7.13 Exercicios sobre Qui-quadrado

Exercício 158. Um modelo de automóvel é vendido em quatro versões: SX, LX, GLX, GTX. Foi feita uma campanha publicitária para melhorar as vendas das versões GLX e GTX. Posteriormente, foi verificada a escolha das versões em 500 vendas escolhidas ao acaso. Os resultados foram:

Versão	SX	LX	GLX	GTX	total
Unidades vendidas	210	125	105	60	500

De acordo com o fabricante, a participação de cada versão nas vendas deste modelo até a realização da campanha era 40% de SX, 30% de LX , 20% de GLX e 10% de GTX.

Utilize Teste Qui-quadrado, com o nível de significância de 2,5%, para verificar se houve ou não mudanças na participação de cada versão nas vendas após a campanha.
Ajuda: *A tarefa, sendo colocada em termos matemáticos precisos, soa assim: Verifique, usando Teste Qui-quadrado, com o nível de significância de 2,5%, se os dados da tabela (a distribuição amostral) sugiram que a participação de cada versão nas vendas (quer dizer, a de todas as vendas, ou, em outras palavras, a distribuição populacional de vendas) após a campanha publicitária continua na mesma proporção que houve antes da campanha.*

Exercício 159. Os dados seguintes representam os resultados de uma investigação da distribuição do sexo das crianças de 96 famílias possuindo cada uma delas 4 crianças (espero que seja calro que trata-se da distribuição amostral construída com base na amostra de 96 famílias escolhidas ao acaso do universo de todas as famílias que possuam 4 crianças).

Número de meninos	0	1	2	3	4	total
Número de famílias	12	30	24	21	9	96

Verifique, usando Teste Qui-quadrado com o nível de significância de 5%, se a amostra comprova o seguinte fato: o número de meninos por família, no universo de famílias com 4 crianças, segue a distribuição binomial com $n = 4$ e $p = 0,50$.
Comentário: *Se supusermos que cada criança, ao nascer numa família, é menino ou é menina com as probabilidades que não dependem dos sexos das crianças já nascidas na família, e se supusemos ainda, que cada uma das duas probabilidades é 1/2, então, a distribuição do número de meninos em famílias com 4 crianças deve ser binomial*$(4; 1/2)$. *Então, se a amostra do presente exercício confirmar tal distribuição, poderemos concluir que as duas suposição são válidas na vida real.*

Exercício 160. Um estudo (Anais Brasileiros de Dermatologia, 2002, 78(3), pp. 283–288) sobre dermatoses infeccionais em pacientes trasplantados renais tem por objetivo verificar se existia uma relação entre a presença do fungo *Pitiríase versicolor* e o tempo transcorrido desde o transplante.

▷ (*É conhecido que nos pacientes transplantados renais, a imunossupressão crônica acarreta maior suscetibilidade às dermatoses infecciosas, fato que motiva a pergunta se "presença do fungo" e "tempo transcorrido desde transplante" sejam fatores relacionados ou não.*)

Os resultados de acompanhamento de 208 pacientes estão na tabela abaixo.

▷ (*O acompanhamento durou 12 meses, e o* Pitiríase versicolor *declarava-se "presente" para paciente, caso esse fungo foi identificado nos exames dermatológicos do mesmo feitos no decorrer do acompanhamento; caso contrário, o fungo declarava-se "ausente".*)

Pitiríase versicolor	Tempo de transplante menos que 1 ano	de 1 a 5 anos	mais que 5 anos	**Total**
Presente	6	24	7	37
Ausente	36	53	82	171
Total	42	77	89	208

(a) Verifique, usando Teste Qui-quadrado, ao nível de significância de 0,5%, se os dados da tabela indicam a independência entre a presença do fungo *Pitiríase versicolor* e o tempo transcorrido desde o transplante.
(b) Nas conclusões do estudo, os pesquisadores escrevem: *"As dermatoses infecciosas são freqüentes nos pacientes transplantados renais, e sua ocorrência aumenta progressivamente conforme o tempo transcorrido a partir do transplante, sendo importante o acompanhamento dermatológico desses pacientes."* Tal conclusão poderia ser feita com base em somente o resultado da aplicação do Teste Qui-quadrado (na citação, interprete "as dermatoses infecciosas" como "a presença de *Pitiríase versicolor*")?

Exercício 161. Um estudo (Revista de Saúde Pública, 2001; 35(2), pp. 159–164) sobre tabagismo entre adolescentes residentes no município de Petolas, RS, tinha por objetivo saber se existia alguma relação entre faixa etária e condição de fumante. Os dados colhidos para dar o respaldo à conclusão, estão apresentados na tabela abaixo.

	12 a 14 anos	15 a 16 anos	17 a 18 anos	total
fumante	15	26	30	69
não fumante	268	175	118	563
total	283	201	148	632

Verifique, usando Teste Qui-quadrado, ao nível de significância de 0,5%, se os dados da tabela indicam a independência entre faixa etária e condição de fumante na população de adolescentes residentes no município de Petolas, RS. a presença do fungo *Pitiríase versicolor* e o tempo transcorrido desde o transplante.

Exercício 162. Em um estudo realizado no Hospital Universitário (HU) e analisado no CEA (Centro de Estatística Aplicada) do Departamento de Estatística, estudou-se a ocorrência de bebês macrossômicos (bebês que nascem com mais de 4Kg) entre as parturientes do HU. Para verificar se o ganho do peso da mãe durante a gestação está relacionado com a ocorrência de macrossomia, uma amostra de 171 mães foi selecionada (não foram incluídas na amostra gestações múltiplas nem bebês prematuros). Destas mães, observou-se que 53 bebês eram macrossômicos, que 50 mães tiveram ganho excessivo de peso ($>$ 25% do peso pré-gravídico), e que das 121 mães com ganho inferior a 25% do peso pré-gravídico, 90 tiveram bebês sem macrossomia.

Aplique Teste Qui-quadrado aos dados da amostra para concluir, ao nível de significância de 10%, o ganho de peso da mãe e a ocorrência de macrossomia são fatores independentes na população das parturientes do HU que tiveram bebês macrossômicos.

Exercício 163. Oitenta artefatos foram classificados de acordo com o período da provável confecção (A, B e C) e de acordo com o sítio arqueológico (Sítio 1 e 2) em que foram encontrados. Os resultados encontrados foram:

Sítio 1: A B B B C C B A B C C B A A A B B B C C A A A A A B B B A C C B B B A A A A

Sítio 2: C C C C A A B B C C A A B C C C C C B B A B A C B C B C C C C B A B C B C C C C B C

Os oitenta artefatos encontrados sugerem que há associação entre o período de confecção e o sítio arqueológico? Responda através de Teste Qui-quadrado com o nível de significância de 10%.

Exercício 164. Em um estudo sobre alcoolismo, foram sorteadas 500 pessoas adultas de uma certa população e observadas duas variáveis para cada uma delas: dependência de álcool (sim/não) e número de progenitores com dependência (0, 1 ou 2). Entre os 80 portadores de dependência, 10 não tinham qualquer progenitor dependente e 24 tinham 1 progenitor dependente. Das 500 pessoas pesquisadas, 60 tinham 1 progenitor dependente e 50 tinham 2 progenitores dependentes.

Use os dados acima para verificar, via o Teste Qui-quadrado com o nível de significância de 5%, se exista um componente genético que influi no alcoolismo.

Exercício 165. Em um estudo para verificar a relação entre asma e incidência de gripe no outono, 150 crianças foram escolhidas ao acaso, dentre aquelas acompanhadas pelo Posto de Saúde de um bairro, nessa época de ano. Os resultados estão na tabela a seguir.

asma	gripe		total
	sim	não	
sim	27	34	61
não	42	47	89
total	69	81	150

Essa amostra evidencia que na população das crianças acompanhadas no outono pelo Posto de Saúde do bairro, a incidências de gripe e de asma são fatores independentes entre si? Use Teste Qui-quadrado com 5% de nível de significância.
Observação: A solução desse exercício indicará ao quê se refere o termo “graus de liberdade”, e, desse modo, justificará o uso das palavras “grau” e “liberdade”. Toda essa informação é opcional.

Exercício 166. Um estudo tem por objetivo verificar se a preferência por certo tipo de filme está associada com o estado civil ou se a preferência e o estado civil são fatores independentes. Uma amostra apresentou os seguintes resultados:

Estado	Preferência		
civil	Policial	Comédia	Romance
Solteiro	40	25	30
Casado	26	31	33
Divorciado	46	33	31
Viúvo	33	38	34

Use Teste Qui-quadrado com nível de significância 2,5% para deduzir dos dados da amostra se, no nível de população, o estado civil e a preferência por tipo de filme (de acordo com a classificação policial/comédia/romance) são fatores independentes.

Exercício 167. Um estudo tem por objetivo verificar se os aproveitamentos em Física e Matemática de alunos de uma escola de ensino médio estão relacionados. Uma amostra de alunos da escola apresentou os seguintes resultados:

	Matemática		
Física	Notas altas	Notas regulares	Notas baixas
Notas altas	46	71	22
Notas regulares	47	143	58
Notas baixas	29	72	40

Use Teste Qui-quadrado para concluir a partir dos dados de amostra, ao nível de significância 10%, se na população de alunos da escola apresenta a independência entre os aproveitamentos em Física e em Matemática.

7.14 Soluções de exercícios sobre Qui-quadrado

Solução de Exc. 158. Ao encaixar o enredo do exercício no arcabouço do Teste Qui-quadrado quando esse está aplicado para verificação de aderência, traduz-se que a tarefa do exercício da seguinte maneira: é para escolher entre as hipóteses que

H: a campanha publicitária não teve efeito nas vendas das versões desse modelo de automóvel, quer dizer, as proporções, nas quais a população compra as versões do modelo de automóvel em questão após a campanha, continuam ser 40% para SX, 30% para LX , 20% para GLX e 10% para GTX;

A: a campanha publicitária teve efeito nas vendas das versões desse modelo de automóvel, e, consequentemente, as quatro proporções suprareferidas diferem-se dos quarteto 40%, 30%, 20%, 10%, que corresponde às proporções antes da campanha.

As hipóteses H e A podem ser reescritas da seguinte maneira:

H: $p_{SX} = 0,40$, $p_{LX} = 0,30$, $p_{GLX} = 0,20$ e $p_{GTX} = 0,10$;

A: pelo menos uma das igualdades acima não está válida,

em que p_{SX}, p_{LX}, p_{GLX} e p_{GTX} representam a participação nas vendas da versão SX, LX, GLX e GTX, respectivamente, participações essas expressas em proporções e – vale lembrar – correspondem ao que acontece com a população dos compradores do referido automóvel após a campanha publicitária.

Qual das duas maneiras para escrever H e A você vai preferir é algo que não ajuda nem atrapalha a resto de solução. O que é de fato importante são duas perícias: a primeira é enchergar que o exercício solicita a verificação de aderência (e não de independência), e a segunda é nomear corretamente as hipóteses (aquela que alega a aderência deve ser nomeada de nula, isto é, "H").

Para realizar o teste, o passo seguinte após a determinação de hipóteses, é calcular as frequencias esperadas que correspondem ao tamanho de amostra fornecida no enunciado (especificamente falando, trata-se das frequencias que a gente esperaria obter se analizasse 500 compradores, se a aleatoriedade não estivesse presente e se a hipotese nula estivesse válida). O passo próximo é o cálculo do valor da estatística de teste correspondente à amostra. Eis a

Versão	Frequência observada (O_i)	Proporção/probabilidade (p_i) sob a validade de H	Frequência esperada (E_i) correspondente ao tamanho de amostra sob a validade de H
SX	210	$p_1 = 0,4$	$E_1 = n \times p_1 = 500 \times 0,4 = 200$
LX	125	$p_2 = 0,3$	$E_2 = n \times p_2 = 500 \times 0,3 = 150$
GLX	105	$p_3 = 0,2$	$E_3 = n \times p_3 = 500 \times 0,2 = 100$
GTX	60	$p_4 = 0,1$	$E_4 = n \times p_4 = 500 \times 0,1 = 50$
Total	500	1	500

conta:

$$\chi^2_{obs} = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} = \frac{(210-200)^2}{200} + \frac{(125-150)^2}{150} + \frac{(105-100)^2}{100} + \frac{(60-50)^2}{50}$$
$$= 0,5000 + 4,1667 + 0,2500 + 2,0000 = 6,9167.$$

No penúltimo passo de solução, é precisa achar o limiar correspondente ao desejado nível de significância α) do teste proferido. Recordo do enunciado: $\alpha = 2,5\%$. O limiar procurado (denotado por ℓ nas minhas aulas e por χ^2_α na literatura científica e didática) é – pela própria definição – o valor que recorta da distribuição Qui-quadrado α por cento de sua cauda direita. Mas a cauda deve ser recortada da distribuição adequada. A adequação está expressa pelo valor do parámetro chamado *graus de liberdade* e denotado genericamente por q. No caso de teste de aderência, a maniera mais simple e mais prática para a determinação do valor de q é olhar na tabela que apresenta resultados de amostra e contar o número de suas colunas (sem contar a coluna com termos explicatórios, que é "Versão" no caso, e a coluna "Total"); ai então

$$q = \text{o número de colunas} - 1 \tag{7.29}$$

No presente caso, $q = 4 - 1 = 3$. Então, vamos á linia "3" da Tabela "Chi-Square Distribution Table". Na coluna $\chi^2_{.025}$ achamos $9,348$, que é o valor do limiar procurado no presente caso.

O último passo de solução é a comparação entre os valores χ^2_{obs} e χ^2_α e a dedução de conclusão final que vai de acordo com a seguinte regra:

$$\begin{array}{l} \text{se } \chi^2_{obs} \text{ está menor que } \chi^2_\alpha \text{ aceita-se } H \\ \text{se } \chi^2_{obs} \text{ está maior que ou igual a } \chi^2_\alpha \text{ rejeita-se } H \end{array} \tag{7.30}$$

Aproveito o momento para chamar sua atenção ao fato que a troca de nomes para os hipóteses faz com que a regra supracitada leve à decisão oposta àquela para a qual ela foi desenvolvida. Isso é uma das duas razões pelas quais você não pode errar na atribuição de nomes às hipóteses. (A segunda das duas razões é que com nomes alterados, você não conseguiria as frequencias esperadas, pois o cálculo dessas é feito sob a hipótese nula.)

No caso do presente exercício, $\chi^2_{obs} = 6,9167$ enquanto que $\chi^2_\alpha = 9,348$. Seguindo a regra (7.30), aceita-se a hipótese nula. É assim que soa a resposta na linguagem profana. Já na linguagem cientificamente correta, tal resposta coloca-se da seguinte maneira: não temos evidências suficientes na amostra, ao nível de significância de 2,5%, que a hipótese nula deve ser rejeitada, ou seja, que a campanha publicitária teve efeito nas vendas desse modelo de automóvel.

Solução do Exc. 159. Espero que sejá entendido que o enunciado orienta você a fazer o teste de aderência. Após se convencer disso, você começa a procurar pelas proporções/probabilidades em relação as quais a aderência deve ser testada. Tais probabilidades/proporções estão embutidas na distribuição binomial referida no enunciado. Em outras palavras, antes testar a aderência, é preciso calcular as probabilidades teóricas às quais a aderência deve ser testada com base na amostra fornecida (observe que o fato de ser "teórico" repercute-se na notação a ser usada abaixo). O cálculo está abaixo; ele usa (para facilitar o entendimento) a variável aleatória auxiliar X que tem a distribuição Bin$(4, 1/2)$.

k:	A probabilidade teórica de uma família com 4 filhos ter k meninos
0	$p_0^{teorica} = I\!P[X = 0] = 0,0625$
1	$p_1^{teorica} = I\!P[X = 1] = 0,25$
2	$p_2^{teorica} = I\!P[X = 2] = 0,375$
3	$p_3^{teorica} = I\!P[X = 3] = 0,25$
4	$p_4^{teorica} = I\!P[X = 4] = 0,0625$

Quanto ao resto de solução, noto que seu respaldo teórico está muito parecido com o da solução do Exc. 158, fato que permite-me omitir certos detalhes nas explicações a seguir.

As hipóteses apropriadas são:

H: o número de meninos por família, em famílias com 4 crianças, segue uma distribuição binomial com $n = 4$ e $p = 0,50$;

A: o número de meninos por família, em famílias com 4 crianças, não segue uma distribuição binomial com $n = 4$ e $p = 0,50$.

Ou, equivalentemente, mas com maior precisão:

H: $p_0^{populacional} = 0,0625$, $p_1^{populacional} = 0,2500$, $p_2^{populacional} = 0,3750$, $p_3^{populacional} = 0,2500$ e $p_4^{populacional} = 0,0625$;

A: pelo menos uma das igualdades acima não está válida,

em que $p_k^{populacional}$ denota a proporção das famílias que tem k meninos entre todas as famílias possuindo 4 crianças ($k = 0, 1, 2, 3, 4$).

Vamos à aplicação de Teste Qui-quadrado para o caso. A tabela abaixo dá as frequencias esperadas (denotadas por E_0, E_1, E_2, E_3, E_4) em amostra de tamanho 96 (especificamente falando, as frequencias que esperamos observar em 96 famílias, com 4 crianças cada, caso a hipótese nula fosse verdade e caso não havesse flutuações aleatórias):

Número de meninos (k)	Frequencia observada O_k	Frequencia esperada E_k
0	$O_0 = 12$	$E_0 = 6 = 96 \times p_0^{teorica} = 96 \times 0,0625$
1	$O_1 = 30$	$E_1 = 24 = 96 \times p_1^{teorica} = 96 \times 0,25$
2	$O_2 = 24$	$E_2 = 36 = 96 \times p_2^{teorica} = 96 \times 0,375$
3	$O_3 = 21$	$E_3 = 24 = 96 \times p_3^{teorica} = 96 \times 0,25$
4	$O_4 = 9$	$E_4 = 6 = 96 \times p_4^{teorica} = 96 \times 0,0625$

Diretamente da tabea temos

$$\chi^2_{obs} = \sum_{i=0}^{4} \frac{(O_i - E_i)^2}{E_i} = \frac{(12-6)^2}{6} + \frac{(30-24)^2}{24} + \frac{(24-36)^2}{36} + \frac{(21-24)^2}{24} + \frac{(9-6)^2}{6}$$
$$= 6 + 1,5 + 4 + 0,375 + 1,5 = 13,3750.$$

O numero de graus de liberdade (q) da distribuição qui-quadrado apropriada para o caso é $5 - 1 = 4$. O valor que corta dessa distrubuição cauda de tamanho 5% é 9,488, quer dizer, $\chi^2_{0,05} = 9,488$.

Conclusão: como $\chi^2_{obs} > \chi^2_{0,05}$ devemos rejeitar H, quer dizer, a conclusão é que a amostra não dá evidências suficientes para que possamos consluir, ao nível de signiicância 5%, que no âmbito populacional, quer dizer, no mundo de todas as famílias com 4 crianças, o número de meninos por família segue uma distribuição binomial com $n = 4$ e $p = 0,50$.

Resultado de Exc. 160. Minhas contas deram o seguinte: $\chi^2_{obs} = 15,7718$, $q = (2-1) \times (3-1) = 2$ e $\chi^2_{0,005} = 10,597$. Como $\chi^2_{obs} > \chi^2_{0,005}$ então a hipótese de independência está rejeitada pelo Teste Qui-quadrado, ao nível de significância 0,5%.

Solução de Exc. 162. O exercício apresenta uma dificuldade que não tem nada a ver com o Qui-quadrado: é a construção de tabela de contingência a partir de dados parciais fornecidos no enunciado do exercício. Eis os valores fornecidos:

Enquanto que a tabela completa está abaixo. Confirme para si que os valores nela presentes estão de fato unicamente definidos pelos valores presentes na tabela acima.

	Macrossomia		
Ganho de Peso	não	sim	total
$< 25\%$ peso gravitício	90		
$> 25\%$ peso gravitício			50
total		53	171

	Macrossomia		
Ganho de Peso	não	sim	total
$< 25\%$ peso gravitício	90	31	121
$> 25\%$ peso gravitício	28	22	50
total	118	53	171

Agora vamos à aplicação do Teste Qui-quadrado. Em primeiro lugar, reformulamos a pergunta do exercício em forma de "competição" de duas hipóteses. Preste a atenção: aquela das duas que aforma a independência deve receber o nome "nula" (e denotada por H de acordo com nosso acordo segundo o qual H leia-se "nula" e A leia-se "alternativa"):

H: O ganho de peso da mãe durante a gestação e a ocorrência de macrossomia são variáveis independentes.

A: O ganho de peso da mãe durante a gestação e a ocorrência de macrossomia são variáveis dependentes.

No segundo passo de solução, é preciso calcular as frequencias esperadas (precisamente: as frequencias que veriam na amostra de 171 pessoas entre as parturientes do HU caso a hipótese nula fose válida e caso não houvesse flutuações açeatórias). Essas frequencias são os valores colocados entre parênteses na tabela abaixo; ao lado de cada um delas, está a frequencia observada (a mesma que a já vista na tabela acima).

	Macrossomia		
Ganho de Peso	não	sim	total
$< 25\%$ peso gravitício	90 (83,50)	31 (37,50)	121
$> 25\%$ peso gravitício	28 (34,50)	22 (15,50)	50
total	118	53	171

No terceiro passo de solução calcula-se

$$\begin{aligned}\chi^2_{obs} &= \tfrac{(90-83{,}50)^2}{83{,}50} + \tfrac{(31-37{,}50)^2}{37{,}50} + \tfrac{(28-34{,}50)^2}{34{,}50} + \tfrac{(22-15{,}50)^2}{15{,}50} \\ &= 0,506461 + 1,127593 + 1,225636 + 2,728774 = 5,5884.\end{aligned}$$

No quarto passo acha-se $\chi^2_{0,1}$. Para tal, calcula-se primeiramente o número de graus de liberdade q:

$$q = (\text{número de linhas da tabela} - 1) \times (\text{número de colunas da tabela} - 1) = (2-1)(2-1) = 1$$

(Observe que contam-se só as linhas que representam as frequencias; não se contam a linha e a coluna que contém os totais.) Agora, usando a linha "df=1" da tabela "Chi-Square Distribution Table", achamos o limiar correspondente ao desejado valor 10%: $\chi^2_{0,1} = 2,706$.
Conclusão: como $\chi^2_{obs} > \chi^2_{0,1}$ conclua-se que os dados não sugerem, ao nível de significância de 10%, que no universo das parturientes do HU, o ganho do peso da mãe durante a gestação esteja independente da ocorrência de macrossomia.

Solução de Exc. 163 (a). As hipóteses são

H: O período de confecção de artefatos e o sítio arqueológico são variáveis independentes.

A: O período de confecção de artefatos e o sítio arqueológico são variáveis dependentes.

Os dados fornecidos no enunciado do exercício estão apresentados na tabela a seguir em forma de frequencias observadas e esperadas (as frequencias esperadas estão entre parênteses):

	Período			
Sítio	A	B	C	Total
sítio 1	15 (10,45)	15 (12,825)	8 (14,725)	38
sítio 2	7 (11,55)	12 (14,175)	23 (16,275)	42
Total	22	27	31	80

Os dados da tabela dão: $\chi^2_{obs} = 10,326$.

O número de graus de liberdade, q, é igual a (número de linhas -1)$\times$(número de colunas -1) $(2-1)(3-1) = 2$. Para a distribuição Qui-quadrado com $q = 2$, tem-se: $\chi^2_{0,1} = 4,605$.
Conclusão: Como $\chi^2_{obs} > \chi^2_{0,1}$ devemos rejeitar H, isto é, os dados de estudo não apresentam evidências para confirmar, ao nível de significância 10%, que o período de confecção de artefatos e o sítio arqueológico são variáveis independentes.

Uma observação partes da solução de Exc. 165. Na tabela abaixo, há frequecias observadas e esperadas no formato O_{ij} (E_{ij}):

asma	gripe		total
	sim	não	
sim	27 (28,06)	34 (32,94)	61
não	42 (40,94)	47 (48,06)	89
total	69	81	150

Recorde que $\chi^2_{obs} = \sum_{i=1}^{2}\sum_{j=1}^{2}\frac{(O_{ij}-E_{ij})^2}{E_{ij}}$. Eu agora gostaria de apresentar o resultado do passo intermediário desse calculo, quer dizer, os valores de $O_{ij} - E_{ij}$:

asma	gripe		total
	sim	não	
sim	-1,06	1,06	61
não	1,06	-1,06	89
total	69	81	150

A igualdade que você observa entre os valores absolutas das diferenças presentes na tabela não é uma mera coincidência. A igualdade é uma regra, que diz que, no caso de tabela 2×2, ocorre que uma das diferenças determina todas as outras três (até o sinal). Então, no caso de tabela 2×2, podemos dizer que as quatro diferenças têm somente um grau de liberdade. Uma regra semelhante vale para qualquer tabela $n \times m$ (n linhas e m colunas, quer dizer) e alega que os valores de $(n-1)(m-1)$ diferenças determinam unicamente (com até o sinal) os valores das demais diferenças. Isso é o motivo para chamar a expressão $(n-1)(m-1)$ por **número de graus de liberdade**.

Alguns partes da solução de Exc. 166. Na tabela abaixo, há frequecias observadas e esperadas no formato O_{ij} (E_{ij}):

Estado civil	Preferência Policial	Comédia	Romance	Total
Solteiro	40 (34,4375)	25 (30,1625)	30 (30,4)	95
Casado	26 (32,625)	31 (28,575)	33 (28,8)	90
Divorciado	46 (39,875)	33 (34,925)	31 (35,2)	110
Viúvo	33 (38,0625)	38 (33,3375)	34 (33,6)	105
Total	145	127	128	400

$q = (4-1)(3-1) = 6$, $\chi^2_{0,025} = 14,449$, $\chi^2_{obs} = 6,72309$.

Como $\chi^2_{obs} < \chi^2_{0,025}$, conclua-se que os dados indicam, ao nível de significância de $2,5\%$, que o estado civil e a preferência pelo tipo de filme (na classificação policial/comédia/romance) são fatores independentes.