6 Comparando Dois Grupos

ESQUEMA DO CAPÍTULO

6.1 Introdução

A ***comparação*** entre técnicas atuais e métodos alternativos está presente em ***diversas*** áreas do conhecimento humano, como, por exemplo:

- Na agricultura, onde são buscadas variedades mais adequadas e produtivas;
- Na engenharia, onde produtos e processos inovadores, mais eficientes e eficazes são desenvolvidos;
- Na área da saúde, onde procuram-se por drogas mais seguras, com maior poder de cura e tolerabilidade.

Entretanto, como há ***variabilidade*** nos dados obtidos de tais comparações, ***métodos estatísticos*** precisam ser usados

6.2 Conceitos fundamentais

Teste de Hipóteses:

Este procedimento estatístico é usado ***amplamente*** em situações em que as variáveis de interesse possuem variabilidade;

Exemplo 6.1: Eficácia do AZT

Os dados centrais de um estudo que comprovou a eficácia de zidovudina (AZT) no tratamento da AIDS são apresentados na Tabela 6.1, em uma tabela 2 × 2 na ***forma padrão***.

Tabela 6.1: Número de sobreviventes tratados com AZT e placebo

Grupo	Situação				Total	
	Vivo		Morto			
AZT	$a=$	144	$b=$	1	$n_1=a+b=$	145
Placebo	$c=$	121	$d=$	16	$n_2=c+d=$	137
Total	$m_1=a+c=$	265	$m_2=b+d=$	17	$N=n_1+n_2=$	282

6.2 Conceitos fundamentais (cont.)

Exemplo 6.1: Eficácia do AZT (cont.)

A análise da Tabela 6.1 consiste basicamente em comparar ***duas*** proporções:

- A proporção de sobreviventes no grupo de ***controle***:

$$p_C = \frac{121}{137} \cong 0{,}883$$

- A proporção de sobreviventes no grupo ***tratamento***:

$$p_T = \frac{144}{145} \cong 0{,}993$$

A diferença observada nessas proporções é ***significativa***?

6.2 Conceitos fundamentais (cont.)

Hipóteses a serem testadas:

1. ***Hipótese nula***, H_0, que é fixada como a ***inexistência de diferenças*** (igualdade) entre os dois tratamentos comparados; no Exemplo 6.1, a hipótese nula é:

$$H_0 : p_C = p_T$$

2. ***Hipótese alternativa***, H_1, que usualmente é a ***inexistência de igualdade*** (diferença) entre os tratamentos; no Exemplo 6.1, a hipótese alternativa será:

$$H_1 : p_C \neq p_T$$

6.2 Conceitos fundamentais (cont.)

Critério de decisão entre as hipótese H_0 e H_1:

- A ***Teoria Estatística*** já desenvolveu critérios, para um grande número de situações;
- O critério de decisão é baseado numa ***estatística de teste***, que mede a discrepância entre o que foi observado na amostra e o que seria esperado se a hipótese nula H_0 fosse verdadeira;
- Se a estatística de teste é "grande", ***rejeita-se*** H_0 em favor de H_1;
- Caso contrário, ***não*** se rejeita H_0.

6.2 Conceitos fundamentais (cont.)

Erros do tipo I e II, nível de significância e poder do teste:

- Em um teste de hipóteses, há ***erros*** possíveis associados a cada decisão tomada, conforme sintetizado na Tabela 6.2;
- A probabilidade do erro tipo I é denominada α, ou ***nível de significância***;
- A probabilidade do erro tipo II é denominada β;

Tabela 6.2: Erros possíveis associados a um teste de hipóteses

Conclusão do teste	Situação real	
	H_0 verdadeira	H_0 falsa
Não rejeitar H_0	decisão correta	erro tipo II
Rejeitar H_0	erro tipo I	decisão correta

6.2 Conceitos fundamentais (cont.)

Erros do tipo I e II, nível de significância e poder do teste (cont.):

- Para um mesmo tamanho de amostra, é ***impossível*** controlar (manter baixos) os dois erros, tipo I e II;
- Convencionou-se que o erro tipo I é o mais ***sério***;
- Assim, os critérios de decisão devem manter α sob ***controle*** (valores baixos, por exemplo, α = 0,10, 0,05 ou 0,01);
- Em um ***segundo*** momento, calcula-se o tamanho da amostra para controlar (manter baixo) também o β;
- A capacidade de um teste identifica diferenças que realmente existem é denominada ***poder do teste***, (1-β).

6.2 Conceitos fundamentais (cont.)

Probabilidade de significância (valor-p):

- Duas formas de se chegar a uma conclusão final sobre um teste de hipóteses:
  1. Compara-se o valor da ***estatística de teste*** com o valor obtido a partir da sua ***distribuição teórica***, para um valor prefixado (por exemplo, 10%, 5% ou 1%);
  2. Utiliza-se a ***probabilidade de significância***, ou ***valor-p***, que é calculado sob H_0 verdadeiro; quanto menor o valor-p, maior é a evidência para rejeitar H_0;

Obs.: Em pesquisas na área de saúde, um valor-p menor ou igual a 0,05 indica que há diferenças significativas entre os grupos).

6.3 Resposta dicotômica: amostras independentes

- A variável de interesse é dita dicotômica, pois só pode assumir ***dois*** valores possíveis: a ocorrência ou a não ocorrência do evento de interesse;
- O problema é ***comparar*** as probabilidades de sua ocorrência do evento de interesse, em dois grupos estudados (p_1 e p_2), o que pode ser formulado através das seguintes ***hipóteses estatísticas***:

$$\begin{cases} H_0 : p_1 = p_2 \\ H_1 : p_1 \neq p_2 \end{cases}$$

6.3 Resposta dicotômica: amostras independentes (cont.)

Teste qui-quadrado (χ^2)

A Tabela 6.3 apresenta dados genéricos, na ***forma padrão***, de uma situação envolvendo a comparação de dois grupos e que a resposta de interesse é ***dicotômica***: a ocorrência ou não de um evento.

Tabela 6.3: Distribuição quanto à ocorrência de um evento

	Ocorrência da evento		
Grupo	Sim	Não	Total
I	a	b	$n1 = a + b$
II	c	d	$n_2 = c + d$
Total	$m_1 = a + c$	$m_2 = b + d$	$N = n_1 + n_2$

Se ***não há diferenças*** entre as proporções nos dois grupos, então vale:

$$\frac{a}{n_1} = \frac{c}{n_2} = \frac{a+c}{n_1+n_2} = \frac{m_1}{N} \Rightarrow a = \frac{m_1 \times n_1}{N},\ b = \frac{m_2 \times n_1}{N},\ c = \frac{m_1 \times n_2}{N} \text{ e } d = \frac{m_2 \times n_2}{N}$$

6.3 Resposta dicotômica: amostras independentes (cont.)

Teste qui-quadrado (χ^2) (cont.)

Tem-se portanto dois conjuntos de valores:

- os valores realmente ***observados*** (O_i) e
- os valores que seriam ***esperados*** (E_i), se o H_0 fosse verdadeiro, que são calculados a partir das equações anteriores;

Pearson propôs medir a discrepância entre os esses dois valores, O_i e E_i, através da seguinte ***estatística*** (expressão):

$$X^2 = \sum_{i=1}^{4} \frac{\left(O_i - E_i\right)^2}{E_i}$$

Na ***prática***, o valor de X^2 pode ser calculado através da expressão:

$$X^2 = \frac{N\left(ad - bc\right)^2}{m_1 m_2 n_1 n_2}$$

6.3 Resposta dicotômica: amostras independentes (cont.)

Exemplo 6.2: Eficácia do AZT (cont.)

Supondo que o AZT e o placebo fossem equivalentes (H_0 verdadeiro), os valores ***esperados*** são mostrados na Tabela 6.4.

Tabela 6.4: Valores esperados de sobreviventes sob hipótese de equivalência entre o AZT e placebo

	Situação		
Grupo	Vivo	Morto	Total
AZT	136,26	8,74	145
Placebo	128,74	8,26	137
Total	265	17	282

A Tabela 6.5 mostra os ***cálculos*** necessários para determinar a medida de discrepância de Pearson, ou seja, $X^2 = 15{,}01$.

Tabela 6.5: Cálculos necessários para a construção do teste χ^2

i	O_i	Ei	$O_i - E_i$	$(O_i - E_i)^2$	$(O_i - E_i)^2/E_i$
1	144	136,26	7,74	59,91	0,44
2	121	128,74	-7,74	59,91	0,47
3	1	8,74	-7,74	59,91	6,85
4	16	8,26	7,74	59,91	7,25
Total	282	282	0	239,63	15,01

6.3 Resposta dicotômica: amostras independentes (cont.)

Exemplo 6.2: Eficácia do AZT (cont.)

O valor de X^2 pode ser obtido mais ***facilmente*** através da expressão abaixo e da Tabela 6.1:

$$X^2 = \frac{N(ad-bc)^2}{m_1 m_2 n_1 n_2} = \frac{282\times(144\times16-1\times121)^2}{265\times17\times145\times137} \cong 15{,}01$$

Agora, é preciso ***decidir*** se $X^2 = 15{,}01$ é ou não um valor "grande". Considera-se que X^2 é grande se ocorrer $X^2 > \chi^2_{1;\alpha}$, em que $\chi^2_{1,\alpha}$ é o valor da distribuição qui-quadrado, com 1 grau de liberdade (Tabela A4), em que $\Pr(X \geq \chi^2_{1;\alpha}) = \alpha$.

Tabela A4: Distribuição do χ^2: $\Pr(X \geq x)$

Graus de liberdade	$\Pr(X \geq x)$						
	0,750	0,250	0,100	0,050	0,025	0,010	0,005
1	0,102	1,323	2,706	3,841	5,024	6,635	7,879

Assim, para um ***nível de significância*** de $\alpha = 0{,}05$ (5%), rejeita-se H_0, pois $X^2 = 15{,}01$ supera $\chi^2_{1;0,05} = 3{,}841$ (veja Tabela A4), decidindo-se em favor do efeito do AZT, com um ***nível de certeza*** de $1-\alpha = 0{,}95$ (95%).

6.3 Resposta dicotômica: amostras independentes (cont.)

Exemplo 6.2: Eficácia do AZT (cont.)

Outra forma de decidir sobre a eficácia ou não do AZT é através da ***probabilidade de significância***, ou valor-p, usualmente obtida em ***softwares estatísticos***. Veja a seguir exemplo no R:

- Carregue os dados:

```
dados<-matrix(c(144,121,1,16),nc=2)
```

- Rode o teste qui-quadrado:

```
chisq.test(x, correct=FALSE)
```

- Resultado:

```
Pearson's Chi-squared test
data:  dados
X-squared = 15.017, df = 1, p-value = 0.0001066
```

Como o valor-p é menor que 0,05, conclui-se pela ***eficácia*** do AZT ao nível de significância de 5%.

6.3 Resposta dicotômica: amostras independentes (cont.)

Exemplo 6.3: Fatores de risco para AVC

A Tabela 6.6 apresenta as ***proporções*** de presença de alguns fatores de risco nos indivíduos que sofreram AVC e naqueles que não foram acometidos pela doença. Os ***valores-p*** tem interpretação conforme descrito anteriormente.

Tabela 6.6: Percentuais e valor-p para a comparação do grupo que sofreu com o que não sofreu AVC

	AVC		
Fator	Sim	Nâo	Valor-p
AVC como causa principal de morte			
na mãe	29,8	11,2	**0,0005**
no pai	7,0	7,5	0,72
AVC como causa principal ou contribuindo para a causa da morte			
na mãe	29,8	14,2	**0,002**
no pai	7,0	9,2	0,59
Doença coronariana	7,0	6,3	0,83
Sinais eletrocardiográficos			
hipertrofia do ventrículo esquerdo	3,5	1,0	0,08
Doença coronariana	10,5	6,5	0,23

6.3 Resposta dicotômica: amostras independentes (cont.)

Teste qui-quadrado com correção de continuidade

Yates (1934) demonstrou que o valor obtido pela expressão abaixo é mais apropriado para comparação com o χ^2, conhecida como ***correção de continuidade*** ou correção de Yates:

$$X^2 = \frac{N\left(|ad - bc| - \frac{N}{2}\right)^2}{m_1 m_2 n_1 n_2}$$

6.3 Resposta dicotômica: amostras independentes (cont.)

Exemplo 6.4: Contraceptivos orais e infarto do miocárdio
Em um estudo sobre a associação entre o uso corrente de contraceptivos e o infarto do miocárdio, Shapiro et al. (1979) observaram os resultados entre pacientes com idade entre 30 e 34 anos, mostrados na Tabela 6.7.

Tabela 6.7: Distribuição de uso de contraceptivo oral segundo grupo que sofreu ou não infarto do miocárdio

	Uso recente		
Grupo	Sim	Não	Total
Casos	9	12	21
Controles	33	390	423
Total	42	402	444

A proporção de uso recente entre os casos é $p_T = 9/21 \approx 0{,}43$ e entre os controles $p_C = 33/423 \approx 0{,}08$, que parece ser menor. Para determinar que essa diferença não ocorreu por acaso, calcula X^2 com correção de continuidade, que dá 24,76, o que ***confirma*** a associação, com um valor-p < 0,001:

```
dados<-matrix(c(9,33,12,390),nc=2)
chisq.test(dados,correct=TRUE)
X-squared = 24.76, df = 1, p-value = 6.494e-07 (< 0,05)
```

6.3 Resposta dicotômica: amostras independentes (cont.)

Teste exato de Fisher

- Há dificuldade técnica na aplicação do teste qui-quadrado, com ou sem correção de Yates, quando o valor esperado em alguma ***casela*** na tabela 2×2 é ***menor*** que 5.
- A ***alternativa*** é usar o teste exato de Fisher, disponível em *softwares* estatísticos, como o R.

6.3 Resposta dicotômica: amostras independentes (cont.)

Exemplo 6.5: Tratamento de pneumonia bacteriana
Pastenark et al. (1992) avaliaram a eficácia e a segurança de dois antibióticos. A avaliação da resposta terapêutica foi baseada na evolução do quadro clínico e está apresentada na Tabela 6.8. Note que há caselas na tabela 2 × 2 com contagens ***inferiores*** a 5.

Tabela 6.8: Percentuais de curas e de efeitos colaterais do cefadroxil e da cefalexina

Quadro	Cefadroxil	Cefalexina	Total
Cura	31	28	59
Efeitos adversos	**1**	**3**	4
Total	32	31	63

Para comparar os percentuais de cura e de ocorrência de efeitos adversos, obtiveram-se os resultados abaixo, pelo teste exato de Fisher, que ***não rejeita*** a hipótese H_0 de igualdade entre as proporções e, consequentemente, ***não confirma*** que há diferenças entre as proporções de cura e de efeitos adversos nos dois antibióticos.

```
dados<-matrix(c(31,1,28,3),nc=2)
fisher.test(dados,alternative="two.sided",conf.level=0.95)
Fisher's Exact Test for Count Data
p-value = 0.3547 (> 0,05)
```

6.3 Resposta dicotômica: amostras independentes (cont.)

Teste Z para comparação de proporções

- É um teste alternativo aproximado para o qui-quadrado, que requer ***amostras grandes***;
- Sejam p_1 e p_2 as proporções de sucesso referentes aos tratamentos a ser comparados, estimado por $\hat{p}_1$ e $\hat{p}_2$, baseadas em amostras de tamanhos n_1 e n_2, respectivamente;
- Além disso, $n_1\hat{p}_1$ e $n_2\hat{p}_2$ devem ***exceder 5***;
- É um teste baseado na estatística:

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\dfrac{\hat{p}_1\left(1-\hat{p}_1\right)}{n_1} + \dfrac{\hat{p}_2\left(1-\hat{p}_2\right)}{n_2}}}$$

6.3 Resposta dicotômica: amostras independentes (cont.)

Teste Z para comparação de proporções (cont.)

- Alternativamente, pode-se ***estimar*** a probabilidade de sucesso nas duas amostras combinadas, denotada por $\hat{p}$ e calculada como

$$\hat{p} = \frac{m_1 + m_2}{n_1 + n_2}$$

onde m_1 e m_2 são respectivamente o número de sucessos de cada amostra;

- Nesse caso, calculamos:

$$Z^* = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\left(1-\hat{p}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

- O valor de Z (ou Z^*) deve ser comparado ao percentil de ordem $(1-\alpha/2)$ da distribuição ***gaussiana***, denotado $Z_{\alpha/2}$.

6.3 Resposta dicotômica: amostras independentes (cont.)

Exemplo 6.6: Comparação de drogas contra náusea

Com o objetivo de comparar a ***eficácia*** de dois preventivos contra a náusea, 400 marinheiros foram divididos em dois grupos de 200. Um grupo recebeu a pílula A e o outro a B. Os resultados são apresentados na tabela abaixo. Há indicações de que a eficácia das pílulas A e B é a mesma?

	Náusea		
Pílula	Não	Sim	Total
A	152	48	200
B	132	68	200
Total	284	116	400

Temos que:

$$\hat{p}_A = \frac{152}{200} \cong 0,76, \ \hat{p}_B = \frac{132}{200} \cong 0,66 \text{ e } \hat{p} = \frac{152+132}{200+200} \cong 0,71$$

6.3 Resposta dicotômica: amostras independentes (cont.)

Exemplo 6.6: Comparação de drogas contra náusea (cont.)

Calculando-se os valores das duas ***estatísticas***, Z e Z^*, os valores são quase iguais:

$$Z = \frac{0,76-0,66}{\sqrt{\frac{0,76(1-0,76)}{200}+\frac{0,66(1-0,66)}{200}}} \cong 2,22$$

$$Z^* = \frac{0,76-0,66}{\sqrt{0,71(1-0,71)\left(\frac{1}{200}+\frac{1}{200}\right)}} \cong 2,20$$

Fixando-se um nível de significância de 5% ($\alpha = 0,05$), tem-se que $Z_{\alpha/2} = Z_{0,025} \approx$ 1,96. Como Z e Z^* são maiores que $Z_{0,05}$, rejeita-se o H_0 (as pílulas têm eficácias ***diferentes***). Alternativamente, os valores-p são, respectivamente, 0,026 e 0,028, o que conduz à mesma conclusão de rejeitar H_0 (valor-p $< 0,05$).

6.4 Resposta dicotômica: amostras pareadas

O exemplo a seguir ilustra a necessidade de um teste ***específico***.

- Suponha que dois patologistas examinaram, separadamente, o material de 100 tumores e os classificaram como benignos e malignos. A questão de interesse é saber se os patologistas ***diferem*** nos seus critérios de decisão. A Tabela 6.9 apresenta do dados na forma apropriada para análise.

Tabela 6.9: Classificação de dois patologistas (A e B) quanto à malignidade de tumores

Diagnóstico de B	Diagnóstico de A		
	Malignos	Benignos	Total
Malignos	9	1	10
Benignos	9	81	90
Total	18	82	100

- É importante observar que a unidade de análise é o tumor e, embora tenhamos 200 análises, há apenas 100 tumores.

6.4 Resposta dicotômica: amostras pareadas (cont.)

Teste de McNemar

Os dados a serem analisados no processo de comparação podem ser ***resumidos*** no formato da Tabela 6.10.

Tabela 6.10: Apresentação de dados obtidos em uma classificação de dados pareados

Controle	Tratamento		
	Sucesso	Fracasso	Total
Sucesso	k	r	n_1
Fracasso	s	l	n_2
Total	m_1	m_2	N

- A estatística de teste, com a correção de continuidade é:

$$X^2_{McN} = \frac{\left(\left|r - \frac{r+s}{2}\right| - \frac{1}{2}\right)^2}{\frac{r+s}{2}} + \frac{\left(\left|s - \frac{r+s}{2}\right| - \frac{1}{2}\right)^2}{\frac{r+s}{2}} = \frac{\left(|r-s|-1\right)^2}{r+s}$$

6.4 Resposta dicotômica: amostras pareadas (cont.)

Teste de McNemar (cont.)

- O teste consiste em rejeitar H_0 ***quando***:

$$X^2_{McN} = \frac{\left(|r-s|-1\right)^2}{r+s} > \chi^2_{1;\alpha}$$

em que que $\chi^2_{1;\alpha}$ é o valor da distribuição qui-quadrado, com 1 grau de liberdade, em que $\Pr(X \geq \chi^2_{1,\alpha}) = \alpha$; por exemplo, para $\alpha = 0{,}05$, $\chi^2_{1;0{,}05} = 3{,}841$ (veja Tabela A4).

Tabela A4: Distribuição do χ^2: $\Pr(X \geq x)$

Graus de liberdade	$\Pr(X \geq x)$						
	0,750	0,250	0,100	0,050	0,025	0,010	0,005
1	0,102	1,323	2,706	3,841	5,024	6,635	7,879

6.5 Resposta contínua: amostras independentes

Teste *t*

Este é o teste ***mais*** conhecido, para testar a diferenças entre duas amostras, adequado para situações em que as respostas aos dois tratamentos são variáveis quantitativas com distribuição gaussiana com diferentes médias e mesmo desvio-padrão, ou seja (μ_1,σ) e (μ_2,σ), respectivamente. Seguem-se os ***passos***:

- Coletam-se amostras de tamanho n_1 e n_2 dos grupos 1 e 2, respectivamente;
- Calculam-se as médias ($\bar{x}_1$ e $\bar{x}_2$) e os desvios-padrão (s_1 e s_2) das medidas dos dois grupos;
- Calcula-se a estimativa do desvio-padrão da diferença dessas medidas:

$$DP = \sqrt{s_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)},\ s_p^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$$

6.5 Resposta contínua: amostras independentes (cont.)

Teste *t* (cont.)

Seguem-se os passos (cont.):

- Calcula-se a ***estatística de teste***:

$$T = \left| \frac{\bar{x}_1 - \bar{x}_2}{DP} \right|$$

- Finalmente, rejeita-se H_0 se T é ***grande***, o que ocorre quando ocorre que $T > t_{n_1+n_2-2;\alpha}$, em que $t_{n_1+n_2-2;\alpha}$ é o percentil de ordem $1-\alpha/2$ da distribuição t com n_1+n_2-2 graus de liberdade (veja Tabela A5);
- Entretanto, em lugar de comparar o T com um valor tabelado, é mais frequente o uso do valor-p, via ***softwares*** estatísticos.

6.5 Resposta contínua: amostras independentes (cont.)

Exemplo 6.9: Comparação da tianeptina com placebo

O ensaio consistiu em administrar a droga a um dos ***dois grupos*** de pacientes, compostos de forma aleatória, e quantificar a depressão através da escala de Montgomery-Asberg (MADRS), em que os valores maiores indicam maior gravidade da depressão. A Tabela 6.14 apresenta os escores finais dos pacientes dos dois grupos.

Tabela 6.14: Escore final na MADRS de pacientes dos dois grupos admitidos em Belo Horizonte

Grupo	Escores							
Placebo	6	33	21	26	10	29	33	29
	37	15	2	21	7	26	13	
Tianeptina	10	8	17	4	17	14	9	4
	21	3	7	10	29	13	14	2

6.5 Resposta contínua: amostras independentes (cont.)

Exemplo 6.9: Comparação da tianeptina com placebo (cont.)

- Para fazer o **teste *t*** é preciso usar as seguintes informações:

$$n_1 = 15;\ \bar{x}_1 = 20,53;\ s_1 = 11,09;$$

$$n_2 = 16;\ \bar{x}_2 = 11,37;\ \text{e}\ s_2 = 7,26$$

- Estima-se a ***variância combinada*** das amostras, $s_p{}^2$:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{(15-1)(11,09)^2 + (16-1)(7,26)^2}{15+16-2} = (9,31)^2$$

- Estima-se o ***desvio-padrão*** da diferença das medidas, *DP*:

$$DP = \sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \sqrt{(9,31)^2\left(\frac{1}{15} + \frac{1}{16}\right)} = 3,34$$

6.5 Resposta contínua: amostras independentes (cont.)

Exemplo 6.9: Comparação da tianeptina com placebo (cont.)

- Calcula-se a ***estatística de teste*** *T*:

$$T = \left| \frac{\bar{x}_1 - \bar{x}_1}{DP} \right| = \left| \frac{20,53 - 11,37}{3,34} \right| = 2,74$$

- Para um α = 0,05, obtém-se na Tabela A5 o valor de $t_{n_1+n_2-2;\alpha} = t_{15+16-2;0,05} = t_{29;0,05}$, ou seja, aproximadamente, $t_{30;0,05} = 2,042$.

Tabela A5: Distribuição t de Student : $\Pr(X \leq -x) + \Pr(X \geq x)$

Graus de liberdade	$\Pr(X \leq -x) + \Pr(X \geq x)$						
	0,50	0,40	0,20	0,10	0,05	0,02	0,01
1	1,000	1,376	3,078	6,314	12,706	31,821	63,657
2	0,817	1,061	1,886	2,920	4,303	6,965	9,925
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
20	0,687	0,860	1,325	1,725	2,086	2,528	2,845
30	0,683	0,854	1,310	1,697	2,042	2,457	2,750
40	0,681	0,851	1,303	1,684	2,021	2,423	2,705

6.5 Resposta contínua: amostras independentes (cont.)

Exemplo 6.9: Comparação da tianeptina com placebo (cont.)

- Como é verdade que $T > t_{30;0,975}$, pois 2,74 > 2,042, **rejeita-se** H_0 e conclui-se que os dois medicamentos têm **diferentes** eficácias;
- Alternativamente, pode-se determinar o valor-p, através de um ***softwares*** estatístico (por exemplo, o R), o que conduziria à mesma conclusão, pois tem-se que o valor-p ≈ 0,01 é inferior a $\alpha = 0{,}05$ (5%):

```
# Carregue os dados:
placebo<-c(6,33,21,26,10,29,33,29,37,15,2,21,7,26,13)
tianeptina<-c(10,8,17,4,17,14,9,4,21,3,7,10,29,13,14,2)
# Rode o teste t
t.test(placebo,tianeptina,var.equal=TRUE)
Two Sample t-test
data:  placebo and tianeptina
t = 2.7383, df = 29, p-value = 0.01045 (< 0,05)
```

6.5 Resposta contínua: amostras independentes (cont.)

Teste *Z* para comparação de médias

No caso de ***amostras grandes*** (n_1 e $n_2 \geq 30$), há o teste Z, aplicável a amostras com variâncias σ_1 e σ_2 iguais ou diferentes;

- No teste Z, a estatística de teste é:

$$Z = \frac{\left(\bar{X}_1 - \bar{X}_2\right)}{\sqrt{s_1^2 / n_1 + s_2^2 / n_2}}$$

- Se H_0 for verdadeira, essa estatística de teste Z segue aproximadamente uma distribuição ***gaussiana padronizada***;
- Assim, ***rejeita-se*** H_0 quando $|Z| > Z_{\alpha/2}$, em que $Z_{\alpha/2}$ é o percentil de ordem $(1-\alpha/2)$ da distribuição gaussiana padrão $N(0,1)$, Tabela A3.

6.5 Resposta contínua: amostras independentes (cont.)

Exemplo 6.11: Efeito do halotano em cirurgias cardíacas
A fim de estudar o efeito de ***dois*** anestésicos, halotano e morfina, foram registradas as pressões sanguíneas antes da indução da anestesia, após a anestesia e em outros momentos importantes durante a operação. Os resultados obtidos são apresentados na Tabela 6.16.

Tabela 6.16: Média e desvio-padrão da pressão sanguínea (*mmHg*) segundo o tipo de anestesia

Informações sobre a amostra	Anestesia	
	Halotano	Morfina
Média	66,9	73,2
Desvio-padrão	12,2	14,4
n	61	61

6.5 Resposta contínua: amostras independentes (cont.)

Exemplo 6.11: Efeito do halotano em cirurgias cardíacas (cont.)

- Nas condições do problema, testaremos as ***hipóteses*** estatísticas H_0: $\mu_1 = \mu_2$ *versus* H_1: $\mu_1 \neq \mu_2$;
- Como as amostras são ***grandes*** ($\geq$ 30), pode-se usar o teste Z;
- Calcula-se a ***estatística de teste*** Z:

$$Z = \frac{\left(\bar{X}_1 - \bar{X}_2\right)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(66{,}9 - 73{,}2)}{\sqrt{\frac{12{,}2^2}{61} + \frac{14{,}4^2}{61}}} = \frac{-6{,}30}{\sqrt{5{,}84}} \cong -2{,}61$$

6.5 Resposta contínua: amostras independentes (cont.)

Exemplo 6.11: Efeito do halotano em cirurgias cardíacas (cont.)

- Adotando-se um nível de ***significância*** $\alpha = 0{,}05$ (5%), tem-se que $Z_{\alpha/2} = Z_{0,025} = 1{,}96$, pois $P(Z < 1{,}96) = 1 - 0{,}025 = 0{,}9750$ (ver Tabela A3):

Tabela A3: Distribuição gaussiana para valores positivos: $Pr(X \leq x)$ (continuação)

x	Segunda casa decimal de x									
	0	1	2	3	4	5	6	7	8	9
0,0	5000	5040	5080	5120	5160	5199	5239	5279	5319	5359
0,1	5398	5438	5478	5517	5557	5596	5636	5675	5714	5753
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1,9	9713	9719	9726	9732	9738	9744	9750	9756	9761	9767

- Logo ***rejeita-se*** H_0, pois $|Z| > Z_{\alpha/2} \rightarrow |-2{,}61| > 1{,}96$, e conclui-se que os dois anestésicos ***não são*** equivalentes;
- Alternativamente, pode-se obter o valor-p (por exemplo, no R), e chegar-se à mesma conclusão que H_0 é rejeitada (valor-p < 0,05):

```
1-pnorm(2.61)
[1] 0.004527111 (< 0,05)
```

6.6 Resposta contínua: amostras pareadas

Teste *T* para amostras pareadas

- Considera-se aqui a situação em que as amostras são n ***pares*** de observações, $(x_{11}, x_{21}), (x_{12}, x_{22}), \ldots, (x_{1n}, x_{2n})$, antes, x_{1i}, e depois, x_{2i}, do tratamento;
- Para cada par, toma-se a ***diferença*** das duas observações, $d_1 = x_{11} - x_{21}$, $d_2 = x_{12} - x_{22}$, ..., $d_n = x_{1n} - x_{2n}$;
- A partir das diferenças d_i, calculam-se a sua ***média***, $\bar{d}$, e o seu ***desvio-padrão***, s_d:

$$\bar{d} = \frac{\sum_{i=1}^{n} d_i}{n}; \quad s_d = \sqrt{\frac{\sum_{i=1}^{n} \left(d_i - \bar{d}\right)^2}{n-1}}$$

6.6 Resposta contínua: amostras pareadas (cont.)

Teste *T* para amostras pareadas

- Calcula-se a ***estatística de teste***, baseada na distância entre a média das diferenças ($\bar{d}$) e zero:

$$T_p = \frac{\bar{d} - 0}{s_d / \sqrt{n}} = \frac{\bar{d}}{s_d / \sqrt{n}}$$

- A ***regra*** do teste é rejeitar H_0 se:

$$\left| T_p \right| \geq t_{n-1;\alpha}$$

em que $t_{n-1;\alpha}$ é o percentil de ordem $(1-\alpha/2)$ da distribuição t, com $n-1$ graus de liberdade (Tabela A5).

6.6 Resposta contínua: amostras pareadas (cont.)

Exemplo 6.12: Programa para redução do nível de colesterol

A Tabela 6.17 mostra os níveis de colesterol de 12 participantes no início e no final de um programa para avaliar a ***efetividade*** de uma dieta combinada com um programa de exercícios físicos na redução do nível de colesterol. Verificar se foi efetivo o programa, com significância de 5%.

Tabela 6.17: Níveis de colesterol no início e no final do programa

	Programa		Diferença	Desvio	Desvio ao quadrado
Participante	Início (x_{1i})	Final (x_{2i})	$d = x_{1i} - x_{1i}$	$d - \bar{d}$	$(d - \bar{d})^2$
1	201	200	1	-19	367,36
2	231	236	-5	-25	633,36
3	221	216	5	-15	230,03
4	260	233	27	7	46,69
5	228	224	4	-16	261,36
6	237	216	21	1	0,69
7	326	296	30	10	96,69
8	235	195	40	20	393,36
9	240	207	33	13	164,69
10	267	247	20	0	0,03
11	284	210	74	54	2898,03
12	201	209	-8	-28	793,36
Soma	2931	2689	242	0	5885,67

6.6 Resposta contínua: amostras pareadas (cont.)

Exemplo 6.12: Programa para redução do nível de colesterol (cont.)

- As médias ***antes*** e ***depois*** foram, respectivamente, $\bar{x}_1 = 2931/12 = 244{,}25$ e $\bar{x}_2 = 2689/12 = 224{,}08$;
- A ***diferença*** entre as médias foi $\bar{d} = 244{,}25 - 224{,}08 \approx 20{,}17$;
- O ***desvio-padrão*** das diferenças foi $s_d = \sqrt{5.885{,}67/(12\text{-}1)} \approx 23{,}13$;
- A ***estatística de teste*** T_p é:

$$T_p = \frac{20{,}17}{23{,}13/\sqrt{12}} \cong 3{,}02$$

6.6 Resposta contínua: amostras pareadas (cont.)

Exemplo 6.12: Programa para redução do nível de colesterol (cont.)

- Assumindo-se um $\alpha = 0{,}05$ (5%), ***rejeita-se*** H_0, pois tem-se que $|T_p| = |3{,}02| > t_{n-1;\alpha} = t_{11;0{,}05} = 2{,}201$ (Tabela A5), e conclui-se que o programa altera (reduz) o colesterol;

Tabela A5: Distribuição t de Student : $\Pr(X \leq -x) + \Pr(X \geq x)$

Graus de	$\Pr(X \leq -x) + \Pr(X \geq x)$						
liberdade	0,50	0,40	0,20	0,10	0,05	0,02	0,01
1	1,000	1,376	3,078	6,314	12,706	31,821	63,657
2	0,817	1,061	1,886	2,920	4,303	6,965	9,925
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
10	0,700	0,879	1,372	1,812	2,228	2,764	3,169
11	0,697	0,876	1,363	1,796	2,201	2,718	3,106
12	0,695	0,873	1,356	1,782	2,179	2,681	3,055

- No próximo capítulo será visto como ***quantificar*** essa redução.

6.7 Testes não-paramétricos

- Os testes ditos ***paramétricos***, apresentados nas Seções 6.5 e 6.6, foram derivados para variáveis contínuas com ***distribuição gaussiana***;
- Os testes ***não-paramétricos*** são apropriados para a comparação de dois grupos, para dados com ***distribuição desconhecida***;
- Os seguintes testes não-paramétricos são bem populares para variáveis contínuas:
  - Teste de Mann-Whitney, para amostras independentes;
  - Teste de Wilcox, para amostras pareadas.

6.7 Testes não-paramétricos (cont.)

Exemplo 6.13: Comparação da tianeptina com placebo (cont.)
Para os dados do Exemplo 6.9, apresentado na Tabela 6.14, pode-se comparar os resultados do grupo que recebeu tianeptina com o que recebeu placebo através do teste de Mann-Witney, que ***rejeita*** H_0, ao nível de 5%.

Tabela 6.14: Escore final na MADRS de pacientes dos dois grupos admitidos em Belo Horizonte

Grupo	Escores							
Placebo	6	33	21	26	10	29	33	29
	37	15	2	21	7	26	13	
Tianeptina	10	8	17	4	17	14	9	4
	21	3	7	10	29	13	14	2

```
# Carregue os dados:
placebo<-c(6,33,21,26,10,29,33,29,37,15,2,21,7,26,13)
tianeptina<-c(10,8,17,4,17,14,9,4,21,3,7,10,29,13,14,2)
# Rode o teste de Mann-Whitney, para amostras independentes
wilcox.test(placebo,tianeptina,paired=FALSE)
Wilcoxon rank sum test with continuity correction
data: placebo and tianeptina W = 176.5, p-value = 0.02655 (< 0,05)
```

6.7 Testes não-paramétricos (cont.)

Exemplo 6.14: Evolução do tratamento com tianeptina

A Tabela 6.19 mostra os escores dos pacientes do grupo tianeptina, no primeiro, x_{1i}, e no último dia, x_{4i}. Como o escore MADRS é uma medida ordinal, o uso de teste não-paramétrico é mais apropriado. O teste ***rejeita*** H_0 ao nível de 5% (valor-p < 0,05).

Tabela 6.19: Escores dos (***mesmos***) pacientes do grupo tianeptina no primeiro, x_{1i}, e no último dia, x_{4i}

Nº	x_{1i}	x_{4i}	Nº	x_{1i}	x_{4i}
1	24	6	9	35	27
2	46	33	10	30	15
3	26	21	11	38	2
4	44	26	12	38	21
5	27	10	13	31	7
6	34	29	14	**27**	*
7	33	33	15	**34**	*
8	25	29	16	32	26

```
# Carregue os dados:
primeiroDia<-c(24,46,26,44,27,34,33,25,35,30,38,38,31,32)
ultimoDia<- c( 6,33,21,26,10,29,33,29,37,15, 2,21, 7,26)
# Rode o teste de Wilcoxon, para amostras dependentes
wilcox.test(primeiroDia,ultimoDia,paired=TRUE)
Wilcoxon signed rank test with continuity correction
data: primeiroDia and ultimoDia
V = 88, p-value = 0.003305 (< 0,05)
```