



ESCOLA SUPERIOR DE
TECNOLOGIA DA SAÚDE
DE LISBOA
INSTITUTO POLITÉCNICO DE LISBOA

Mestrado Farmácia

Qui-Quadrado

Carina Silva

carina.silva@estesl.ipl.pt

SUMÁRIO

Testes de hipóteses não-paramétricos

Teste de Independência do Qui-quadrado

Teste de Homogeneidade do Qui-quadrado

Teste Exato de Fisher

Teste do Qui-quadrado

Uma das preocupações de muitos investigadores é saber que variáveis estão associadas, de que forma e intensidade. O teste do Qui-quadrado é uma forma de analisar a relação entre variáveis qualitativas (exclusivamente). Quando as variáveis são quantitativas, pode-se proceder a uma categorização e assim aplicar o teste do Qui-quadrado.

Este teste é adequado para testar se:

- Duas ou mais variáveis são independentes – Teste do qui-quadrado para a independência
- A homogeneidade de duas ou mais populações relativamente a uma variável – Teste do qui-quadrado para a homogeneidade

Ambas as situações têm as mesmas propriedades, o que difere é a formulação das hipóteses e como consequência as conclusões.

Os dados são apresentados numa tabela de contingência de dupla entrada ou bidimensional (por se estar a considerar apenas duas variáveis).

Teste do Qui-quadrado

Tabela do tipo r×c:

		Variável 1					n _{j.}
		B ₁	...	B _j	...	B _c	
Variável2	A ₁	O ₁₁	...	O _{1j}	...	O _{1c}	O _{1.}

	A _k	O _{k1}	...	O _{kj}	...	O _{kc}	O _{k.}

	A _r	O _{r1}	...	O _{rj}	...	O _{rc}	O _{r.}
	n _{.k}	O _{.1}	...	O _{.j}	...	O _{.c}	n

Onde n é a dimensão da amostra e $n = \sum_{j=1}^c O_{.j} = \sum_{k=1}^r O_{k.} = \sum_{k=1}^r \sum_{j=1}^c O_{kj}$, O_{kj} são as frequências observadas.

As frequências relativas são dadas por $f_{kj} = \frac{O_{kj}}{n}$ $f_{k.} = \frac{O_{k.}}{n}$ $f_{.j} = \frac{O_{.j}}{n}$

Para cada célula pode-se calcular as frequências esperadas: $\hat{e}_{kj} = \frac{O_{k.} \times O_{.j}}{n}$

A discrepância entre as frequências observadas e as esperadas são dadas pela estatística:

$$\chi^2 = \sum_{k=1}^r \sum_{j=1}^c \frac{(o_{kj} - e_{kj})^2}{e_{kj}} \cap \chi^2_{(r-1)(c-1)}$$

Teste do qui-quadrado para a independência

Hipóteses:

H₀: As variáveis X e Y são independentes vs. H₁: As variáveis X e Y não são independentes

Variáveis independentes \Leftrightarrow não existe associação entre X e Y

Teste do qui-quadrado para a homogeneidade

Hipóteses:

H₀: As populações (grupos) são homogêneas relativamente ao critério X

vs.

H₁: As populações (grupos) não são homogêneas relativamente ao critério X

Estatística de teste sob as condições da H_0 :

$$\chi_0^2 = \sum_{k=1}^r \sum_{j=1}^c \frac{(o_{kj} - e_{kj})^2}{e_{kj}}$$

Condições de aplicabilidade:

Numa tabela **2x2** todas as frequências esperadas são maiores ou iguais a 5. Caso esta condição não se verifique pode-se aplicar o **Teste Exato de Fisher**.

- Numa tabela **2xc** ou **rx2** todas as frequências esperadas são maiores ou iguais a 1.
- Numa tabela **rx c** menos de 20% do n.º total das células da tabela com frequências esperadas inferiores a 5 e nenhuma célula com frequência esperada inferior a 1.

Sempre que as condições não se verifiquem, o investigador, poderá agregar categorias das variáveis sempre que ache adequado.

Regra de decisão clássica:

se $\chi_0^2 \geq \chi_{(r-1)(c-1); 1-\alpha}^2$ então rejeitamos H_0 ao nível de significância α fixado

Regra de decisão que utiliza o valor_p: $valor_p = P(\chi^2 \geq \chi_0^2) = 1 - P(\chi^2 < \chi_0^2)$

Exemplo 1:

Teste do Qui-quadrado

O CHORO DOS BEBÉS

A experiência tem demonstrado que os bebés prematuros choram muito. Uma das razões sugeridas para tal facto é que estes bebés sentem falta do ambiente pré-natal.

Uma prática recente consiste em gravar os sons do interior do corpo da mãe antes do parto e posteriormente passar a gravação sempre que o bebé chora.

Interessado em estudar a relação entre a idade de um bebé e a sua reacção ao ouvir a gravação, um obstetra realizou a experiência em bebés com 1, 5, 10 e 20 semanas de idade, obtendo os

idade do bebé * reacção Crosstabulation

			reação		Total
			continuou a chorar	parou imediatament e de chorar	
idade do bebé	1 semana	Count	7	34	41
		Expected Count	16,9	24,1	41,0
	5 semanas	Count	15	29	44
		Expected Count	18,2	25,8	44,0
	10 semanas	Count	17	19	36
		Expected Count	14,9	21,1	36,0
	20 semanas	Count	23	6	29
		Expected Count	12,0	17,0	29,0
Total	Count	62	88	150	
	Expected Count	62,0	88,0	150,0	

Teste ao nível de significância de 5%, se a idade do bebé é independente da reacção.

RESOLUÇÃO:

Teste do Qui-quadrado

HIPÓTESES:

H0 : A idade do bebé e a reacção são independentes

vs.

H1 : A idade do bebé e a reacção não são independentes

Caminho no SPSS: analyse→descriptive statistics→crosstabs →statistics →chi-square...

Output:

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	28,666 ^a	3	,000
Likelihood Ratio	30,108	3	,000
Linear-by-Linear Association	28,311	1	,000
N of Valid Cases	150		

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 11,99.

Exemplo 2:

Pretende-se estudar o efeito de duas vacinas da gripe em dois grupos distintos de indivíduos, um grupo do Norte do país e o outro do Sul. Após se ter aplicado a vacina registaram-se os dados relativamente ao facto de terem ou não contraído o vírus da gripe durante o Inverno:

gripe * grupo Crosstabulation

			grupo		Total
			norte	sul	
gripe	não	Count	3	18	21
		Expected Count	8,9	12,1	21,0
		% of Total	3,8%	23,1%	26,9%
	sim	Count	30	27	57
		Expected Count	24,1	32,9	57,0
		% of Total	38,5%	34,6%	73,1%
Total	Count	33	45	78	
	Expected Count	33,0	45,0	78,0	
	% of Total	42,3%	57,7%	100,0%	

Para um nível de significância de 5%, diga se os dois grupos são homogéneos relativamente ao sucesso da vacina.

RESOLUÇÃO:

Teste do Qui-quadrado

HIPÓTESES:

H0 : A contração do vírus da gripe é homogénea nas duas zonas do país

H1 : A contração do vírus da gripe não é homogénea nas duas zonas do país

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	9,245 ^b	1	,002		
Continuity Correction ^a	7,741	1	,005		
Likelihood Ratio	10,192	1	,001		
Fisher's Exact Test				,004	,002
Linear-by-Linear Association	9,126	1	,003		
N of Valid Cases	78				

a. Computed only for a 2x2 table

b. 0 cells (,0%) have expected count less than 5. The minimum expected
8,88.

Exercícios 5 e 6 da Ficha 3
Base de dados: GLICEMIA_CEREBRO.sav