = 170 \text{ cm} \quad s = 10 \text{ cm}$$

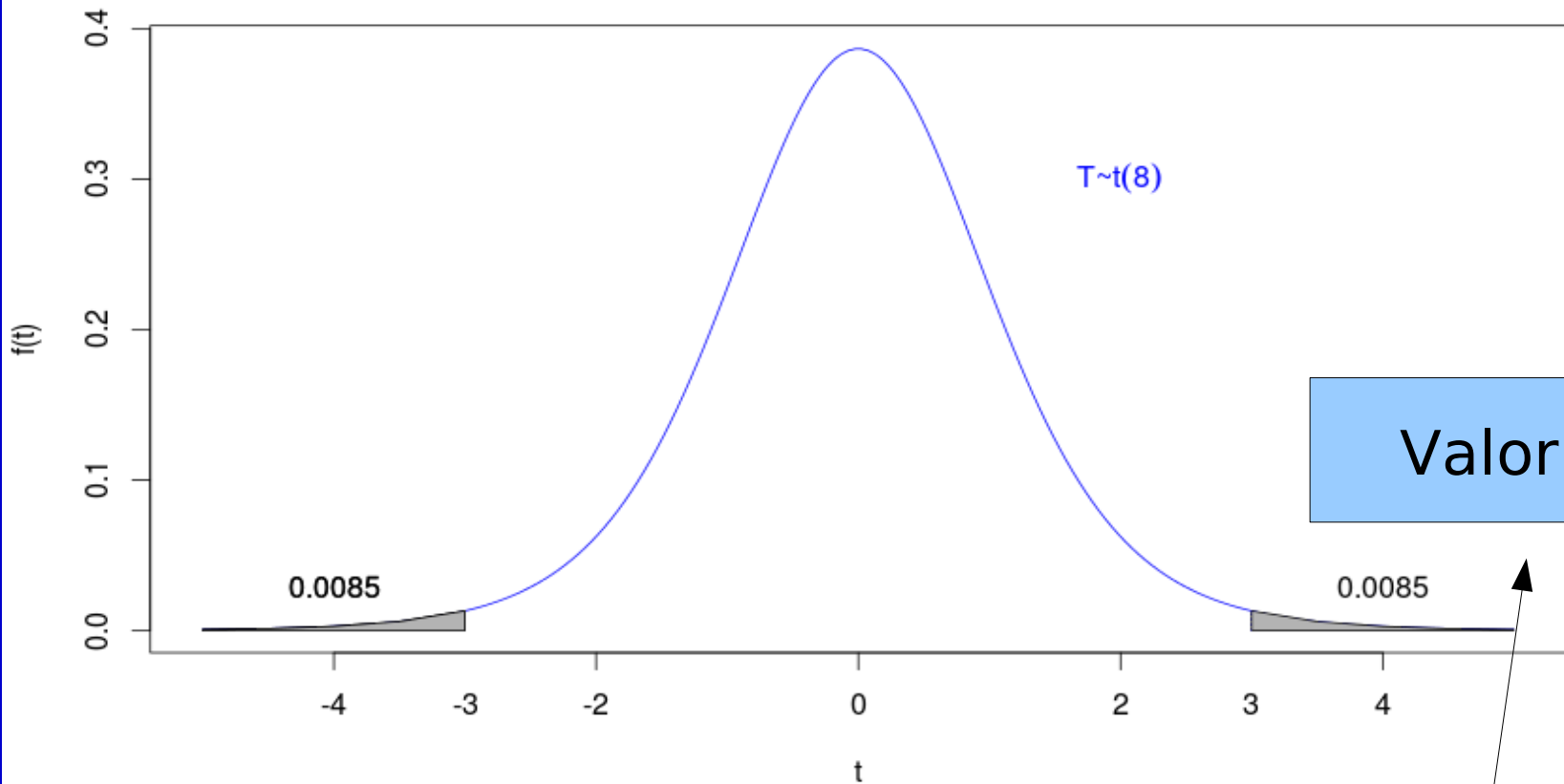
$$T = \left| \frac{\bar{X} - \mu}{s/\sqrt{n}} \right| = \left| \frac{170 - 160}{10/\sqrt{9}} \right| = 3 \sim t_{(9-1)} = t_8$$

Teste t para uma média



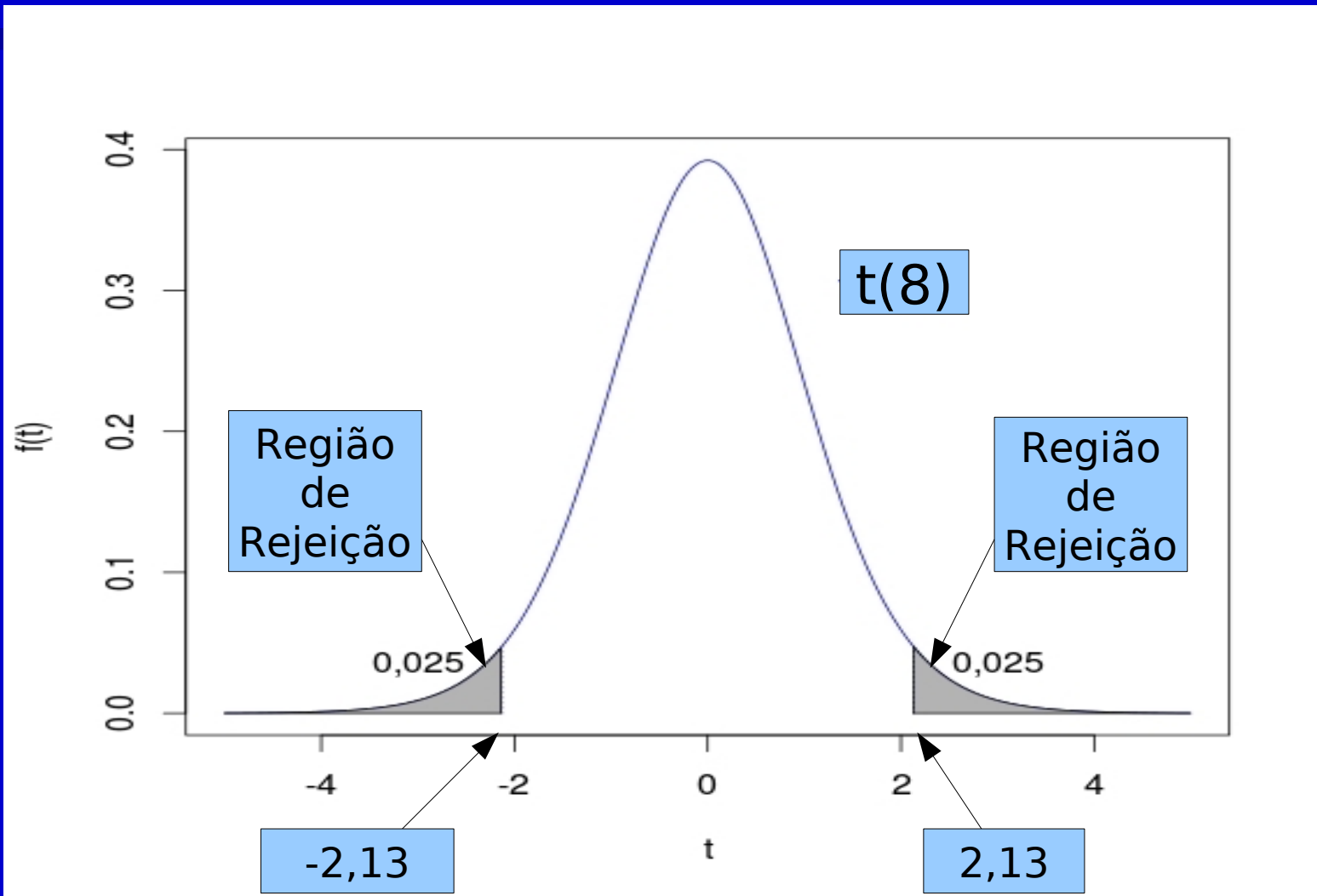
$$P(T < -3) + P(T > 3) = 2 \times 0,0085 = 0,017$$

Teste t para uma média



$$P(T < -3) + P(T > 3) = 2 \times 0,0085 = 0,017$$

Teste t para uma média



Teste t para uma média

□ Suposição:

- Distribuição normal ou aproximadamente normal da variável de interesse

Teste t para uma média

1. Especificar H_0 e H_A

$$H_0: \mu = \mu_0$$

$$H_A: \mu \neq \mu_0$$

2. Escolher o nível de significância ($\alpha = 5\%$)

3. Calcular a estatística de teste

$$T = \left| \frac{\bar{X} - \mu_0}{s/\sqrt{n}} \right|$$

4. Comparar o valor de T com uma distribuição de t com n-1 graus de liberdade

5. Calcular o valor de p e comparar com α

6. Descrever os resultados e conclusões estatísticas

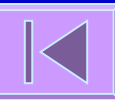
Tipos de Erros

- ▣ Erro tipo I (α)

Probabilidade de rejeitar a H_0
quando H_0 é verdadeira

- ▣ Erro tipo II (β)

Probabilidade de não rejeitar a H_0
quando H_0 é falsa



Exemplo:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Birthweight	462	3263,23	553,516	25,752

Valor de p

H0: $\mu = 3500$ g; HA: $\mu \neq 3500$ g

One-Sample Test

	Test Value = 3500					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Birthweight	-9,194	461	,000	-236,77	-287,38	-186,17

Teste t para a diferença entre duas médias

1. Especificar H_0 e H_A

$$H_0: \mu_1 = \mu_2 \quad H_A: \mu_1 \neq \mu_2$$

$$H_0: \mu_1 - \mu_2 = 0 \quad H_A: \mu_1 - \mu_2 \neq 0$$

2. Escolher o nível de significância ($\alpha = 0,05$ ou 5%)

3. Calcular a estatística e a estatística de teste

Média das duas amostras

$$t = [(\text{Média 1} - \text{Média 2}) - (\mu_1 - \mu_2)] / [s_{(\text{Média 1} - \text{Média 2})}]$$

4. Comparar o valor de t com uma distribuição de t com $(n_1 + n_2 - 2)$ graus de liberdade

5. Calcular o valor de p

6. Comparar p e α

7. Descrever os resultados e conclusões estatísticas

Teste t para a diferença entre duas médias

□ Suposições:

- Distribuição normal ou aproximadamente normal da variável nos dois grupos
- Independência entre os grupos



Exemplo:

Group Statistics

	Premature birth?	N	Mean	Std. Deviation	Std. Error Mean
Birthweight	No	401	3367,13	442,718	22,108
	Yes	59	2558,98	697,190	90,766

Valor de p

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Birthweight	Equal variances assumed	22,954	,000	12,014	458	,000	808,15	67,268	675,959	940,344
	Equal variances not assumed			8,651	65,053	,000	808,15	93,420	621,582	994,722

Teste t para a diferença entre duas médias

Group Statistics




	Sex of baby	N	Mean	Std. Deviation	Std. Error Mean
Birthweight	Male	250	3290,02	580,145	36,692
	Female	212	3231,63	519,954	35,711

Valor de p

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Birthweight	Equal variances assumed	1,265	,261	1,130	460	,259	58,39	51,663	-43,138	159,913
	Equal variances not assumed			1,140	458,577	,255	58,39	51,201	-42,229	159,005

Exemplo: Birthweight (cont.)

- Dados > Modificação de variáveis... > Converter variável numérica... 
- Estatísticas > Variâncias > Teste de Levene 
- Estatísticas > Médias > Teste t para amostras independentes 

Rcmdr: Convertendo variável numérica



Rcmdr: Teste de Levene



Rcmdr: Teste t para amostras independentes



Teste t para dados pareados

1. Especificar H_0 e H_A

$$H_0: \mu_d = 0 \quad H_A: \mu_d \neq 0$$

2. Escolher o nível de significância ($\alpha = 0,05$ ou 5%)

3. Calcular a estatística e a estatística de teste

Média das duas amostras

$$t = (\text{Média das diferenças} - \mu_d) / S_{(\text{diferenças})}$$

4. Comparar o valor de t com uma distribuição de t com (n-1) graus de liberdade

5. Calcular o valor de p

6. Comparar p e α

7. Descrever os resultados e conclusões estatísticas

Teste t para dados pareados

▮ Assume-se

- Distribuição normal ou aproximadamente normal das diferenças
- Dependência (correlação) entre os grupos

Teste t para dados pareados

Exemplo:

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Score na escala de depressão antes do tratamento	62,10	10	7,249	2,292
	Score na escala de depressão depois do tratamento	55,80	10	11,545	3,651

Valor de p

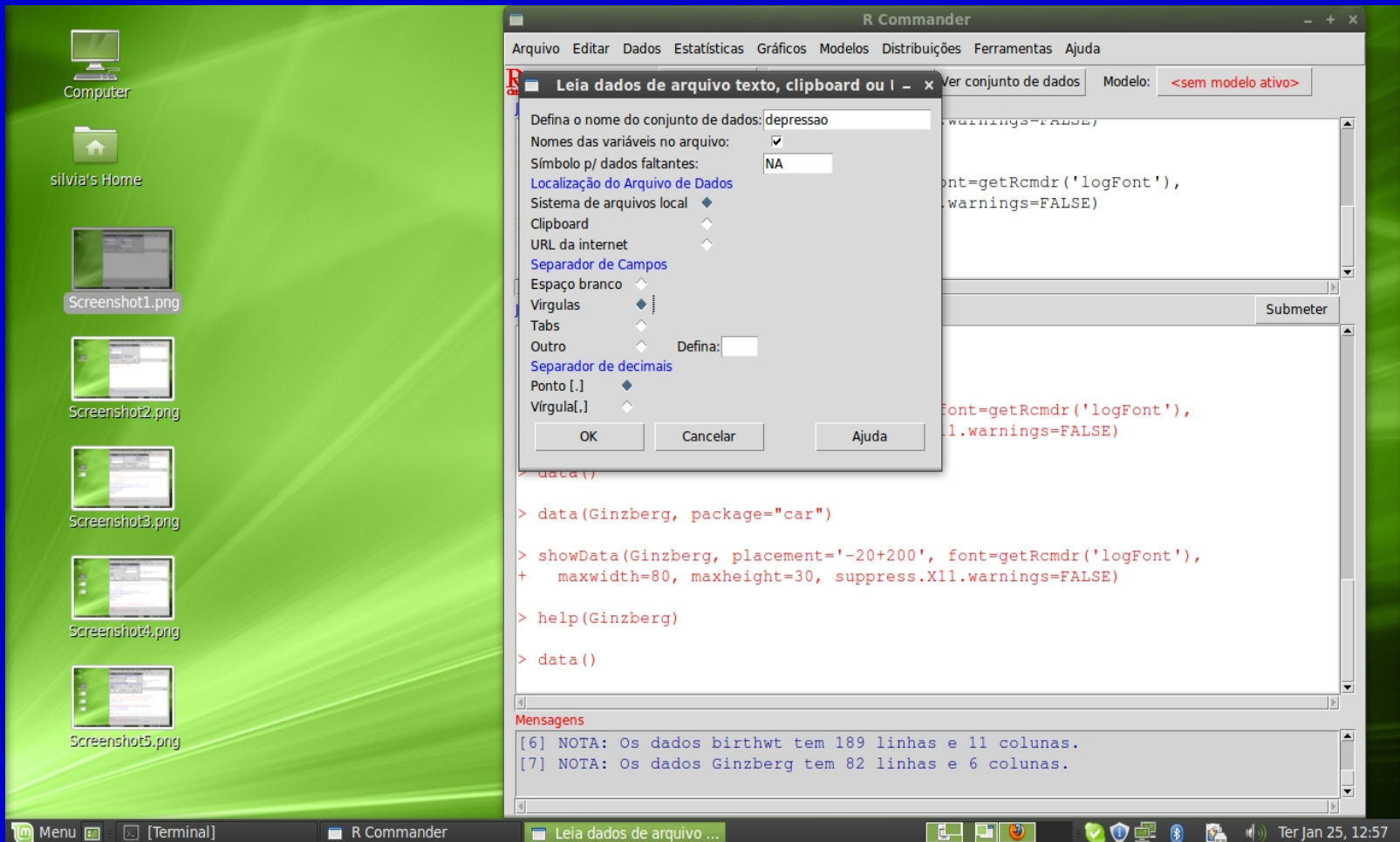
Paired Samples Test

		Paired Differences				t	df	Sig. (2-tailed)	
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower				Upper
Pair 1	Score na escala de depressão antes do tratamento - Score na escala de depressão depois do tratamento	6,30	9,298	2,940	-,35	12,95	2,143	9	,061

Exemplo: Escores de depressão

- ▮ Dados > Importar arquivos de dados > de arquivo texto...
- ▮ Estatísticas > Médias > Teste t (dados pareados)

Rcmdr: Lendo banco de dados de arquivo texto



The image shows a screenshot of the R Commander interface. A dialog box titled "Leia dados de arquivo texto, clipboard ou URL" is open, allowing the user to load data from a file, clipboard, or URL. The dialog box has the following fields and options:

- Nome do conjunto de dados:
- Localização do Arquivo de Dados: Local Clipboard URL da internet
- Sistema de arquivos local: Windows Linux
- Separador de Campos: Espaço branco Virgulas Tabs Outro (Defina:)
- Separador de decimais: Ponto [.] Virgual [.]

The main window shows the following R code and output:

```
> data()
> data(Ginzberg, package="car")
> showData(Ginzberg, placement='-20+200', font=getRcmdr('logFont'),
+   maxwidth=80, maxheight=30, suppress.X11.warnings=FALSE)
> help(Ginzberg)
> data()
```

Mensagens

```
[6] NOTA: Os dados birthwt tem 189 linhas e 11 colunas.
[7] NOTA: Os dados Ginzberg tem 82 linhas e 6 colunas.
```

Rcmdr: Teste t para dados pareados

The screenshot displays the R Commander interface. A dialog box titled "Teste-t pareado" is open, showing the configuration for a paired t-test. The first variable is "dia1" and the second is "dia42". The confidence level is set to 0.95. The alternative hypothesis is "Bilateral".

The console window shows the following R code and output:

```
> showdata(depressao, placement= 201200, font=getRcmdr('logFont',  
+ maxwidth=80, maxheight=30, suppress.X11.warnings=FALSE)  
> t.test(depressao$dia1, depressao$dia42, alternative='two.sided',  
+ conf.level=.95, paired=TRUE)
```

Paired t-test

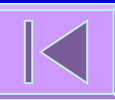
data: depressao\$dia1 and depressao\$dia42
t = 4.0702, df = 13, p-value = 0.001325
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
5.630646 18.369354
sample estimates:
mean of the differences
12

Mensagens

```
[7] NOTA: Os dados Ginzberg tem 82 linhas e 6 colunas.  
[8] NOTA: Os dados depressao tem 16 linhas e 2 colunas.
```

ANOVA

Análise de variância



ANOVA

- Comparação de médias de 2 grupos

Teste t

$$H_0: \mu_1 = \mu_2$$

$$\text{Erro tipo I } (\alpha) = 1 - 0,95 = 0,05$$

- Mais de 2 grupos:

$$\text{Ex: } H_0: \mu_1 = \mu_2 = \mu_3$$

$$(1) H_0: \mu_1 = \mu_2 \quad (2) H_0: \mu_1 = \mu_3 \quad (3) H_0: \mu_2 = \mu_3$$

$$\text{Erro tipo I} = 1 - 0,95^3 = 0,14$$

- Comparação de médias de mais de 2 grupos

ANOVA

$$H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$$

ANOVA

- Considere um conjunto de k grupos, com n_i indivíduos cada um, um total de n indivíduos, uma média de cada grupo \bar{x}_i e uma média comum \bar{X}

Ex: Considere os pesos (em kg) de 3 grupos de indivíduos de grupos étnicos diferentes (caucasianos, latinos e asiáticos).

Grupo 1: 80; 75; 82; 68; 76; 86; 78; 90; 85; 64 $\bar{x}_1 = 78,40$ kg

Grupo 2: 65; 84; 63; 54; 86; 62; 73; 64; 69; 81 $\bar{x}_2 = 70,10$ kg

Grupo 3: 58; 59; 61; 63; 71; 53; 54; 72; 61; 57 $\bar{x}_3 = 60,90$ kg

$\bar{X} = 69,80$ kg $k = 3$

$n_1 = 10$ $n_2 = 10$ $n_3 = 10$ $n = 30$

ANOVA

Fontes de variação:

- **Intra-grupos** - Variabilidade das observações em relação à média do grupo

▫ **Within group SS**
(sum of squares)

▫ **Within group DF**
(degrees of freedom)

▫ **Within group MS**
(mean square = variance)

$$\sum_{i=1}^k \left[\sum_{j=1}^{n_i} (x_{ij} - \bar{X}_i)^2 \right]$$

$$\sum_{i=1}^k (n_i - 1) = n - k$$

$$\frac{\text{Withingroup SS}}{\text{Withingroup DF}}$$

ANOVA

Fontes de variação:

- **Entre-grupos** - Variabilidade entre os grupos.
Dependente da média do grupo em relação à média conjunta

Between group SS

$$\sum_{i=1}^k n_i (\bar{X}_i - \bar{X})^2$$

Between group DF

$$k-1$$

Between group MS

$$\frac{\text{Between group SS}}{\text{Between group DF}}$$

ANOVA

- A variabilidade observada num conjunto de dados deve-se a:
 - Variação em relação à média do grupo - Within group MS
 - Variação da média do grupo em relação à média comum - Between group MS

ANOVA

- Prova-se que se $\mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$, então, Between MS e Within MS serão ambas estimativas de σ^2 - a variância comum aos k grupos - logo, Between MS \approx Within MS
- Se pelo contrário $\mu_1 \neq \mu_2 \neq \mu_3 \neq \dots \neq \mu_k$, então, Between MS será maior que Within MS
- Assim, para testar a Hipótese nula
H₀: $\mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$ calcula-se a estatística F

$$F = \frac{\text{Between group MS}}{\text{Within group MS}}$$

ANOVA

- A estatística F tem uma distribuição teórica conhecida - Distribuição F - dependente dos graus de liberdade Between DF e Within DF
- O cálculo da estatística F e seu enquadramento na distribuição adequada permite-nos conhecer um valor de p - probabilidade de obter um F tão ou mais extremo que o calculado se a hipótese nula for verdadeira
- O valor de p é subsequentemente comparado com o grau de significância (α) à partida estabelecido e
 - **Se $p \leq \alpha$, rejeita-se a $H_0 \Rightarrow$ Existem diferenças estatisticamente significativas entre as médias dos grupos**
 - **Se $p > \alpha$, aceita-se a $H_0 \Rightarrow$ Não existem diferenças estatisticamente significativas entre as médias dos grupos**

ANOVA

- ▮ Suposições:
 - Normalidade
 - Igualdade das variâncias dos grupos
- ▮ Funciona melhor se:
 - Igual tamanho dos grupos
 - Igualdade dos grupos exceto na variável de interesse

Exemplo:

Descriptives

Peso do indivíduo (Kg)

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Caucasiano	10	78,40	8,06	2,55	72,64	84,16	64	90
Latino	10	70,10	10,61	3,35	62,51	77,69	54	86
Asiático	10	60,90	6,38	2,02	56,33	65,47	53	72
Total	30	69,80	10,98	2,00	65,70	73,90	53	90

Test of Homogeneity of Variances

Peso do indivíduo (Kg)

Levene	df1	df2	Sig.
Statistic			
1,862	2	27	,175

ANOVA

Valor de p

ANOVA

Peso do indivíduo (Kg)

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1532,600	2	766,300	10,534	,000
Within Groups	1964,200	27	72,748		
Total	3496,800	29			

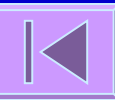
Exemplo: Peso x raça

- Crie banco de dados do exemplo acima numa planilha e salve como txt
- Converter grupo em fator
- Realizar teste de Levene
- Fazer a Anova

peso	grupo
80	1
75	1
82	1
68	1
76	1
86	1
78	1
90	1
85	1
64	1
65	2
84	2
63	2
54	2
86	2
62	2
73	2

Testes Não Paramétricos

Mann-Whitney Test; Wilcoxon
Signed Ranks Test; Kruskal-
Wallis Test



Mann-Whitney Test

- Análogo ao teste t para a diferença entre duas médias
- Quando as condições necessárias para a utilização do teste t não são cumpridas (normalidade e igualdade de variâncias) tem que se optar pelos testes análogos não paramétricos
- Não faz condições sobre a distribuição da variável
- Faz uso das posições ordenadas dos dados (ranks) e não dos valores da variável obtidos

Mann-Whitney Test

- **EX:** Para investigar se os mecanismos envolvidos nos ataques fatais de asma provocados por alergia à soja são diferentes dos mecanismos envolvidos nos ataques fatais de asma típica compararam-se o número de células T CD3+ na submucosa de indivíduos destes dois grupos.



Mann-Whitney Test

- Ex: situações possíveis (dois grupos A e B de 5 elementos cada um):

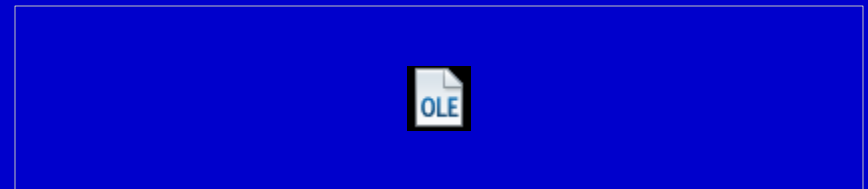
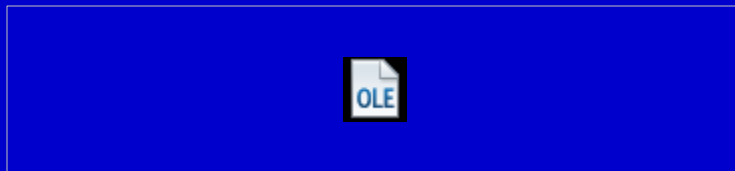
A A A A A B B B B
1º 2º 3º 4º 5º 6º 7º 8º 9º 10º

A e B diferentes

A B A B A B A B A B
1º 2º 3º 4º 5º 6º 7º 8º 9º 10º

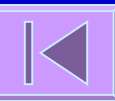
Não há diferenças entre A e B

- São calculadas as seguintes estatísticas:



R_1 = soma das posições no grupo 1

R_2 = soma das posições no grupo 2



Mann-Whitney Test

- A maior destas estatísticas é comparada com uma distribuição adequada (distribuição da estatística U ou aproximação normal)
- Obtem-se um valor de p - probabilidade de se obter uma estatística tão ou mais extrema do que a verificada caso a hipótese nula seja verdadeira
- O valor de p é subsequentemente comparado com o grau de significância (α) à partida estabelecido e
 - **Se $p \leq \alpha$, rejeita-se a $H_0 \Rightarrow$ Existem diferenças estatisticamente significativas relativamente à distribuição da variável entre os grupos**
 - **Se $p > \alpha$, aceita-se a $H_0 \Rightarrow$ Não existem diferenças estatisticamente significativas relativamente à distribuição da variável entre os grupos**

Mann-Whitney Test

Exemplo:

Ranks				
	Grupo	N	Mean Rank	Sum of Ranks
Número de células T CD3+ na submucosa (células/mm ²)	Grupo de alergia à soja	7	4,57	32,00
	Grupo de asma típica	10	12,10	121,00
	Total	17		

Valor de p

Test Statistics ^b	
	Número de células T CD3+ na submucosa (células/mm ²)
Mann-Whitney U	4,000
Wilcoxon W	32,000
Z	-3,033
Asymp. Sig. (2-tailed)	,002
Exact Sig. [2*(1-tailed Sig.)]	,001 ^a

a. Not corrected for ties.
b. Grouping Variable: Grupo



Wilcoxon Signed Ranks Test

- ▮ Análogo do teste t para pares emparelhados ou teste t para a diferença entre 2 médias de grupos dependentes
- ▮ **Ex:** Num ensaio de um fármaco antidepressivo obtêm-se os seguintes scores numa escala de depressão, antes e depois do tratamento:



Wilcoxon Signed Ranks Test

- ▣ Posicionam-se os valores absolutos das diferenças de forma ascendente e atribui-se o sinal da diferença à posição
- ▣ Calculam-se as seguintes estatísticas:
 T_+ = soma das posições com sinal positivo
 T_- = soma das posições com sinal negativo
- ▣ Utiliza-se a menor destas estatísticas, sendo esta comparada com uma distribuição adequada (distribuição da estatística T ou aproximação normal)



Wilcoxon Signed Ranks Test

- Obtem-se um valor de p - probabilidade de se obter uma estatística tão ou mais extrema do que a verificada caso a hipótese nula seja verdadeira
- O valor de p é subsequentemente comparado com o grau de significância (α) à partida estabelecido e
 - Se $p \leq \alpha$, rejeita-se a $H_0 \Rightarrow$ Existem diferenças estatisticamente significativas relativamente à distribuição da variável entre os grupos
 - Se $p > \alpha$, aceita-se a $H_0 \Rightarrow$ Não existem diferenças estatisticamente significativas relativamente à distribuição da variável entre os grupos

Wilcoxon Signed Ranks Test

Exemplo:

Ranks				
		N	Mean Rank	Sum of Ranks
Score na escala de depressão depois do tratamento - Score na escala de depressão antes do tratamento	Negative Ranks	7 ^a	6,43	45,00
	Positive Ranks	3 ^b	3,33	10,00
	Ties	0 ^c		
	Total	10		

a. Score na escala de depressão depois do tratamento < Score na escala de depressão antes do tratamento

b. Score na escala de depressão depois do tratamento > Score na escala de depressão antes do tratamento

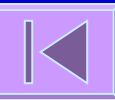
c. Score na escala de depressão antes do tratamento = Score na escala de depressão depois do tratamento

Valor de p

Test Statistics ^b	
7	Score na escala de depressão depois do tratamento - Score na escala de depressão antes do tratamento
	-1,786 ^a
Asymp. Sig. (2-tailed)	,074

a. Based on positive ranks.

b. Wilcoxon Signed Ranks Test



Kruskal-Wallis Test

- Análogo da Análise de Variância (ANOVA) para a comparação das médias de 3 ou mais grupos
- Ex: Pesos em Kg de 3 grupos de indivíduos de grupos étnicos diferentes (caucasianos, latinos e asiáticos).

Grupo 1: 80; 75; 82; 68; 76; 86; 78; 90; 85; 64

Grupo 2: 65; 84; 63; 54; 86; 62; 73; 64; 69; 81

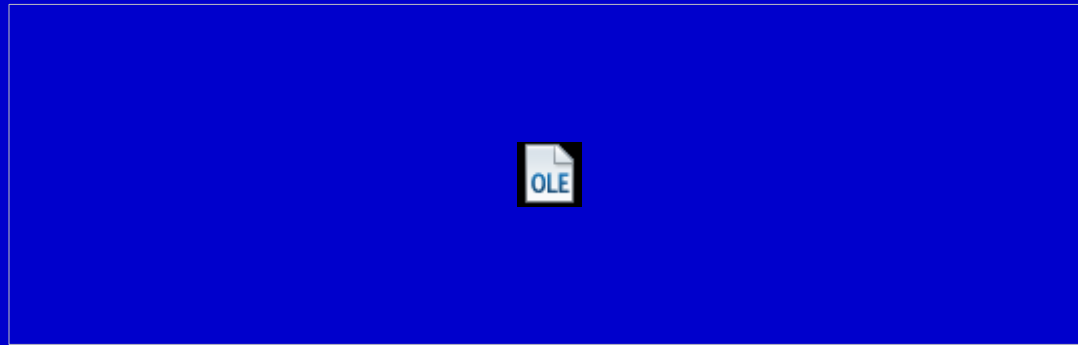
Grupo 3: 58; 59; 61; 63; 71; 53; 54; 72; 61; 57

Organizam-se todos os valores por ordem crescente de modo a cada valor ter uma posição atribuída



Kruskal-Wallis Test

- Calcula-se a estatística:



- \mathbf{N} = nº total de indivíduos; \mathbf{n}_i = nº de indivíduos no grupo i e \mathbf{R}_i = soma das posições no grupo i
- Esta estatística será comparada com uma distribuição adequada (distribuição de Qui-quadrado com $k-1$ graus de liberdade)



Kruskal-Wallis Test

- Obtem-se um valor de p - probabilidade de se obter uma estatística tão ou mais extrema do que a verificada caso a hipótese nula seja verdadeira
- O valor de p é subsequentemente comparado com o grau de significância (α) à partida estabelecido e
 - Se $p \leq \alpha$, rejeita-se a $H_0 \Rightarrow$ **Existem diferenças estatisticamente significativas relativamente à distribuição da variável entre os grupos**
 - Se $p > \alpha$, aceita-se a $H_0 \Rightarrow$ **Não existem diferenças estatisticamente significativas relativamente à distribuição da variável entre os grupos**

Kruskal-Wallis Test

Exemplo:

Ranks			
	Grupo étnico	N	Mean Rank
Peso do indivíduo (Kg)	Caucasiano	10	22,40
	Latino	10	16,20
	Asiático	10	7,90
	Total	30	

Valor de p

Test Statistics ^{a,b}	
	Peso do indivíduo (Kg)
Chi-Square	13,675
df	2
Asymp. Sig.	,001

a. Kruskal Wallis Test
b. Grouping Variable: Grupo étnico



Tabelas de Contingência e Teste Qui-quadrado

Tabelas de contingência; teste qui-quadrado; teste exato de Fisher; correção de Yates; teste de McNemar; teste qui-quadrado para tendências



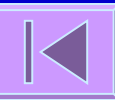
Tabelas de Contingência

- Forma de representar a relação entre duas variáveis categóricas. Distribuição das frequências das categorias de uma variável em função das categorias de uma outra variável.

		Race of Respondent				
		White	Black	Other	Total	
Region of the United States	North East	Count	582	82	15	679
		% within Region of the United States	85,7%	12,1%	2,2%	100,0%
		% within Race of Respondent	46,0%	40,2%	30,6%	44,8%
		% of Total	38,4%	5,4%	1,0%	44,8%
	South East	Count	307	94	14	415
		% within Region of the United States	74,0%	22,7%	3,4%	100,0%
		% within Race of Respondent	24,3%	46,1%	28,6%	27,4%
		% of Total	20,2%	6,2%	,9%	27,4%
	West	Count	375	28	20	423
		% within Region of the United States	88,7%	6,6%	4,7%	100,0%
		% within Race of Respondent	29,7%	13,7%	40,8%	27,9%
		% of Total	24,7%	1,8%	1,3%	27,9%
Total	Count	1264	204	49	1517	
	% within Region of the United States	83,3%	13,4%	3,2%	100,0%	
	% within Race of Respondent	100,0%	100,0%	100,0%	100,0%	
	% of Total	83,3%	13,4%	3,2%	100,0%	

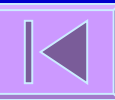
Teste Qui-quadrado

- Quando estamos perante duas variáveis categóricas podemos usar o teste qui-quadrado para testar a hipótese da existência de uma associação entre as variáveis na população.
- As hipóteses nula e alternativa que serão testadas são:
 - H_0 : Não existe uma associação entre as categorias de uma variável e as da outra variável na população ou as proporções de indivíduos nas categorias de uma variável não variam em função das categorias da outra variável na população
 - H_A : Existe uma associação entre as categorias de uma variável e as da outra variável na população ou as proporções de indivíduos nas categorias de uma variável variam em função das categorias da outra variável na população



Teste Qui-quadrado

- Podem-se apresentar os dados numa tabela de contingência $r \times c$ (r - nº de linhas; c - nº de colunas). As entradas da tabela são frequências e cada célula contém o nº de indivíduos que pertencem simultaneamente àquela linha e coluna.
- Calcula-se as frequências esperadas caso a hipótese nula fosse verdadeira. A frequência esperada numa determinada célula é o produto do total da linha e do total da coluna dividido pelo total global.
- Baseada na estatística de teste (χ^2): discrepância entre as **frequências observadas** e as **frequências esperadas**, caso a H_0 seja verdadeira, em cada célula da tabela. Se a discrepância for grande é improvável que a hipótese nula seja verdadeira.



Teste Qui-quadrado

- A estatística de teste calculada (χ^2) tem a seguinte forma genérica:



O - frequência observada na célula e E - frequência esperada na célula, caso a H_0 seja verdadeira.

- A tabela de contingência tem a seguinte forma genérica:



Teste Qui-quadrado

- A estatística de teste segue a Distribuição de Qui-quadrado com $(r-1) \times (c-1)$ graus de liberdade.
- O cálculo da estatística χ^2 e seu enquadramento na distribuição adequada permite-nos conhecer um valor de p (probabilidade de obter um χ^2 tão ou mais extremo que o calculado se a hipótese nula for verdadeira)
- O valor de p é comparado com o grau de significância (α):
 - ▣ **Se $p \leq \alpha$, rejeita-se a H_0 =>** Existe uma associação entre as categorias de uma variável e as da outra variável na população **ou** as proporções de indivíduos nas categorias de uma variável variam em função das categorias da outra variável na população
 - ▣ **Se $p > \alpha$, não rejeita-se a H_0 =>** Não existe evidência suficiente de uma associação entre as categorias de uma variável e as da outra variável na população

Teste Qui-quadrado

- Ex: Num ensaio clínico compara-se a eficácia de um Medicamento X (n=30 indivíduos) em relação ao placebo (n=32 indivíduos) na melhoria do estado clínico dos doentes 6 meses após o tratamento (melhorado, agravado, falecido).

Estado clínico 6 meses após o tratamento * Tratamento efectuado Crosstabulation

		Tratamento efectuado			Total
		Placebo	Medicamento X		
Estado clínico 6 meses após o tratamento	Melhorado	Count	9	17	26
		Expected Count	13,4	12,6	26,0
	Agravado	Count	12	9	21
		Expected Count	10,8	10,2	21,0
	Falecido	Count	11	4	15
		Expected Count	7,7	7,3	15,0
Total	Count	32	30	62	
	Expected Count	32,0	30,0	62,0	

$$E_{11} = (26 \cdot 32) / 62 = 13,4$$

$$E_{12} = (26 \cdot 30) / 62 = 12,6$$

$$E_{21} = (21 \cdot 32) / 62 = 10,8$$

$$E_{22} = (21 \cdot 30) / 62 = 10,2$$

$$E_{31} = (15 \cdot 32) / 62 = 7,7$$

$$E_{32} = (15 \cdot 30) / 62 = 7,3$$

Teste Qui-quadrado

Ex: (continuação)

Valor de p

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	6,099 ^a	2	,047
Likelihood Ratio	6,264	2	,044
Linear-by-Linear Association	5,947	1	,015
N of Valid Cases	62		

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 7,26.



Teste Qui-quadrado

- ▮ $p = 0,047$ Logo, $p < \alpha \Rightarrow$ Rejeita-se a H_0 .
- ▮ Existem uma associação entre o estado clínico 6 meses após o tratamento (melhorado, agravado, falecido) e o tipo de tratamento efectuado (placebo ou medicamento X) **ou** Existem diferenças estatisticamente significativas quanto ao estado clínico 6 meses após o tratamento entre



Teste Qui-quadrado

□ Assume-se:

- **Independência dos grupos**

Caso as variáveis em análise sejam dependentes deverá ser usado o **Teste de McNemar**.

- **Pelo menos 80% das frequências esperadas têm valores ≥ 5**

No caso de existirem mais de 20% de células com valores esperados < 5 deve **reduzir-se a tabela**, através da fusão de colunas ou linhas (esta fusão deve fazer sentido no contexto da análise que está a ser feita), até ter pelo menos 80% das frequências esperadas com valor ≥ 5 .

Se numa tabela de 2×2 (corresponde à fusão máxima possível) existir uma ou mais frequências esperadas com valor < 5 , então deverá ser usado o **Teste Exato de Fisher**.

Teste Qui-quadrado

- Teste Exato usado em tabelas de 2×2 (faz o cálculo das probabilidades exatas e não faz uso da distribuição de qui-quadrado como aproximação para o cálculo de probabilidades).
- Utiliza-se no caso de uma tabela de contingência de 2×2 , uma ou mais frequências esperadas < 5 .
- Ex: num outro ensaio clínico comparou-se a mortalidade no grupo tratado com placebo e tratado com o medicamento X e obtiveram-se os seguintes resultados:

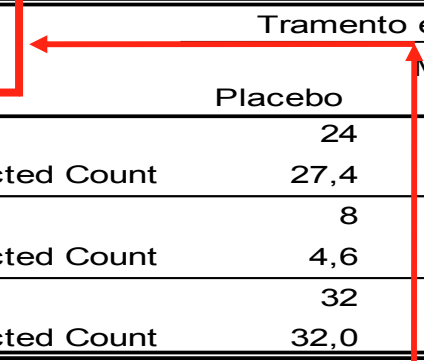


Teste Exato de Fisher

Mortalidade 6 meses após o tratamento * Tratamento efectuado Crosstabulation

		Tratamento efectuado			
		Placebo	Medicamento X	Total	
Mortalidade 6 meses após o tratamento	Vivo	Count	24	29	53
		Expected Count	27,4	25,6	53,0
	Morto	Count	8	1	9
		Expected Count	4,6	4,4	9,0
Total	Count	32	30	62	
	Expected Count	32,0	30,0	62,0	

Valor de p



Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	5,858 ^b	1	,016		
Continuity Correction ^a	4,242	1	,039		
Likelihood Ratio	6,606	1	,010		
Fisher's Exact Test				,027	,017
Linear-by-Linear Association	5,763	1	,016		
N of Valid Cases	62				



a. Computed only for a 2x2 table

b. 2 cells (50,0%) have expected count less than 5. The minimum expected count is 4,35.



Correção de Yates

- Correção para a continuidade em tabelas de 2x2:



Valor de p

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	5,858 ^D	1	,016		
Continuity Correction ^a	4,242	1	,039		
Likelihood Ratio	6,606	1	,010		
Fisher's Exact Test				,027	,017
Linear-by-Linear Association	5,763	1	,016		
N of Valid Cases	62				

a. Computed only for a 2x2 table

b. 2 cells (50,0%) have expected count less than 5. The minimum expected count is 4,35.



Teste de McNemar

- Análogo ao teste qui-quadrado mas para variáveis dependentes.



Teste de McNemar

Ex:

Tosse antes do tratamento * Tosse depois do tratamento Crosstabulation

		Tosse depois do tratamento			
		Ausente	Presente	Total	
Tosse antes do tratamento	Ausente	Count	44	0	44
		Expected Count	34,8	9,2	44,0
	Presente	Count	5	13	18
		Expected Count	14,2	3,8	18,0
Total	Count	49	13	62	
	Expected Count	49,0	13,0	62,0	

Valor de p

Chi-Square Tests

	Value	Exact Sig. (2-sided)
McNemar Test		,063 ^a
N of Valid Cases	62	

a. Binomial distribution used.



Teste Qui-quadrado para Tendências

Ex:

		Grupo etário * Estado clínico 6 meses após o tratamento Crosstabulation				
		Estado clínico 6 meses após o tratamento				
		Melhorado	Agravado	Falecido	Total	
Grupo etário	20-35 anos	Count	14	4	3	21
		Expected Count	9,5	6,0	5,5	21,0
		% within Grupo etário	66,7%	19,0%	14,3%	100,0%
	36-50 anos	Count	13	6	3	22
		Expected Count	9,9	6,3	5,8	22,0
		% within Grupo etário	59,1%	27,3%	13,6%	100,0%
	51-65 anos	Count	6	7	7	20
		Expected Count	9,0	5,8	5,3	20,0
		% within Grupo etário	30,0%	35,0%	35,0%	100,0%
	>65 anos	Count	3	6	8	17
		Expected Count	7,7	4,9	4,5	17,0
		% within Grupo etário	17,6%	35,3%	47,1%	100,0%
Total		Count	36	23	21	80
		Expected Count	36,0	23,0	21,0	80,0
		% within Grupo etário	45,0%	28,8%	26,3%	100,0%



Teste Qui-quadrado para Tendências

Valor de p

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	14,083 ^a	6	,029
Likelihood Ratio	14,681	6	,023
Linear-by-Linear Association	12,144	1	,000
N of Valid Cases	80		

a. 2 cells (16,7%) have expected count less than 5. The minimum expected count is 4,46.



Testes Qui-quadrado no R

- ▣ `chisq.test()`
- ▣ `fisher.test()`
- ▣ `mcnemar.test()`
- ▣ `prop.trend.test()`

Quadros de Síntese

Estatística; testes de hipóteses; testes de hipóteses para variáveis quantitativas; testes de hipóteses para variáveis categóricas; outros métodos

