Universidade Federal do Paraná

Testes t para comparação de médias de dois grupos pareados.

Gabryela Nagazawa Hayashi
Patricia Souza
Patrine Thiele

Amostras pareadas

- Quando temos duas amostras e cada observação da primeira está pareada com uma da segunda.
- Exemplos de amostras pareadas:
  - Estudo feito do antes e depois no mesmo indivíduo
  - Estudo feito com gêmeos
- Não são consideradas estatisticamente independentes, pois as duas observações são mais prováveis de serem similares
- Assume-se distribuição normal ou aproximadamente normal das diferenças
- Com dados pareados podemos usar as seguintes notações:
  - x_{1i} = medida 1 do par i
  - x_{2i} = medida 2 do par i
  - $d_i = x_{2i} - x_{1i}$
  - Um conjunto de d_i forma uma amostra de diferenças para usarmos o teste t para uma amostra pareada

Exemplo

- Exemplo de amostras pareadas transformada em uma amostra de diferenças (d_i) para efetuar o teste t para uma única amostra, porém esta pareada

Peso antes e depois em um mesmo indivíduo		
Peso antes (x_1)	Peso depois (x_2)	d_i
10	9	-1
15	12	-3
13	10	-3
14	9	-5
12	8	-4

Teste t para dados pareados

1) Especificar H_0 e H_1
   - H_0: $\mu_d = 0$
   - H_1: $\mu_d \neq 0$

2) Escolher nível de significância
   - $\alpha = 0.05$ ou 5%

Teste t para dados pareados

3) Calcular a estatística e a estatística teste média das duas amostras

- $$t = \frac{\bar{d} - \mu_d}{\mathrm{SE}(\bar{d})}$$

- $\bar{d}$ = média das diferenças (d_i)
- μ_d = valores de H_0 e H_1
- $\mathrm{SE}(\bar{d})$ = erro padrão das diferenças (d_i)
- $\mathrm{SE}(\bar{d}) = \frac{s}{\sqrt{n}}$, em que s é o desvio padrão amostral

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

Teste t para dados pareados

5. Calcular o valor de p

6. Comparar p e α

7. Descrever resultados e conclusões estatísticas

Exemplo 1

- A mudança nos valores de IMC de indivíduos do início ao final de seis meses de tratamento foram

-1,5 ; -0,6 ; -0,3 ; 0.2 ; -2 ; - 1,2

Média = -0,9

Desvio padrão = 0,81

Erro padrão = 0,33

1) Especificar H_0 e H_1
   - H_0: $\mu_d = 0$ (perda média de IMC é zero)
   - H_1: $\mu_d \neq 0$

Exemplo 1

2) Escolher nível de significância

- $\alpha = 0.05$ ou 5%

3) Calcular a estatística e a estatística teste média das duas amostras

- $$t = \frac{\bar{d} - \mu_d}{\mathrm{SE}(\bar{d})} = \frac{-0.9 - 0}{0.33} = -2.73$$

- Nota-se que t deu um valor negativo porque a mudança média observada foi a redução do IMC (um valor positivo seria o aumento do IMC)

Exemplo 1

4) Comparar o valor de t encontrado com uma distribuição de t com (n – 1) graus de liberdade

	5%	4%	3%	2%	1%	0.5%	0.2%	0.1%
2	4.303	4.849	5.643	6.965	9.925	14.089	22.327	31.599
3	3.182	3.482	3.896	4.541	5.841	7.453	10.215	12.924
4	2.776	2.999	3.298	3.747	4.604	5.598	7.173	8.610
5	2.571	2.757	3.003	3.365	4.032	4.773	5.893	6.869
6	2.447	2.612	2.829	3.143	3.707	4.317	5.208	5.959

- Graus de liberdade = 5

5) Calcular o valor de p

- 0,05 > p > 0,04

6) Comparar p e α

- p < α

Exemplo 1

7) Descrever resultados e conclusões estatísticas

- Rejeitamos H_0 ao nível de 5%, pois p não equivale, ou é maior, ao valor encontrado na coluna de 0,05 e linha 5

- Podemos concluir que existem ao nível de 5% de que há uma redução média de IMC durante o período de seis meses em indivíduos sujeitos ao tratamento

Exemplo 2

- A fim de determinar a eficiência de um medicamento antitérmico, a temperatura corporal (em graus Celsius) de 7 indivíduos foi medida. Em seguida, foi administrado o medicamento e após uma hora a temperatura foi medida novamente. Os resultados podem ser encontrados na tabela abaixo:

	Temperatura	
Indivíduo	Antes	Depois
1	37,5	37,8
2	36	36,4
3	39	37,6
4	38	37,2
5	37,8	36,9
6	38,5	37,7
7	39,3	38

	Temperatura		
Indivíduo	Antes	Depois	Diferença
1	37,5	37,8	-0,3
2	36	36,4	0,4
3	39	37,6	-1,4
4	38	37,2	-0,8
5	37,8	36,9	-0,9
6	38,5	37,7	-0,8
7	39,3	38	-1,3

Exemplo 2

- $\bar{d}$ = -0,7286
- $$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2} = 0{,}61$$
- $$\mathrm{SE}(\bar{d}) = \frac{s}{\sqrt{n}} = 0{,}23$$

1) Especificar H_0 e H_1
   - H_0: $\mu_d = 0$ (Perda de temperatura igual a zero)
   - H_1: $\mu_d \neq 0$

2) Escolher nível de significância
   - $\alpha = 0.05$ ou 5%

Exemplo 2

3) Calcular a estatística e a estatística teste média das duas amostras

- $t = \dfrac{\bar{d} - \mu_d}{\mathrm{SE}(\bar{d})}$ = -3,17
- Redução de temperatura

4) Comparar o valor de t encontrado com uma distribuição de t com (n – 1) graus de liberdade

-

	5%	4%	3%	2%	1%	0.5%	0.2%	0.1%
2	4.303	4.849	5.643	6.965	9.925	14.089	22.327	31.599
3	3.182	3.482	3.896	4.541	5.841	7.453	10.215	12.924
4	2.776	2.999	3.298	3.747	4.604	5.598	7.173	8.610
5	2.571	2.757	3.003	3.365	4.032	4.773	5.893	6.869
6	2.447	2.612	2.829	3.143	3.707	4.317	5.208	5.959

- Graus de liberdade = 6

Exemplo 2

5) Calcular o valor de p
   - $0{,}02 > p > 0{,}01$

6) Comparar p e α
   - $p < \alpha$

7) Descrever resultados e conclusões estatísticas
   - Rejeitamos H_0 ao nível de 5%, pois p não equivale, ou é maior, ao valor encontrado na coluna de 0,05 e linha 6
   - Podemos concluir que existem ao nível de 5% de que há uma redução média de temperatura durante o período de uma hora em indivíduos sujeitos ao antitérmico.

Referências

- http://leg.ufpr.br/~silvia/CE001/node48.html
- http://leg.ufpr.br/~silvia/CE001/node50.html
- http://www.leg.ufpr.br/lib/exe/fetch.php/disciplinas:ce001:estatistica.pdf

Obrigada!