

# Estatística Aplicada às Ciências Sociais

## Sexta Edição

Pedro Alberto Barbetta

Florianópolis: Editora da UFSC, 2006

## Cap. 11 – Testes de comparação entre duas amostras

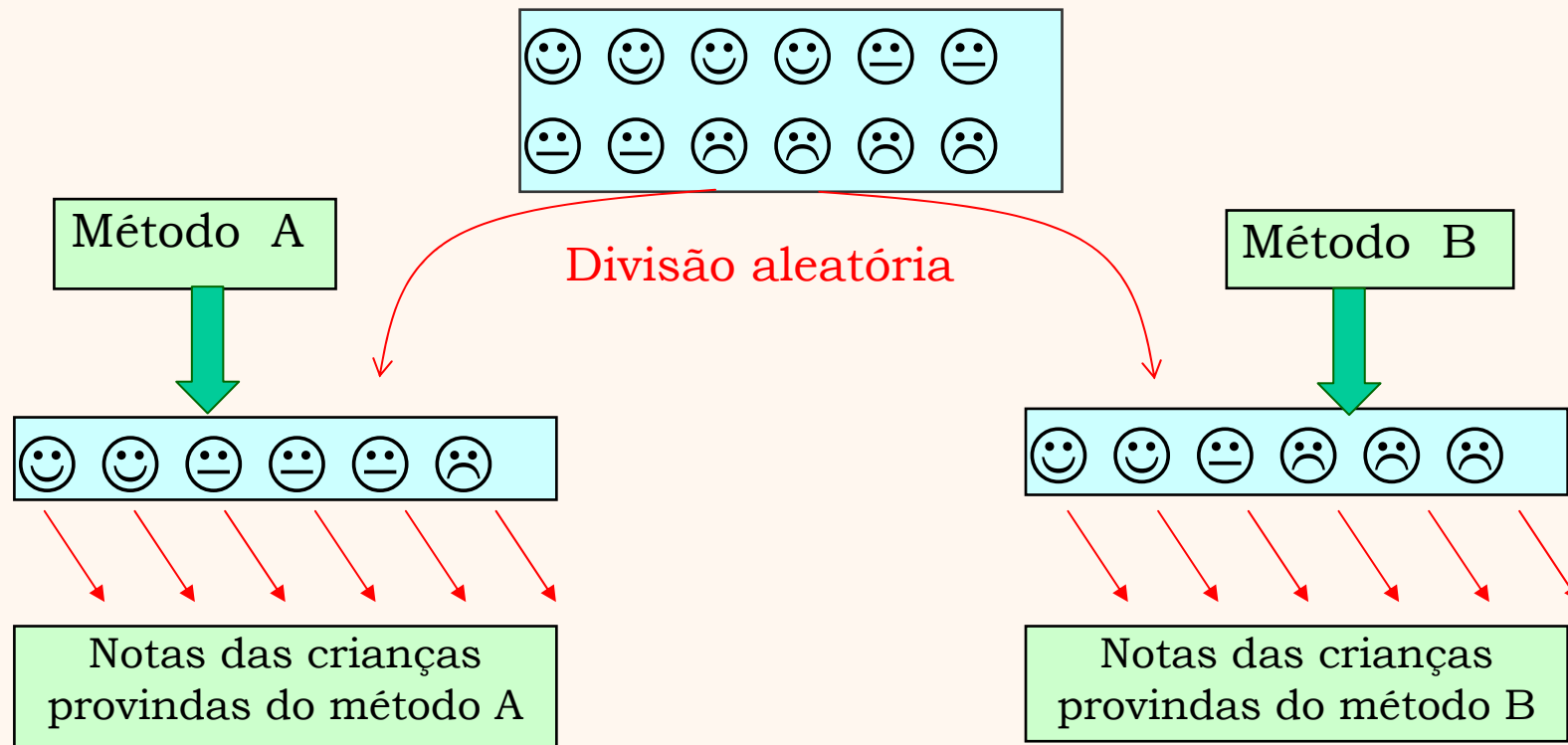
# Planejamento da pesquisa e análise estatística

- A análise estatística dos dados depende:
  - da forma como a pesquisa foi conduzida (p. ex., experimentos que geram amostras independentes ou pareadas) e
  - do tipo de dados gerado pela pesquisa (variável-resposta quantitativa ou categórica)

# Planejamento do experimento: amostras independentes

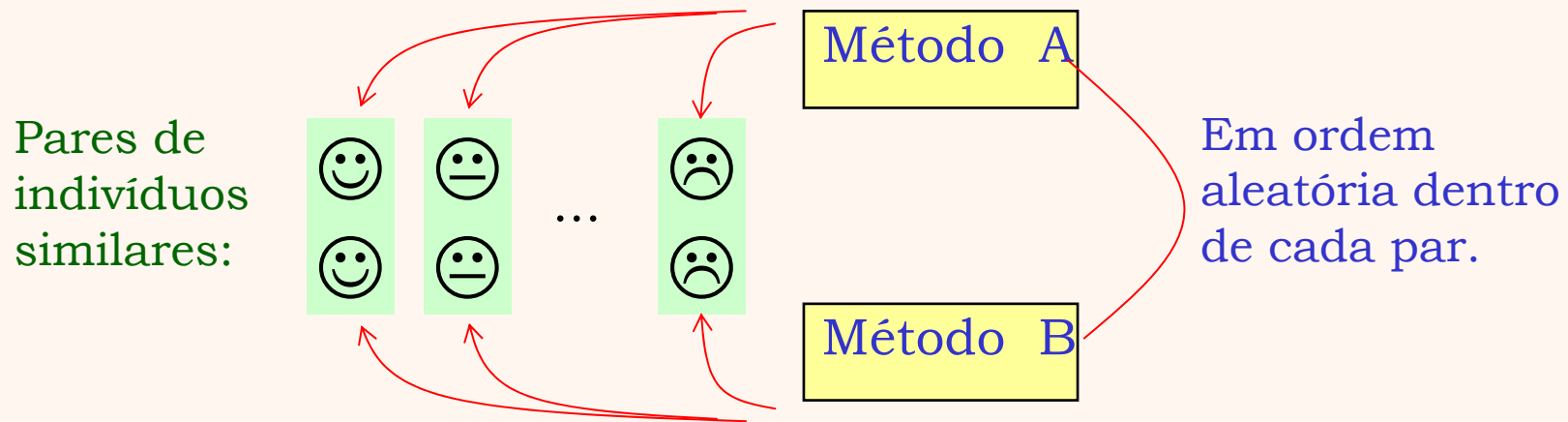
- Exemplo: experimento para comparação de dois métodos de ensino.

**Crianças selecionadas para o experimento:**



# Planejamento do experimento: amostras pareadas

- Exemplo: experimento para comparação de dois métodos de ensino.



# Planejamento do experimento:

## amostras independentes X amostras pareadas

- Dados gerados

Amostras independentes

Nota (método A)	Nota (método B)
X <sub>11</sub>	X <sub>21</sub>
X <sub>12</sub>	X <sub>22</sub>
...	...
X <sub>1n<sub>1</sub></sub>	X <sub>1n<sub>2</sub></sub>

Amostras pareadas

Par	Nota (método A)	Nota (método B)
1	X <sub>11</sub>	X <sub>21</sub>
2	X <sub>12</sub>	X <sub>22</sub>
...	...	...
n	X <sub>1n</sub>	X <sub>2n</sub>

# Testes estatísticos

- Os testes estatísticos permitem avaliar se as diferenças observadas entre os dois grupos podem ser meramente justificadas por fatores casuais ( $H_0$ ), ou se tais diferenças são reais ( $H_1$ ).

# Testes dos sinais

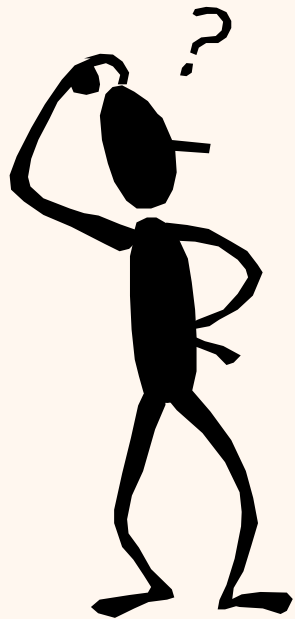
- Amostra pareada
- Em cada par tem-se apenas uma avaliação qualitativa:
  - + (diferença no par compatível com  $H_1$ ) ou
  - (diferença no par ao contrário do que afirma  $H_1$ )
- Hipóteses ( $\pi$  = probab. de + num dado par):
  - $H_0: \pi = 0,5$
  - $H_1: \pi \neq 0,5$  (podendo também ser unilateral)

## Exemplo 11.3(a)

- Com o objetivo de avaliar o efeito de um programa de treinamento sobre a produtividade dos funcionários de uma certa empresa, fez-se um estudo em que se observou a produtividade de uma amostra de funcionários antes e depois do programa de treinamento.



## Exemplo 11.3(a)



**ANTES**



**TREINAMENTO**



**Melhorou  
ou piorou?**

**DEPOIS**

## Exemplo 11.3(a)

- Hipóteses:

$$H_0: \pi = 0,5$$

$$H_1: \pi > 0,5$$

$\pi =$  probabilidade melhorar (para um dado funcionário)

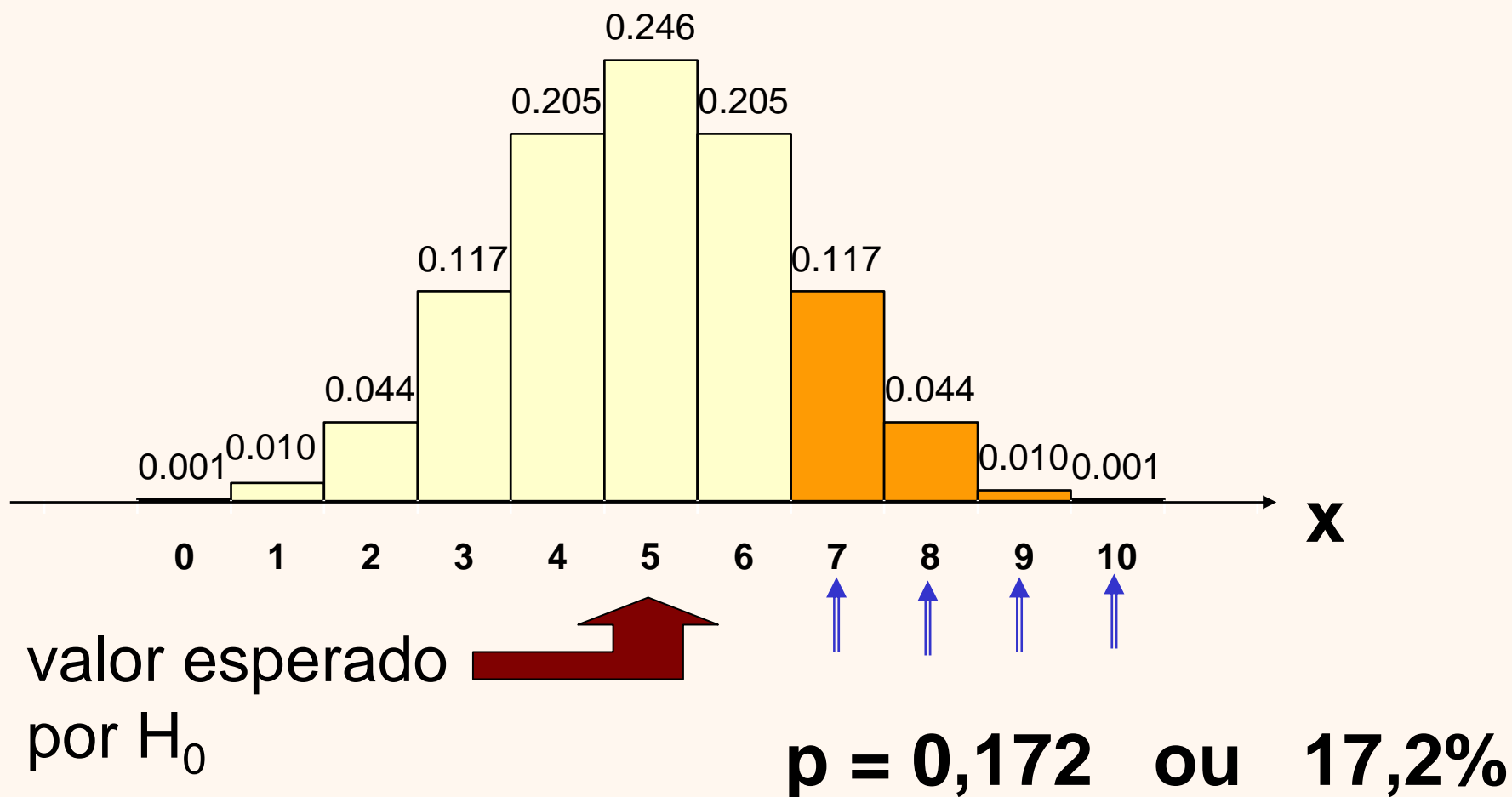
- Experimento com  $n = 10$  funcionários, observados antes e depois do treinamento

## Exemplo 11.3(a)

Resultado do experimento:

- Dos 10 funcionários, 7 melhoraram.
  - Qual é o valor-p?
  - Qual é a conclusão?

## Exemplo 11.3(a) : uso da distribuição binomial



## Exemplo 11.3(a)

- Adotando  $\alpha = 5\%$
- $p = 17,2\% > \alpha \implies$  O teste **aceita**  $H_0$
- Não se pode afirmar que o programa de treinamento realmente aumenta a produtividade dos funcionários

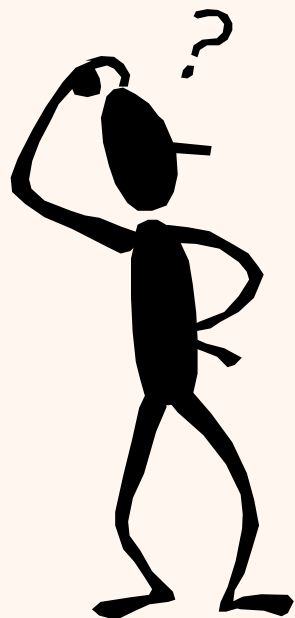
# Teste t para dados pareados

- Amostra pareada
- Em cada par tem-se uma avaliação quantitativa (medidas quantitativas da variável-resposta **X**)
- Supõe-se que a distribuição de **X** seja aprox. normal
- Hipóteses feitas em termos de médias (ou valores esperados):

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2 \text{ (podendo também ser unilateral)}$$

## Exemplo 11.3(b)



**ANTES**



**TREINAMENTO**



**DEPOIS**

De quanto  
foi a  
variação?

## Exemplo 11.3(b)

Hipóteses:

$$H_0: \mu_{\text{depois}} = \mu_{\text{antes}}$$

$$H_1: \mu_{\text{depois}} > \mu_{\text{antes}}$$

$\mu_{\text{antes}}$ : produtividade esperada antes do programa

$\mu_{\text{depois}}$ : produtividade esperada depois do programa



## Exemplo 11.3(b): amostras

Func.	Produtividade		
	antes	depois	diferença
João	22	25	3
Maria	21	28	7
José	28	26	-2
Pedro	30	36	6
Rita	33	32	-1
Joana	33	39	6
Flávio	26	28	2
Paulo	24	33	9
Catarina	31	30	-1
Felipe	22	27	5

## Exemplo 11.3(b): amostras

Func.	Produtividade		
	antes	depois	diferença
João	22	25	3
Maria	21	28	7
José	28	26	-2
Pedro	30	36	6
Rita	33	32	-1
Joana	33	39	6
Flávio	26	28	2
Paulo	24	33	9
Catarina	31	30	-1
Felipe	22	27	5

Média amostral das diferenças: **3,4**  
Desvio padrão amostral das diferenças: **3,81**

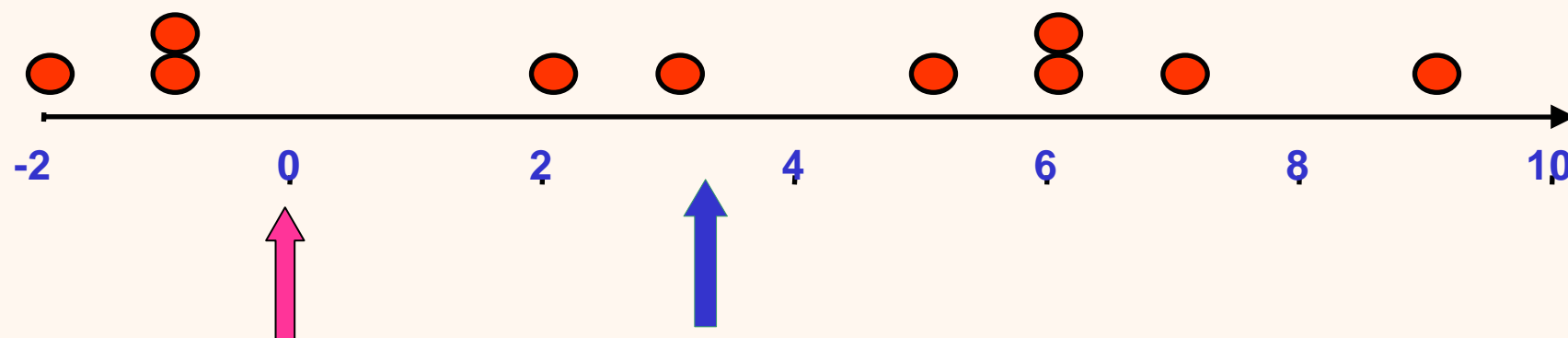
$$\bar{D} = \frac{\sum D}{n}$$

$$S_D = \sqrt{\frac{\sum D^2 - n\bar{D}^2}{n-1}}$$

## Exemplo 11.3(b): amostras

A análise é feita com a variável  
diferença:  $D = \text{depois} - \text{antes}$

$D$ : 3, 7, -2, 6, -1, 6, 2, 9, -1, 5



Valor esperado  
de  $D$ ,  $\mu_D$ , sob  $H_0$

Média dos valores  
observados de  $D$ :

$$\bar{D} = 3,4$$

# Estatística do teste

$$t = \frac{\bar{D} \cdot \sqrt{n}}{S_D}$$

onde

**n**: número de pares (*antes, depois*) observados;

$\bar{D}$ : média das diferenças observadas e

**S<sub>D</sub>**: desvio padrão das diferenças observadas

## Exemplo 11.3(b): resultados

Func.	Produtividade		
	antes	depois	diferença
João	22	25	3
Maria	21	28	7
José	28	26	-2
Pedro	30	36	6
Rita	33	32	-1
Joana	33	39	6
Flávio	26	28	2
Paulo	24	33	9
Catarina	31	30	-1
Felipe	22	27	5
Média amostral das diferenças:			<b>3,4</b>
Desvio padrão amostral das diferenças:			<b>3,81</b>

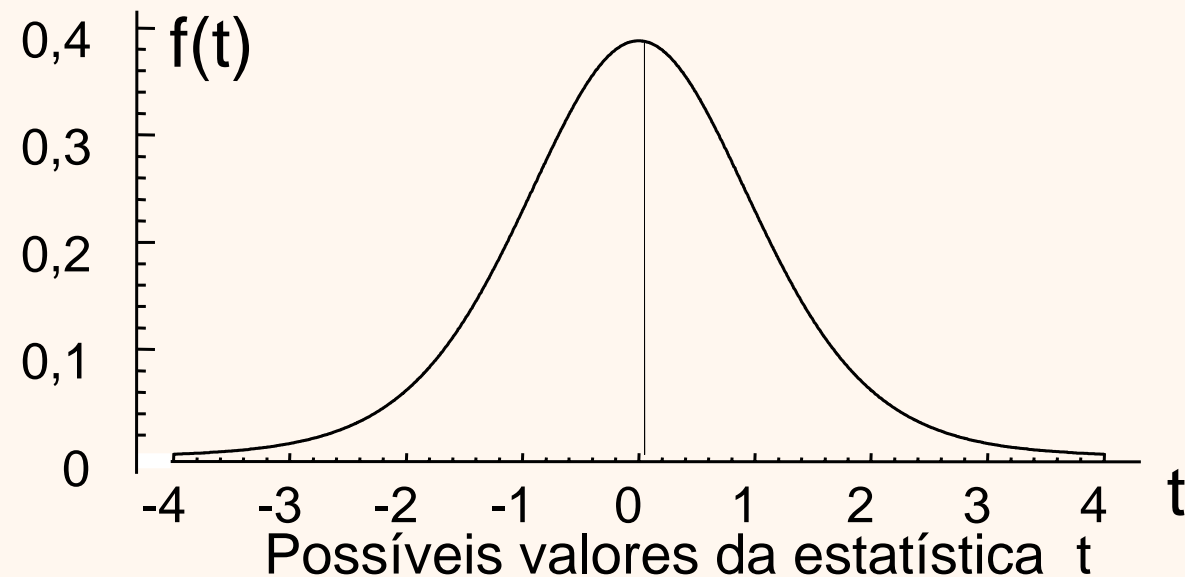
$$\bar{D} = \frac{\sum D}{n} = \frac{34}{10} = 3,4$$

$$S_D = \sqrt{\frac{\sum D^2 - n \cdot \bar{D}^2}{n - 1}} = \sqrt{\frac{246 - (10) \cdot (3,4)^2}{10 - 1}} = 3,81$$

$$t = \frac{\bar{D} \cdot \sqrt{n}}{S_D} = \frac{3,4 \cdot \sqrt{10}}{3,81} = 2,82$$

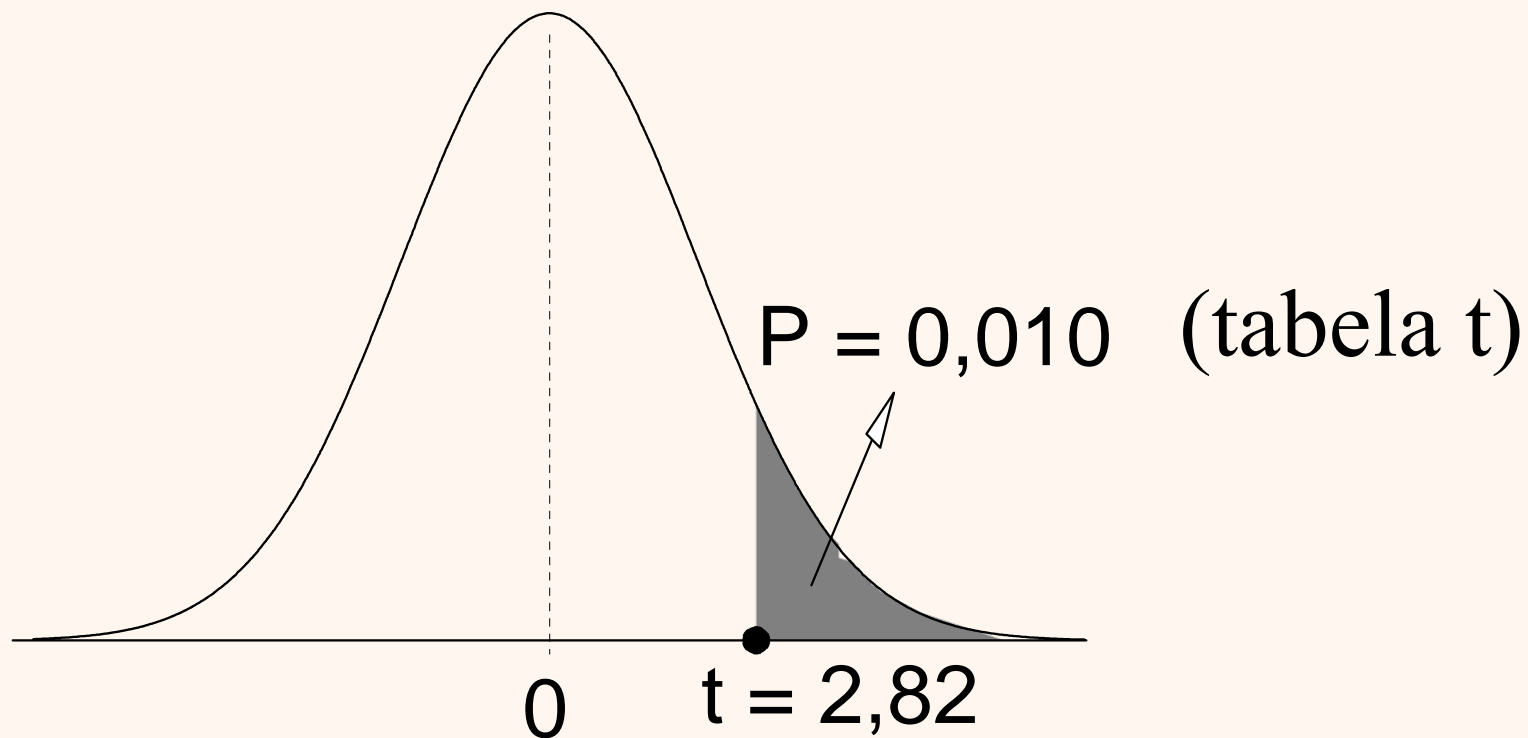
# Distribuição de referência

- Assumindo que  $D$  tenha distrib. normal, sob  $H_0$ ,  $t$  tem distrib. "t" com  $gl = n - 1$ .



## Exemplo 11.3(b): obtenção do valor-p

Distrib. de referência: t com  $gl = n - 1 = 9$



## Exemplo 11.3(b): uso da Tabela t

Amostras	<i>gl</i>	Área na cauda superior					
		0,25	0,1	0,05	0,025	0,01	0,005
	...						
$t = 2,82$	<b>9</b>	0,703	1,383	1,833	2,262	2,821	3,250
$gl = 9$	...						

⇒  $p = 0,01$



## Exemplo 11.3(b): conclusão

- O teste rejeita  $H_0$  ao nível de significância de 5%.
- O programa de treinamento aumenta a produtividade dos funcionários

# Teste t para amostras independentes

- Amostras independentes
- Variável-resposta **X** quantitativa)
- Supõe-se que a distribuição de **X** em cada grupo seja aprox. normal
- Supõe-se que as variâncias sejam aproximadamente iguais nos dois grupos
- Hipóteses feitas em termos de médias (ou valores esperados):

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2 \text{ (podendo também ser unilateral)}$$

## Exemplo 11.7

**Problema:** comparação de dois métodos de ensino.

$H_0$ : Em média, os dois métodos produzem os *mesmos* resultados;

$H_1$ : Em média, os dois métodos produzem resultados *diferentes*.

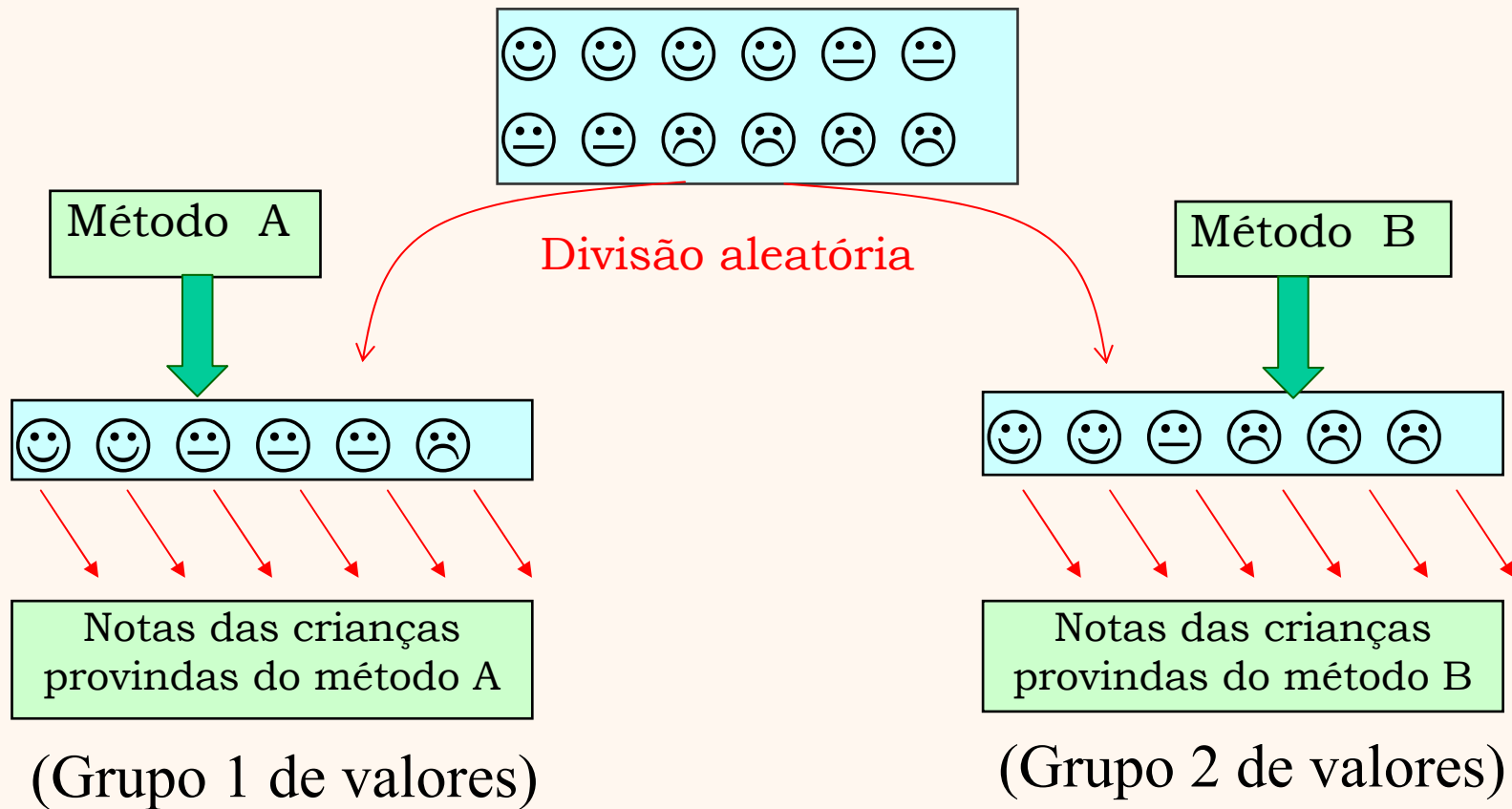
ou  $H_0: \mu_1 = \mu_2$  e  $H_1: \mu_1 \neq \mu_2$

$\mu_1$ : nota média de indivíduos submetidos ao método **A** de ensino;

$\mu_2$ : nota média de indivíduos submetidos ao método **B** de ensino.

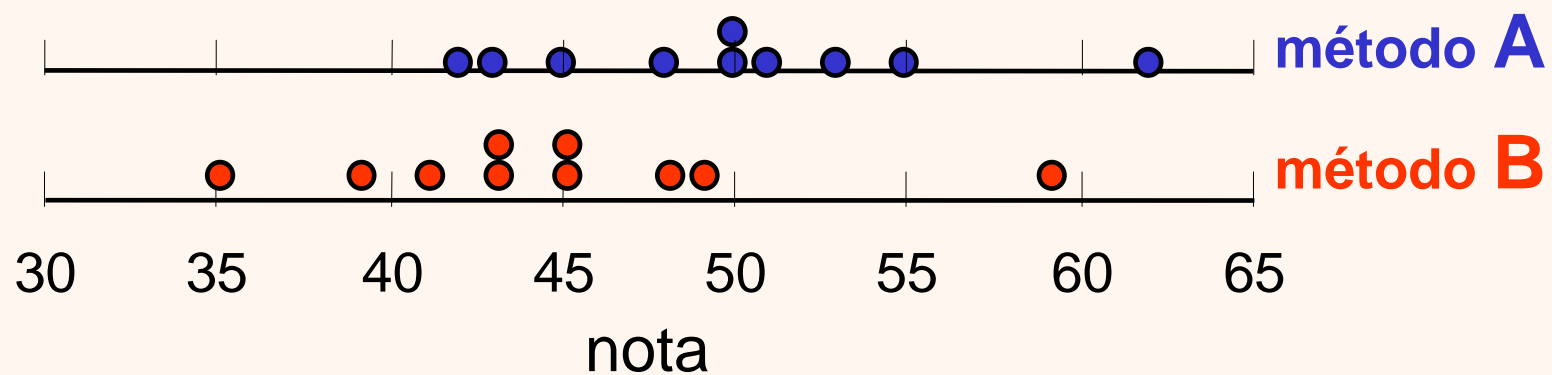
# Exemplo 11.7: o experimento

**Crianças selecionadas para o experimento:**



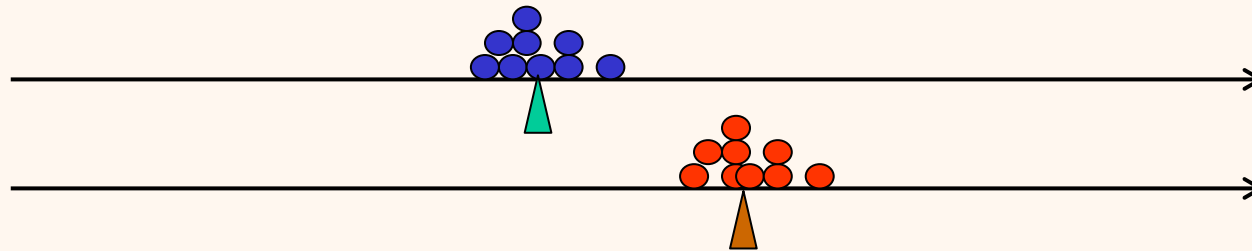
## Exemplo 11.7: amostras

	Grupo 1	Grupo 2
	45	45
	51	35
	50	43
	62	59
	43	48
	42	45
	53	41
	50	43
	48	49
	55	39
Média	49,9	44,7
Variância	35,66	42,23

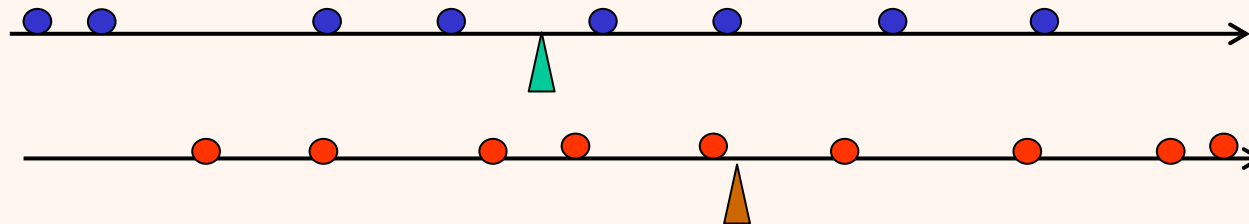


# A importância da análise da variabilidade

## a) Evidência de grupos diferentes



## b) **Não** evidência de grupos diferentes



# A estatística do teste

a) Se  $n_1 = n_2 = n$

$$t = (\bar{X}_1 - \bar{X}_2) \cdot \sqrt{\frac{n}{2S_a^2}}$$

onde

$$S_a^2 = \frac{S_1^2 + S_2^2}{2}$$

e

$$gl = 2n - 2$$

b) Se  $n_1 \neq n_2$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_a \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$S_a^2 = \frac{(n_1 - 1) \cdot S_1^2 + (n_2 - 1) \cdot S_2^2}{gl}$$

$$gl = n_1 + n_2 - 2:$$

## Exemplo 11.7: resultados

	<i>Amostra 1</i>	<i>Amostra 2</i>
n	10	10
média	49,9	44,7
variância	35,66	42,23

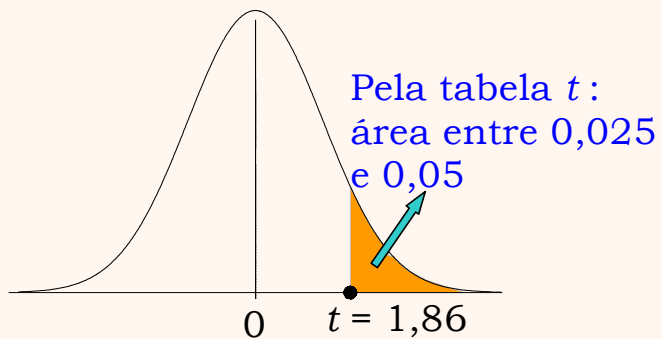
$$S_a^2 = \frac{S_1^2 + S_2^2}{2} = \frac{35,64 + 42,25}{2} = \frac{77,89}{2} = 38,95$$

$$t = (\bar{X}_1 - \bar{X}_2) \cdot \sqrt{\frac{n}{2S_a^2}} = (49,90 - 44,70) \cdot \sqrt{\frac{10}{2(38,95)}} = 1,86$$

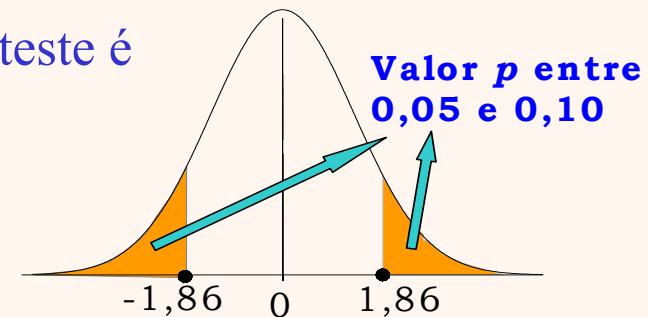


# Exemplo 11.7: uso da tabela t

Amostras		Área na cauda superior					
<i>gl</i>		0,25	0,1	0,05	0,025	0,01	0,005
...							
18	<b><math>t = 1,86</math></b> $gl = 18$	0,688	1,33	1,734	2,101	2,552	2,878
...							



Como o teste é  
bilateral:



## Exemplo 11.7: conclusão

- Valor  $p > 0,05$  (nível de significância  $\alpha$  adotado).
- Então o teste aceita  $H_0$  ao nível de significância de 5%.
- Os dados não comprovam diferença entre os dois métodos de ensino.

# Teste t

- Ver, no livro, exemplo com amostras de tamanho diferentes (Exemplo 11.8).
- Ver, no livro, exemplos de teste t em estudos não-experimentais (amostras obtidas em levantamentos por amostragem) (Exemplos 11.2 e 11.8).