

Qui Quadrado

(Leitura complementar ao [capítulo 3](#))

Sumário:

[Como calcular](#)

[Como usar a tabela](#)

[Correção de Yates \(continuidade\)](#)

[Definição](#)

[Em tabelas de contingência](#)

[Heterogeneidade de amostras](#)

[Modelos "simplificados"](#)

[Teste exato de Fisher](#)

Definição

Qui Quadrado, simbolizado por χ^2 é um teste de hipóteses que se destina a encontrar um valor da dispersão para duas variáveis [nominais](#), avaliando a associação existente entre variáveis qualitativas.

É um teste *não paramétrico*, ou seja, não depende dos parâmetros populacionais, como [média](#) e [variância](#).

O princípio básico deste método é comparar proporções, isto é, as possíveis divergências entre as frequências observadas e esperadas para um certo evento.

Evidentemente, pode-se dizer que dois grupos se comportam de forma semelhante se as diferenças entre as frequências observadas e as esperadas em cada categoria forem muito pequenas, próximas a zero.

Portanto, o teste é utilizado para:

- Verificar se a frequência com que um determinado acontecimento observado em uma amostra se desvia significativamente ou não da frequência com que ele é esperado.
- Comparar a distribuição de diversos acontecimentos em diferentes amostras, a fim de avaliar se as proporções observadas destes eventos mostram ou não diferenças significativas *ou* se as amostras diferem significativamente quanto às proporções desses acontecimentos.

Condições necessárias

Para aplicar o teste as seguintes proposições precisam ser satisfeitas:

- Os grupos são independentes,
- Os itens de cada grupo são selecionados aleatoriamente,
- As observações devem ser frequências ou contagens,

- Cada observação pertence a uma e somente uma categoria e
- A amostra deve ser relativamente grande (pelo menos 5 observações em cada célula e, no caso de poucos grupos - pelo menos 10 - Exemplo: em tabelas 2x 2).

Como calcular

Karl Pearson propôs a seguinte fórmula para medir as possíveis discrepâncias entre proporções observadas e esperadas:

$$\chi^2 = \sum [(o - e)^2 / e]$$

em que

- o = frequência observada para cada classe,
- e = frequência esperada para aquela classe.

Note-se que $(o - e)$ = desvio (d), portanto a fórmula também pode ser escrita como

$$\chi^2 = \sum (d^2 / e)$$

Percebe-se que as frequências observadas são obtidas diretamente dos dados das amostras, enquanto que as frequências esperadas são calculadas a partir destas

É importante notar que $(o - e)$ é a diferença entre a frequência observada e a esperada em uma classe. Quando as frequências observadas são muito próximas às esperadas, o valor de χ^2 é pequeno. Mas, quando as divergências são grandes $(o - e)$ passa a ser também grande e, conseqüentemente, χ^2 assume valores altos.

Hipóteses a serem testadas

O pesquisador trabalha com duas hipóteses:

- Hipótese *nula*: As frequências observadas *não são* diferentes das frequências esperadas. Não existe diferença entre as frequências (contagens) dos grupos. Portanto, não há associação entre os grupos
- Hipótese *alternativa*: As frequências observadas *são* diferentes das frequências esperadas, portanto existe diferença entre as frequências. Portanto, há associação entre os grupos.

Procedimento

É necessário obter duas estatísticas denominadas χ^2 calculado e χ^2 tabelado. (Para ver a tabela de χ^2 clique [aqui](#)).

As frequências observadas são obtidas diretamente dos dados das amostras, enquanto que as frequências esperadas são calculadas a partir destas.

Assim, o χ^2 calculado é obtido a partir dos dados experimentais, levando-se em consideração os valores observados e os esperados, tendo em vista a hipótese.

Já o χ^2 tabelado depende do número de graus de liberdade e do nível de significância adotado.

A **tomada de decisão** é feita comparando-se os dois valores de χ^2 :

- Se χ^2 calculado $>$ ou $=$ χ^2 tabelado: *Rejeita-se* H_0 .
- Se χ^2 calculado $<$ χ^2 tabelado: *Aceita-se* H_0 .

Quando se consulta a tabela de χ^2 observa-se que é determinada uma *probabilidade de*

ocorrência daquele acontecimento.

Portanto, rejeita-se uma hipótese quando a máxima probabilidade de erro ao rejeitar aquela hipótese for baixa (alfa baixo). Ou, quando a probabilidade dos desvios terem ocorrido pelo simples acaso é baixa.

O *nível de significância* (alfa) representa a máxima probabilidade de erro que se tem ao rejeitar uma hipótese.

O número de graus de liberdade, nesse caso é assim calculado:

$$G.L. = \text{número de classes} - 1$$

E, evidentemente, quanto maior for o valor do χ^2 mais significativa é a relação entre a variável dependente e a variável independente.

Exemplo 1:

Se uma moeda não viciada for jogada 100 vezes, espera-se obter 50 caras e 50 coroas, já que a probabilidade de cair cara (p) é = $\frac{1}{2}$ e a de cair coroa (q) também é = $\frac{1}{2}$. Entretanto, na prática, é muito difícil obter valores observados, idênticos aos esperados, sendo comum encontrar valores que se desviam dos teóricos.

Supondo que uma moeda foi jogada 100 vezes e se obteve 60 caras e 40 coroas.

- a. Qual será o valor de χ^2 ?
- b. Como se pode interpretar esse valor?

Resolvendo:

As frequências esperadas em cada classe são calculadas por: $p.N$. Portanto:

$$E_{(\text{cara})} = \frac{1}{2} \cdot 100 \text{ e } E_{(\text{coroa})} = \frac{1}{2} \cdot 100$$

Assim, os valores *esperados* são: cara: 50 e coroa: 50 e os *observados* são: cara: 60 e coroa: 40.

$$\chi^2 = [(60 - 50)^2 / 50] + [(40 - 50)^2 / 50]$$

$$a. \text{ Valor de } \chi^2 = 2 + 2 = 4$$

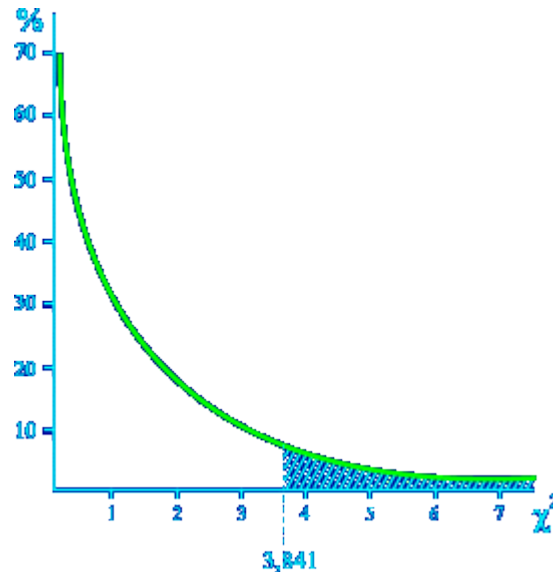
O que significa esse número? Ou seja, *como se analisa* um teste de χ^2 ?

Supondo que em vez de lançarmos 100 moedas uma única vez, tivéssemos feito inúmeros lançamentos de 100 moedas. Se calcularmos o χ^2 a cada 100 lançamentos, e, depois, colocarmos todos os resultados em um gráfico, teria sido obtida a figura ao lado.

Nota-se que os valores pequenos de χ^2 ocorrem mais frequentemente que os grandes, pois se um experimento puder ser representado pelo modelo teórico proposto, pequenos desvios casuais entre proporções esperadas e observadas ocorrerão em maior número do que grandes desvios.

Tomando a *área total sob a curva* como 100%, sabe-se que o valor 3,841 delimita 5% dela. Este é o valor crítico de qui quadrado conhecido como χ^2_c . Portanto, espera-se em experimentos semelhantes, que valores de χ^2 e menores que 3,841 tenham 95% de probabilidade de ocorrência.

Sempre que o valor de χ^2 for menor que 3,841 aceita-se a hipótese de igualdade estatística entre os números de observados e de esperados (H_0). Ou seja, admite-se que os desvios *não são* significativos.



b. Como se pode interpretar esse valor?

No exemplo dado, como o valor de Qui Quadrado obtido (4) para 2 classes foi *maior* que o esperado ao acaso (3,841), *aceita-se a hipótese alternativa* e admite-se que a moeda seja viciada.

Como usar a tabela

Entretanto, é importante notar que esse raciocínio e decisão só são válidos quando há 2 *classes* possíveis de eventos. (Como no exemplo dado, em que o lançamento da moeda pode resultar em 2 acontecimentos: cara ou coroa).

Mas, se tivéssemos lançado um dado seriam 6 classes possíveis. Como faríamos, então?

Deve-se consultar uma tabela de χ^2 .e lembrar que, nesse caso:

$$G.L. = \text{número de classes} - 1$$

A tabela de Qui Quadrado mostra o número de *Graus de liberdade* nas linhas e o valor da *Probabilidade* nas colunas.

Na coluna referente a 5% de probabilidade encontra-se o *valor crítico* de qui quadrado (χ^2_c), com o qual deve ser comparado o valor calculado de χ^2 .

						χ^2_c			
GL \ P	0,99	0,95	0,90	0,80	...	0,05	0,02	0,01	0,001
1	0,0002	0,004	0,016	0,064	...	3,841	5,412	6,635	10,827
2	0,020	0,103	0,211	0,446	...	5,991	7,824	9,210	13,815
3	0,115	0,352	0,584	1,005	...	7,815	9,837	11,345	16,266
4	0,297	0,711	1,064	1,649	...	9,488	11,668	13,277	18,467
5	0,554	1,145	1,610	2,343	...	11,070	13,388	15,080	20,515
...									
	Aceita-se a hipótese de igualdade estatística entre os números de observados e de esperados (H_0).					Rejeita-se H_0 e aceita-se H_1 .			
	Os desvios <i>não</i> são significativos.					Os números de obs e esp são estatisticamente diferentes. Os desvios <i>são</i> significativos.			

Exemplo 2:

Se um dado não viciado for jogado 6 vezes, espera-se obter 1 vez cada face (1, 2, 3, 4, 5 e 6) já que a probabilidade de cair qualquer face é 1/6.

Supondo que um dado foi jogado 186 vezes e se obteve:

Face 1	Face 2	Face 3	Face 4	Face 5	Face 6
34	29	30	32	28	33

- a. Qual será o valor de χ^2 ?
 b. Como se pode interpretar esse valor?

Resolvendo:

As frequências esperadas em cada classe são calculadas por: $p.N$. Portanto:

$$E_{(\text{face 1})} = E_{(\text{face 2})} = E_{(\text{face 3})} = E_{(\text{face 4})} = E_{(\text{face 5})} = E_{(\text{face 6})} = p.N = 1/6 \cdot 186 = 31$$

- a. Qual será o valor de χ^2 ?

Assim, os valores parciais são somados: e chega-se ao valor de χ^2 :

observado	34	29	30	32	28	33
esperado	31	31	31	31	31	31
χ^2 parcial	0,2903	0,1290	0,0322	0,0322	0,2903	0,1290

$$\chi^2 = (0,2903 + 0,1290 + 0,0322 + 0,0322 + 0,2903 + 0,1290) = 0,903$$

- b. Como se pode interpretar esse valor?

Lembrando que G.L. = número de classes -1, como há há 6 classes, G.L. = 5.

Verificando-se a tabela de χ^2 na linha em G.L. = 5 encontra-se χ^2_c igual a 11,070. Como o valor de Qui Quadrado obtido (0,903) foi *menor* que o esperado ao acaso (11,070) admite-se que o dado seja honesto.

Para facilitar os cálculos para *certas proporções* utilize uma planilha especial:

Qui Quadrado, com até 4 classes, havendo proporções esperadas

Copie a planilha comprimida em formato xls ou em ods

<http://www.cultura.ufpa.br/dicas/biome/biozip/quipropo.zip>

Em tabelas de contingência

Até aqui foram analisadas situações em que havia uma hipótese baseada em alguma teoria, gerando proporções esperadas. Por exemplo, efetuar um experimento semelhante ao de Mendel e verificar se a distribuição de uma certa variável obedece a proporção 3 :1.

Entretanto, o teste de χ^2 pode ser aplicado em casos em que não se dispõe de uma teoria que permita efetuar o cálculo de classes esperadas.

Por exemplo, supondo que se deseja verificar se uma característica se distribui igualmente entre os sexos, ou em classes sociais, ou em diferentes grupos raciais, ou em grupos etários, ou

em localizações geográficas ou...

Note-se que não existe um modo de calcular os esperados. Nesses casos constrói-se uma tabela de contingência.

Hipóteses a serem testadas

- Hipótese nula, H_0 : Não há associação entre os grupos, ou seja, as variáveis são independentes.
- Hipótese alternativa, H_a : Há associação entre os grupos, ou seja, as variáveis são dependentes.

Cálculo dos esperados:

A frequência esperada em cada classe é calculada pela multiplicação do total de sua coluna, pelo total de sua linha, dividindo-se o produto pelo total geral da tabela (N).

$$E = \text{total marginal da linha} \times \text{total marginal da coluna} / \text{total (N)}$$

O número de graus de liberdade, quando os dados estão em tabela de contingência é assim calculado:

$$G.L. = \text{número de linhas} - 1 \times \text{número de colunas} - 1$$

Procedimento

É necessário obter duas estatísticas denominadas χ^2_c calculado e χ^2 tabelado. (Para ver a tabela de χ^2 , clique [aqui](#)).

A tomada de decisão é feita comparando-se os dois valores de χ^2 e o χ^2_c . Se χ^2 e calculado > ou = χ^2_c tabelado: *Rejeita-se* H_0 .

* Critério: *

Se χ^2 calculado < χ^2_c tabelado: *Aceita-se* H_0 .

Exemplo:

Os resultados abaixo provêm de um teste sorológico aplicado a indivíduos pertencentes a 3 amostras compostas por indivíduos de provenientes de diferentes faixas etárias (crianças, adolescentes e adultos). Por à prova a hipótese de que a proporção de indivíduos com reação positiva não difere significativamente nas 3 amostras contra a hipótese de que isso não é verdadeiro.

AMOSTRA	Reação +	Reação -	Total
Crianças	25	45	70
Jovens	15	25	40
Adultos	10	30	40
<i>Total</i>	50	100	150

Para calcular os esperados multiplica-se os totais parciais relativos a cada casela e divide-se pelo total geral (N).

Por exemplo, na casela crianças + = $50 \times 70 / 150 = 23,3333$

Depois calcula-se os *qui quadrados parciais*.

Por exemplo, na casela crianças + = $(o-e)^2 / e = [(25 - 23,3333)^2 / 23,3333] = 0,1190$.

Depois, calcula-se a parcela de χ^2_c referente a cada casela.

Ao final, soma-se as parcelas e obtém-se o χ^2_c .

<i>Amostra</i>	<i>Reação +</i>	<i>Reação -</i>	<i>Total</i>
<i>Crianças</i>	25	45	70
Esp	23,3333	46,6667	
$(o-e)^2 / e$	0,1190	0,0595	
<i>Jovens</i>	15	25	40
Esp	13,3333	26,6667	
$(o-e)^2 / e$	0,2083	0,1042	
<i>Adultos</i>	10	30	40
Esp	13,3333	26,6667	
$(o-e)^2 / e$	0,8333	0,4167	
Total	50	100	150

$\chi^2 = 0,1190 + 0,0595 + 0,2083 + 0,1042 + 0,8333 + 0,4167$. Portanto, $\chi^2 = 1,7410$

O número de GL *em tabelas* é assim calculado:

GL = (número de linhas -1) x (número de colunas -1).

Portanto:

GL = (2 - 1) x (3 - 1) = 2

Depois, consulta-se a tabela de Qui quadrado e verifica-se que $\chi^2_c = 5,991$.

Como o valor de χ^2_c obtido é menor que o χ^2_c conclui-se que os desvios não são significativos.

Portanto, os indivíduos pertencentes às 3 amostras (crianças, adolescentes e adultos) reagem do mesmo modo ao teste sorológico, não havendo influência das diferentes faixas etárias sobre o resultado do teste. Assim sendo, o resultado sorológico *independe* dos grupos etários.

Para facilitar os cálculos utilize uma planilha especial:

Qui Quadrado, em tabelas de contingência, com ordem até 5 x 5
Copie a planilha comprimida em formato xls ou em ods

<http://www.cultura.ufpa.br/dicas/biome/biozip/qui55.zip>

Correção de Yates (ou Correção de continuidade)

Ao aplicar o teste de χ^2 supõe-se que o tamanho das amostras seja "grande".

Mas em situações práticas, o valor de χ^2 calculado é *aproximado*, pois

- utiliza-se amostras de tamanho finito,
- o valor da frequência observada só assumir os valores de números inteiros, ou seja nunca haverá por exemplo 2,73 indivíduos observados.

Quando se obtém um valor de χ^2 *significativo* mas nota-se que *a amostra é pequena* e/ou que

a frequência esperada em uma das classes é pequena (tipicamente, quando for menor que 5) a fórmula de obtenção de χ^2 poderá produzir um valor maior que o real.

Alguns autores, entre eles *Ronald Fisher*, recomendam que se observe a seguinte restrição:

O teste de χ^2 pode ser usado se o número de observações em cada casela da tabela for maior ou igual a 5 e a menor frequência esperada for maior ou igual a 5.

Em caso contrário, em cada classe deve ser utilizada a correção de Yates:

$$\chi^2 = \sum [(|o - e| - 0,5)^2 / e]$$

Evidentemente, não é preciso usar a correção de Yates se o valor de χ^2 obtido for menor que χ^2_c , pois o novo valor será menor que o primeiro, continuando a não ser significativo.

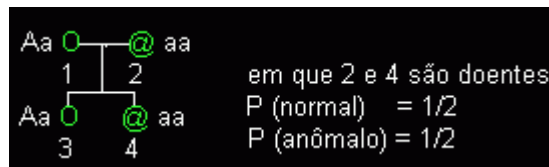
Apesar do assunto ser *controverso*, de modo geral, usa-se a correção de Yates quando:

- o valor de Qui Quadrado obtido é maior que o crítico e
- o valor de N é menor que 40 ou
- há pelo menos uma classe com número de esperados menor que 5.

Exemplo

Supondo uma coleção de irmandades, com $N = 16$, filhos de casais com um cônjuge afetado por uma anomalia endógena. 4 dentre os filhos também apresentam a doença. A característica obedece o padrão de transmissão autossômico, dominante e monogênico?

A genealogia seria:



Em 16 filhos espera-se 8 normais e 8 anormais. Mas foram observados 12 normais e 4 anormais.

Obtém-se o valor de Qui quadrado:

$$\chi^2 = [(12 - 8)^2 / 8 + (4 - 8)^2 / 8] = 2 + 2 = 4$$

Simplemente analisando o valor de χ^2 concluiria-se que como é maior que χ^2_c (3,841) os desvios não são devidos ao acaso. Portanto, conclui-se que a doença não obedece o padrão de transmissão autossômico, dominante e monogênico.

Entretanto, deve-se reparar que:

- N é menor que 20 e
- o valor de Qui Quadrado obtido é maior que o crítico

Portanto, deve-se aplicar a correção de Yates:

$$\chi^2 = [(|o_1 - e_1| - 0,5)^2 / e_1 + (|o_2 - e_2| - 0,5)^2 / e_2]$$

$$\chi^2 = [(|12 - 8| - 0,5)^2 / 8 + (|4 - 8| - 0,5)^2 / 8] = 1,51313 + 1,51313 = 3,062$$

É importante notar que agora, após ter sido aplicada a correção, $\chi^2 < \chi^2_c$, ou seja será alterada a decisão do teste.

Portanto, aceita-se que os desvios são devidos ao acaso. Assim, conclui-se que a doença

obedece o padrão de transmissão autossômico, dominante e monogênico.

Para facilitar os cálculos utilize uma planilha especial:

Qui Quadrado, em tabelas de contingência,
com ordem até 2 x 2, com correção de Yates
Copie a planilha comprimida em formato xls ou em ods

<http://www.cultura.ufpa.br/dicas/biome/biozip/qui22y.zip>

Teste exato de Fisher

Em amostras pequenas o erro do valor de Qui quadrado é alto e, portanto, o teste não é recomendável.

Ronald Fisher apresentou outro teste que permite calcular a *probabilidade de associação das características* que estão em análise, ou seja, a probabilidade de tais características serem independentes, quando o número total de dados é *pequeno*.

Assim, em amostras pequenas deve-se executar esse teste, pois produz erro menor que o teste de Qui Quadrado.

Apesar do assunto ser um pouco *controverso*, de modo geral usa-se o Teste exato de Fisher quando:

- o valor de $N < 20$ ou
 - $20 < N < 40$ e a menor frequência esperada for menor que 5.
- A análise do teste de Fisher é feita como a de χ^2

GL / P	P									
	0,99	0,95	0,90	0,80	...	0,05	0,02	0,01	0,001	
1	
2	
...	
.	Os desvios <i>não</i> são significativos. As variáveis estudadas são independentes. A associação entre as variáveis é devida ao acaso.					Os desvios <i>são</i> significativos. As variáveis não são independentes. A associação entre as variáveis estudadas <i>não</i> é devida ao acaso.				

A probabilidade calculada será igual ao produto dos fatoriais dos totais marginais pelo fatorial do total geral multiplicado pelo inverso do produto dos fatoriais dos valores observados em cada classe.

Para facilitar, identifiquemos as células por letras.

A probabilidade de dependência das duas amostras será definida por:

$$P = \left[\frac{G! H! E! F!}{I!} \right] \times \left[\frac{1}{(a! c! b! d!)} \right]$$

1. Se houver célula com o valor zero

Exemplo

Supondo a presença de uma determinada enzima em pessoas submetidas a uma reação sorológica:

Reação	Enzima		Total
	Presente	Ausente	
+	5	1	6
-	0	3	3
<i>Total</i>	5	4	9

$$P = [(G! H! E! F!) / I!] \times [1 / (a! c! b! d!)]$$

$$P = [(6! 3! 5! 4!) / 9!] \times [1 / (5! 1! 0! 3!)] = 0,0476 = 4,76\%$$

Como esse valor é menor que 5% a hipótese das características serem independentes é rejeitada, dizendo-se que a sua associação não é casual.

2. Se não houver célula com o valor zero

Deve-se:

- calcular a probabilidade identicamente ao escrito acima
- construir outra tabela 2x2, subtraindo-se uma unidade dos valores da diagonal que contiver o menor número de casos e adicionando essa unidade aos valores das caselas da outra diagonal
- calcular novamente a probabilidade
- esse processo continuará até que se atinja o valor 0
- somar todas as probabilidades calculadas

Exemplo: Supondo que os valores obtidos sejam:

Reação	Enzima		Total
	Presente	Ausente	
+	5	3	8
-	2	5	7
<i>Total</i>	7	8	15

Calcularíamos

			Total	
	5	3	8	$P = (8! 7! 7! 8! / 15!) (1 / 5! 3! 2! 5!)$
	2	5	7	$P = 0,1828$
Total	7	8	15	
	6	2	8	$P = (8! 7! 7! 8! / 15!) (1 / 6! 2! 1! 6!)$
	1	6	7	$P = 0,0305$
Total	7	8	15	
	7	1	8	$P = (8! 7! 7! 8! / 15!) (1 / 0! 7! 1! 7!)$

	0	7	7	P = 0,0012
Total	7	8	15	

$$P = 0,1828 + 0,0305 + 0,0012 = 0,2145 = 21,45 \%$$

Nesse caso, como o valor encontrado de P é maior que 5%, a hipótese das características serem independentes é aceita, dizendo-se que a sua associação é casual.

Observação: Em tabelas com mais de 2 colunas ou 2 linhas, ou seja, quando $G.L. > 1$, pode-se utilizar o teste de χ^2 se o número de caselas com frequência esperada inferior a 5 é menor que 20% do total de caselas e nenhuma frequência esperada for zero.

Para facilitar os cálculos utilize uma planilha especial:

Teste Exato de Fisher em tabelas 2 x 2

Copie a planilha comprimida em formato xls ou em ods

<http://www.cultura.ufpa.br/dicas/biome/biozip/fisher1zip>

Heterogeneidade entre amostras

Pode-se testar se amostras diferentes em uma série de experimentos semelhantes são homogêneas ou não.

Nesse caso, calcula-se o χ^2 de cada amostra e o χ^2 do total delas (χ^2_T).

Depois, soma-se os χ^2 obtidos para cada amostra ($\Sigma\chi^2$) e da soma se subtrai o valor obtido para o total de qui quadrados (χ^2_T).

O valor final obtido é o χ^2 de heterogeneidade.

Exemplo

Em 3 amostras de filhos de casais MN, obteve-se:

Amostras	MM	MN	NN	Total
Belém - PA	19	38	23	80
Maceió - AL	18	25	17	60
São Carlos - SP	8	23	9	40
Total	45	86	49	180
Proporções esperadas	1/4	1/2	1/4	

Pergunta-se: As amostras são homogêneas?

Nesse caso, deve-se calcular o χ^2 para cada amostra, a partir de proporções esperadas (MM = 1/4, MN = 1/2 e NN = 1/4), obtendo-se:

Amostras		χ^2	GL
Belém - PA	$(19-20)^2/20 + (38-40)^2/40 + (23-20)^2/20$	0,600	2
Maceió - AL	$(18-15)^2/15 + (25-30)^2/30 + (17-15)^2/15$	1,700	2
São Carlos - SP	$(8-10)^2/10 + (23-20)^2/20 + (9-10)^2/10$	0,950	2
Total	$(45-45)^2/45 + (86-90)^2/90 + (49-45)^2/45$	0,534	2

Para obter a $\Sigma\chi^2$ basta somar os valores de χ^2 obtidos em cada classe e os seus G.L.:

$$0,600 + 1,700 + 0,950 = \Sigma\chi^2 = 3,250 \text{ e } GL = 2 + 2 + 2 = 6.$$

E, então, subtrair $\chi^2_{T.}$ de $\Sigma\chi^2$

$$3,250 - 0,534 = 2,716 \text{ e } GL = 6 - 2 = 4. \text{ Portanto, } \chi^2_4 = 2,716.$$

Na *tabela* verifica-se que χ^2_4 crítico = 9,488.

Conclui-se que como o valor de χ^2 calculado é *menor* que o de χ^2_c aceita-se que os desvios são devidos ao acaso, não sendo significativos. Portanto as amostras são homogêneas.

Teste de hipóteses e a distribuição de Poisson

Quando se quer provar se os dados de uma mostra seguem a distribuição de Poisson deve-se:

- reunir em uma única classe as que tiverem valores esperados menores que 5
- calcular os esperados segundo a distribuição de Poisson

Número de vezes em que o evento ocorre:	0	1	2	3	4
Números esperados de amostras	n/e^u	nu/e^u	$nu^2/2.e^u$	$nu^3/2.3.e^u$	$nu^4/2.3.4.e^u$

- calcular a parcela de χ^2 , ou seja $[(o-e)^2/e]$ referente a cada classe
- GL = número de classes - 2 (o total e a média da amostra)

Exemplo

Foram observados os seguintes números de tripanossomos em 128 campos de um hemocítômetro

Tripanossomos	Freq. Obs.	Freq. Esp.
0	26	21,9
1	37	38,7
2	31	34,1
3	18	20,1
4	10	8,9
5	3	3,1
6	3 6	0,9 4,3
7	0	0,2
>7	0	0,1
Total	128	

Portanto, os dados resultaram em

Tripanossomos	Freq. Obs.	Freq. Esp.	$(o-e)^2/e$
0	26	21,9	0,768
1	37	38,7	0,075
2	31	34,1	0,282

3	18	20,1	0,219
4	10	8,9	0,136
≥ 5	6	4,3	0,672
Total	128	128,0	$\chi^2 = 2,152$

GL = 6 - 2 = 4 e χ^2_{4} crítico = 9,488.

Como o χ^2_4 encontrado é menor que o χ^2_4 crítico conclui-se que os desvios não são significativos. Portanto, pode-se aceitar que o número de tripanossomos encontrado segue a distribuição de Poisson.

Métodos "simplificados" para cálculo em tabelas 2 por x

Até algumas décadas atrás, quando não se usava calculadoras ou computadores, eram utilizadas maneiras de efetuar cálculos em alguns tipos de tabelas que facilitavam em muito a obtenção de resultados, por efetuarem menor número de operações.

Assim, quase que só por curiosidade, alguns estão especificados [aqui](#).

Este "site", destinado prioritariamente aos alunos de Fátima Conti, pretende auxiliar quem esteja começando a se interessar por Bioestatística, computadores e programas, estando em permanente construção. Sugestões e comentários são bem vindos. Agradeço antecipadamente.

Endereço dessa página:

HTML: <http://www.cultura.ufpa.br/dicas/biome/bioqui.htm>

PDF: <http://www.cultura.ufpa.br/dicas/pdf/bioqui.pdf>

Última alteração: 11 out 2009 (Solicito conferir datas. Pode haver atualização só em HTML)