

Unidade 3 – Testes para duas amostras relacionadas

Teste de McNemar

O teste de McNemar pode ser utilizado em situações de “Antes” e “Depois”. Os indivíduos são observados duas vezes e deseja-se verificar a eficácia de um determinado tratamento, por exemplo. O teste tenta captar se indivíduos se deslocam entre duas categorias e se esse deslocamento é significativo.

Os dados necessitam estar em escala nominal ou ordinal (desde que seja dicotômico), pois precisamos de estabelecer as frequências para o teste.

Procedimento:

A) enquadrar as frequências observadas em uma tabela de quatro células (a, b, c, d) na forma:

		DEPOIS		TOTAL
		A	B	
ANTES	A	a	b	a + b
	B	c	d	c + d
TOTAL		a + c	b + d	

Hipóteses:

H_0 : não existe diferença entre o antes e depois;

H_1 : existe diferença entre o antes e depois.

B) determinar as frequências esperadas nas células b e c, onde ocorreram mudanças de antes para depois:

$$fe_i = \frac{1}{2}(b+c)$$

- se as frequências esperadas são inferiores a 5, empregar o teste binomial em substituição ao teste de McNemar, fazendo x igual a menor frequência;

- se as frequências esperadas forem maiores ou iguais a 5, calcular o valor de χ^2_c com o emprego da fórmula:

$$\chi^2_{cal} = \frac{(|b - c| - 1)^2}{b + c}$$

Em que: b = número de casos observados na célula b;

c = número de casos observados na célula c.

Regra de decisão:

Se $\chi^2_{cal} > \chi^2_{\alpha, v}$, então rejeita-se H_0 .

Utilizaremos então a tabela da distribuição qui-quadrado com 1 grau de liberdade (como só temos duas categorias, $v=2-1=1$).

Exemplo:

Dois supermercados disputam a preferência dos consumidores de uma cidade. Um deles (A), para aumentar o seu número de fregueses, lança uma campanha publicitária, através de concursos, com vários brindes. O resultado final do concurso apresentou a seguinte situação, numa amostra tomada ao acaso com 100 consumidores. Verificar se a campanha foi eficiente, ao nível de 5%.

	Depois da campanha		Total
	A	B	
Antes da campanha			
A	37	3	40
B	13	47	60
Total	50	50	100

Resolução: Antes de começarmos, perceba que é uma tabela de dupla entrada. O que acontece antes, está nas linhas e o depois está nas colunas.

Antes da campanha, havia 40 clientes em A e 60 em B, totalizando 100 clientes. Depois da campanha os mesmos 100 clientes se dividiram em 50 em A e 50 em B. Ou seja, parece que houve uma diferença no que aconteceu Antes e Depois. Para verificar se essa diferença é significativa, procederemos com o teste.

H_0 : não existe diferença entre o antes e depois;

H_1 : existe diferença entre o antes e depois.

$$\chi^2_{cal} = \frac{(|b - c| - 1)^2}{b + c} = \frac{(|3 - 13| - 1)^2}{3 + 13} = \frac{81}{16} = 5,06$$

Pela tabela da distribuição qui-quadrado, $\chi^2_{0,05,1} = 3,84$. Como $5,06 > 3,84$, rejeitamos H_0 . Logo, existem evidências de diferenças entre o “Antes” e o “Depois”.

Teste dos Sinais

Para aplicar o teste dos sinais, precisamos de ter a amostra pareada. Os mesmos indivíduos medidos antes e depois.

O teste possui esse nome porque utilizaremos sinais “+” e “-” para aplicá-lo, sendo bastante utilizado em situações que não conseguimos quantificar a variável mas sabemos que uma é maior que a outra.

Hipóteses:

H_0 : os tratamentos não diferem entre si;

H_1 : os tratamentos diferem entre si;

Procedimento:

Determinar o sinal da diferença entre os dois valores de cada par;

Determinar n = número total de diferenças com sinal “+” e “-” (valores nulos não contam);

Sabendo o valor de n , determinar a probabilidade associada: se $n \leq 25$, procedemos como um teste binomial com $p = \frac{1}{2}$, fazendo x igual a menor frequência de sinais (+ ou -). Lembre-se que podemos utilizar o comando `pbinom` no R ou diretamente o comando do teste (`binom.test`).

Se $n > 25$, aproximamos pela normal, fazendo:

$$z = \frac{2x - n}{\sqrt{n}}$$

Em ambos os casos, se o p -valor $\leq \alpha$, rejeitamos a hipótese nula.

Exemplo: uma empresa submeteu oito de seus funcionários a um treinamento intensivo sobre um novo método a ser implantado, visando a um maior rendimento na produção. O resultado em número diário de peças produzido está na tabela abaixo. Aplique o teste dos sinais ao nível de significância de 10%, para decidir se o novo método deve substituir o antigo.

Funcionário	Método antigo	Método novo
1	18	24
2	15	14
3	19	22
4	23	28
5	12	16
6	16	20
7	18	20
8	17	18

Resolução: Para executar o teste, basta fazermos a diferença para cada par e tomarmos o sinal. O importante é manter a ordem até o fim, ou seja, se iniciar do antigo menos o novo, faça assim até o final. Nesse teste, observamos somente o sinal, se negativo ou positivo, não estamos

interessados na magnitude da diferença. Isso torna o teste um pouco impreciso, mas em determinadas situações é o que podemos aplicar.

H_0 : os tratamentos não diferem entre si;

H_1 : os tratamentos diferem entre si;

<i>Funcionário</i>	<i>Método antigo</i>	<i>Método novo</i>	<i>Diferença</i>
1	18	24	(18-24) = -
2	15	14	+
3	19	22	-
4	23	28	-
5	12	16	-
6	16	20	-
7	18	20	-
8	17	18	-

Temos então 7 sinais (-) e 1 sinal (+).

Logo, $n = 8$ e $x = 1$. Com isso recorremos ao teste binomial com $p=0,5$, ou calculamos a probabilidade pelo modelo binomial acumulada até $x=1$. No caso de fazermos pela distribuição binomial, devemos multiplicar o resultado por 2, pois o teste é bilateral.

No R basta fazermos:

```
binom.test(1,8,0.5,alternative="two.sided")
```

Obteremos um p -valor de 0.07 e assim concluiríamos que há diferença entre os métodos ($0.07 < 0.10$).
