

AULA 18

Experimentos Multinomiais e Tabelas de Contingência

Ernesto F. L. Amaral

14 de outubro de 2010
Metodologia de Pesquisa (DCP 854B)

Fonte:

Triola, Mario F. 2008. “Introdução à estatística”. 10^a ed. Rio de Janeiro: LTC. Capítulo 11 (pp.468-505).

ESQUEMA DA AULA

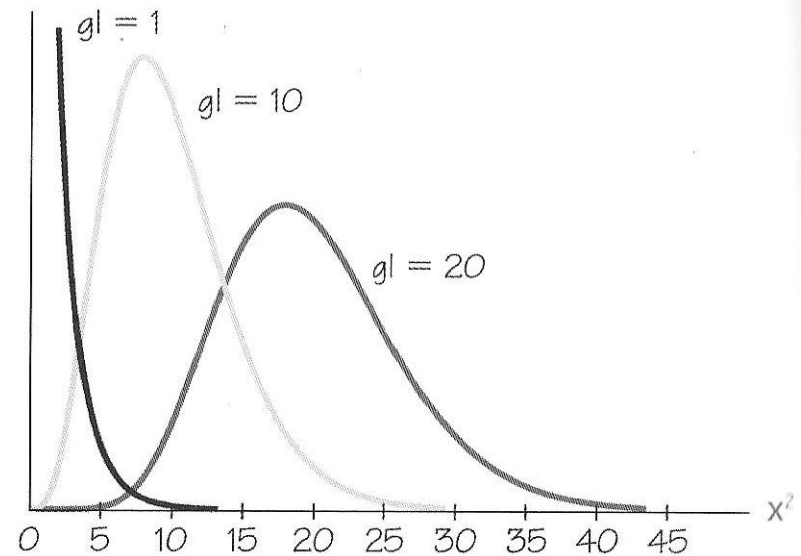
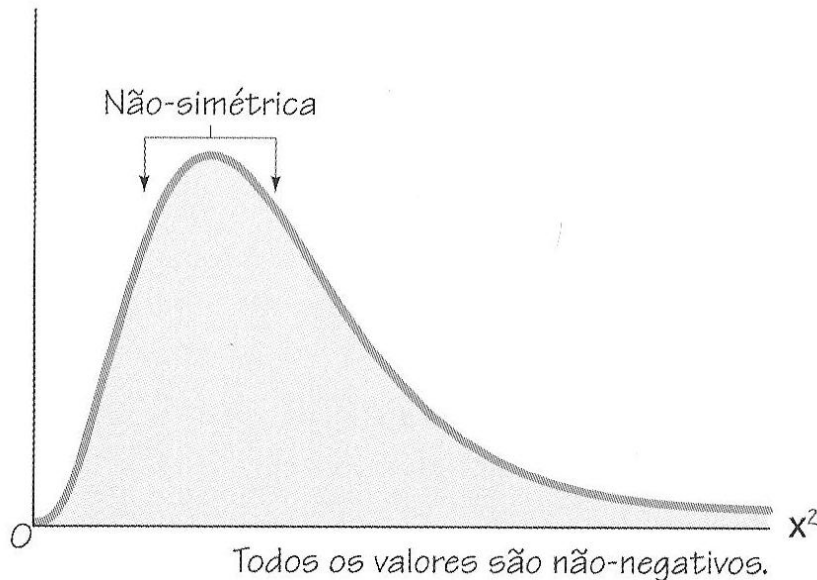
- Experimentos multinomiais: aderência.
- Tabelas de contingência: independência e homogeneidade.
- Teste de McNemar para dados emparelhados.

VISÃO GERAL

- Tratar de dados categóricos (ou qualitativos ou de atributo) que podem ser separados em diferentes células.
- **Objetivo** é testar afirmativas sobre dados categóricos que consistem em contagem de freqüências para as categorias:
 - **Experimentos multinomiais:** contagens de freqüências observadas, arranjadas em uma única linha ou coluna (tabela de freqüência de entrada única) para verificar se tais contagens seguem alguma distribuição alegada.
 - **Tabelas de contingência:** contagens de freqüência arranjadas em uma tabela com, no mínimo, 2 linhas e 2 colunas.
 - **Tabelas de dupla entrada** que envolvem dados emparelhados.

DISTRIBUIÇÃO DE QUI-QUADRADO

- É utilizada a distribuição de qui-quadrado que possui as seguintes propriedades:
 - Não é simétrica.
 - Valores da distribuição podem ser 0 ou positivos, mas não podem ser negativos.
 - É diferente para cada número de graus de liberdade.



EXPERIMENTOS MULTINOMIAIS: ADERÊNCIA

TESTE DE HIPÓTESE

- O teste de hipótese usará a distribuição qui-quadrado com as contagens de frequências observadas e as contagens de frequências esperadas.
- Ou seja, a estatística de teste qui-quadrado é uma medida de discrepância entre as frequências observadas e esperadas.
- O experimento multinomial possui mais de duas categorias, enquanto o experimento binomial tem exatamente duas categorias.

EXPERIMENTO MULTINOMIAL

- Este experimento satisfaz as seguintes condições:
 - Número de tentativas é fixo.
 - Tentativas são independentes.
 - Todos resultados de cada tentativa devem ser classificados em exatamente uma das várias diferentes categorias.
 - Probabilidades para diferentes categorias permanecem constantes para cada tentativa.
- É testada afirmativa de que freqüências observadas nas diferentes categorias se ajustam a uma distribuição alegada.

TESTE DE ADERÊNCIA

- O teste de aderência (bondade de ajuste) é usado para testar a hipótese de que uma distribuição de frequência observada se ajusta (ou concorda com) alguma distribuição teórica especificada.

- Notação:
 - ***O***: frequência observada de um resultado.
 - ***E***: frequência esperada de um resultado.
 - ***k***: número de diferentes categorias ou resultados.
 - ***n***: número de tentativas total.

ENCONTRANDO FREQUÊNCIAS ESPERADAS

- Se todas frequências esperadas **são iguais**:
 - Então cada frequência esperada é a soma de todas frequências observadas dividida pelo número de categorias.
 - $E=n/k$.
- Se as frequências esperadas **não são todas iguais**:
 - Então cada frequência esperada é encontrada multiplicando-se a soma de todas frequências observadas pela probabilidade da categoria.
 - $E=np$ para cada categoria.

ALGUMAS CONSIDERAÇÕES

- **Freqüências observadas** têm que ser números inteiros (contagens reais), mas **freqüências esperadas** não precisam ser números inteiros.
- **Freqüências amostrais** comumente se desviam um pouco dos valores teoricamente esperados.
- Devemos testar se as diferenças entre os valores reais observados (**O**) e os valores teoricamente esperados (**E**) são estatisticamente significativos.

REQUISITOS PARA TESTE DE DIFERENÇAS

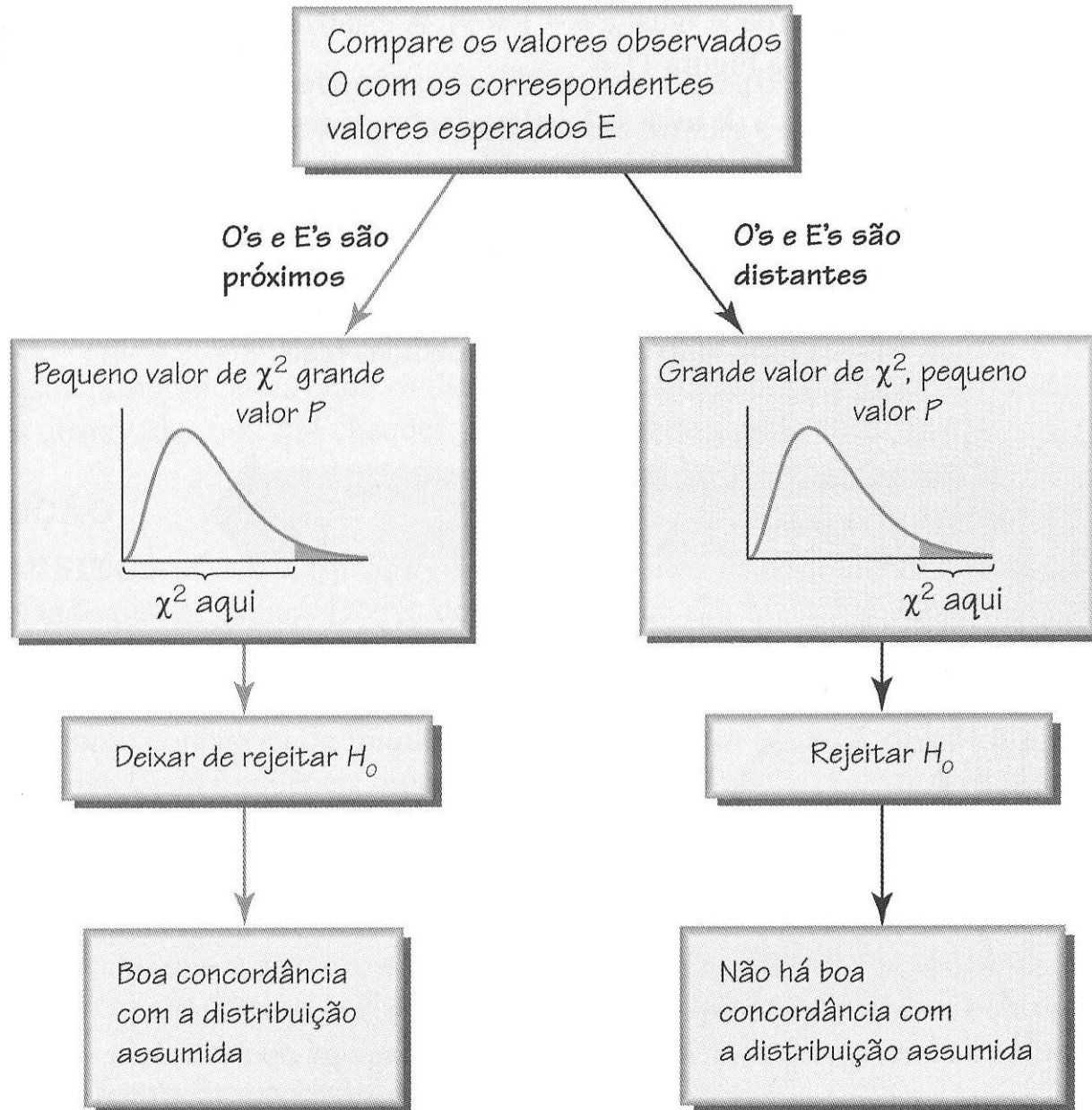
- Dados selecionados aleatoriamente.
- Dados amostrais consistem em contagens de frequências para cada uma das diferentes categorias.
- Para cada categoria, frequência esperada é, no mínimo, 5.
- Valores críticos são encontrados usando-se graus de liberdade $(k-1)$ específicos, sendo k o número de categorias.
- Estes testes são sempre unilaterais à direita.
- Estatística de teste para testes de aderência em experimentos multinomiais:

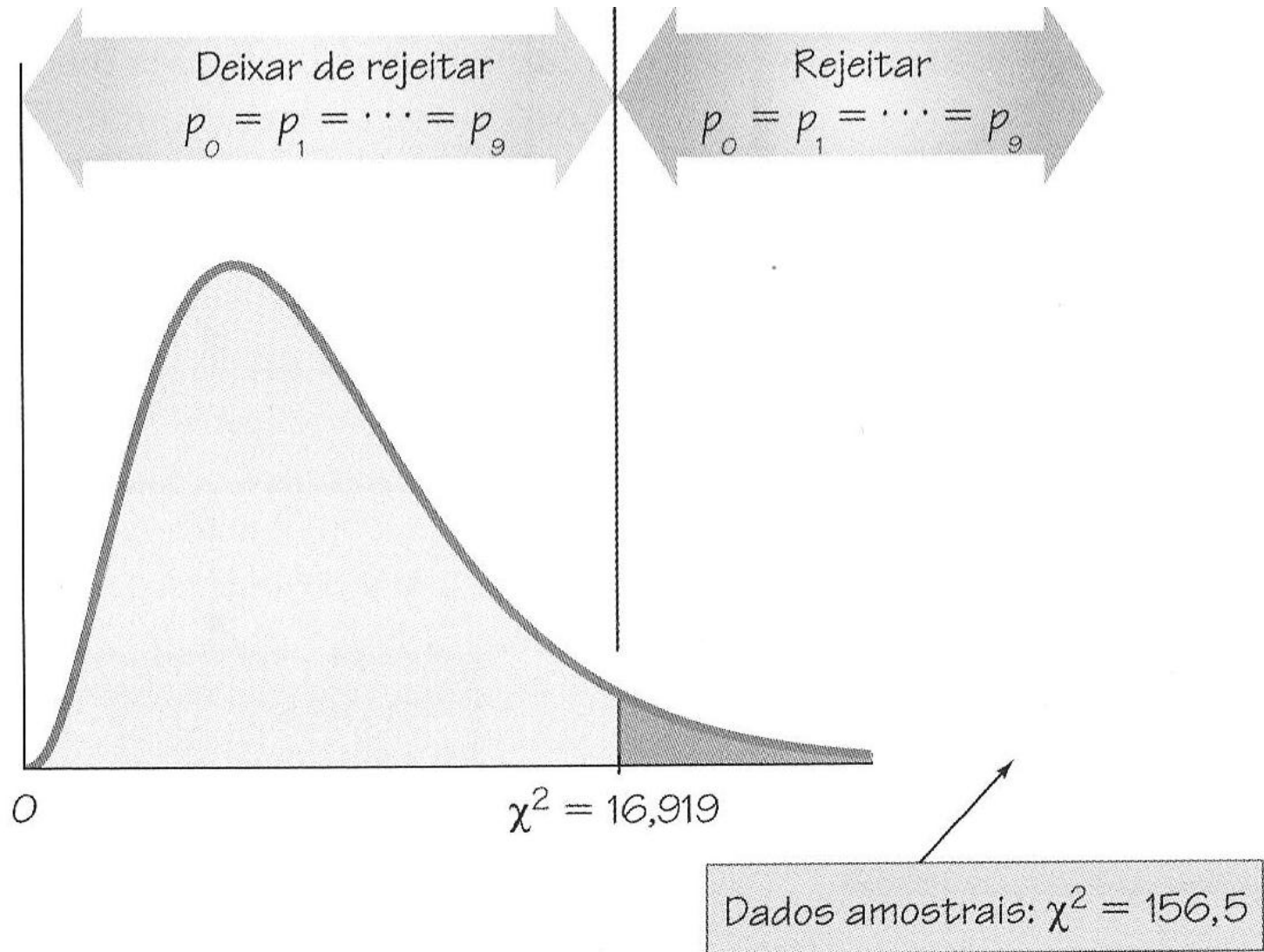
$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

RESULTADOS DA ESTATÍSTICA DE TESTE

- A estatística de teste χ^2 se baseia nas diferenças entre valores observados e esperados.
- Uma concordância entre valores observados e esperados levará a um pequeno valor de χ^2 e a um grande valor P .
- Uma discrepância entre valores observados e esperados levará a um grande valor de χ^2 e a um pequeno valor P .
- O valor crítico e a região crítica se localizam no extremo direito da distribuição (unilateral à direita).

RELAÇÕES ENTRE χ^2 , VALOR P E ADERÊNCIA

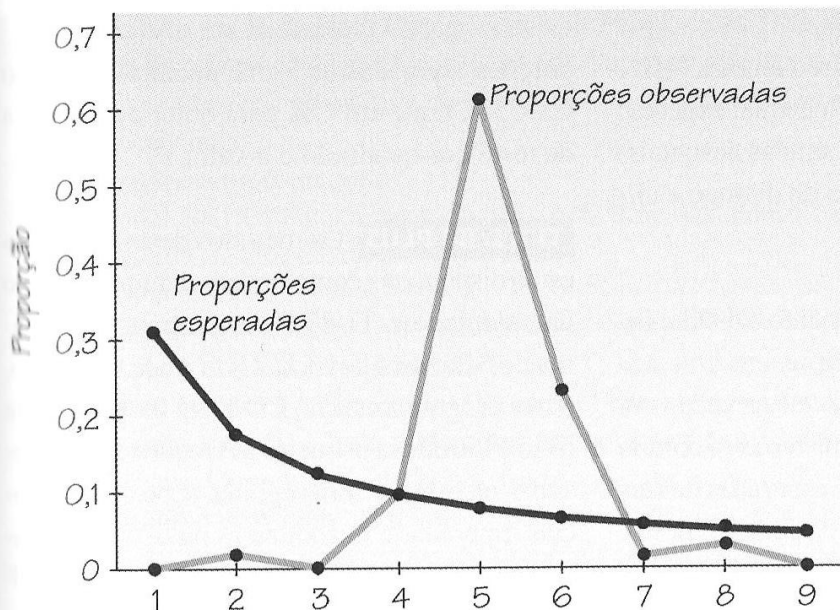


EX.: TESTE DE $p_0 = p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = p_7 = p_8 = p_9$ 

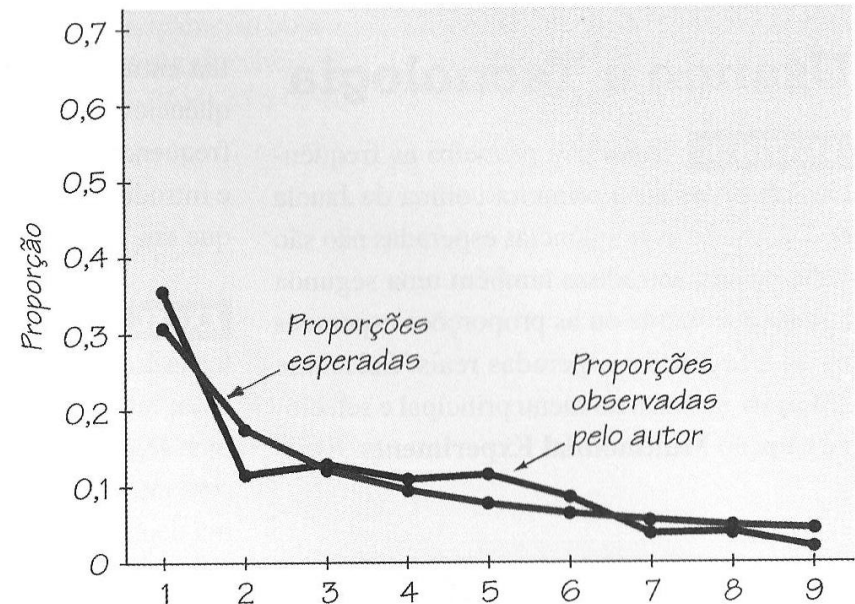
– Graus de liberdade: $k - 1 = 10 - 1 = 9$

EX.: COMPARAÇÃO DE FREQUÊNCIAS

- Gráficos como este abaixo são úteis na comparação visual de frequências esperadas e observadas, bem como na sugestão de quais categorias resultam principais diferenças.



(a) DÍgitos Líder



(b) DÍgitos Líder

VALORES P

- A abordagem do valor P também pode ser usada.
- Os valores P são fornecidos automaticamente por programas estatísticos.

FUNDAMENTOS PARA ESTATÍSTICA DE TESTE

- Mede-se diferença de freqüências observadas e esperadas.
- **Simple soma** das diferenças entre valores observados e esperados não é eficaz, porque soma é sempre zero.
- **Elevação ao quadrado** dos valores de $(O-E)$ fornece uma estatística melhor (como no caso do desvio padrão).
- $\sum(O-E)^2$ mede a magnitude das diferenças.
- $\sum(O-E)^2/E$ mede magnitude das diferenças em relação ao esperado. Distribuição pode ser aproximada pela distribuição χ^2 que é contínua.
- **Graus de liberdade** indicam número de categorias que podemos inferir freqüências, antes que estas sejam determinadas para todas categorias.

TABELAS DE CONTINGÊNCIA: INDEPENDÊNCIA E HOMOGENEIDADE

TABELAS DE CONTINGÊNCIA

- **Tabela de contingência** (ou tabela de frequência de dupla entrada) é uma tabela na qual as frequências correspondem a duas variáveis (linhas e colunas).
- Estas tabelas incluem contagens de frequência para dados categóricos arranjados em uma tabela com pelo menos 2 linhas e 2 colunas.
- **Testes de independência** são usados para determinar se uma variável linha de uma tabela de contingência é independente de sua variável coluna.
- **Testes de homogeneidade** são usados para determinar se populações diferentes têm as mesmas proporções de alguma característica.

TESTE DE INDEPENDÊNCIA

- Um **teste de independência** testa a hipótese nula de que não há associação entre a variável linha e a variável coluna em uma tabela de contingência.
- **Hipótese nula:** variáveis linha e coluna são independentes.

REQUISITOS

- Dados amostrais são selecionados aleatoriamente e são representados como contagens de freqüências em tabela de dupla entrada.
- **Hipótese nula** (H_0) é a afirmativa de que variáveis linha e coluna são **independentes**.
- **Hipótese alternativa** (H_1) é a afirmativa de que as variáveis linha e coluna são **dependentes**.
- Em toda célula da tabela, a freqüência esperada (E) é no mínimo 5.
- Não há exigência quanto à freqüência observada (O).
- Não há exigência de que população deva ter distribuição normal ou qualquer outra.

ESTATÍSTICA DE TESTE

- Estatística de teste para um teste de independência:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

- Permite medir grau de discordância entre freqüências observadas e as teoricamente esperadas, quando as duas variáveis são independentes.
- Grandes valores da estatística de teste refletem diferenças significativas entre as freqüências observadas e esperadas.

VALORES CRÍTICOS

- **Valores críticos** são encontrados com graus de liberdade $[(r-1)(c-1)]$, em que r é o número de linhas e c é o número de colunas.
- Ao saber total de todas freqüências, podemos **associar livremente** freqüências a apenas $r-1$ linhas e a $c-1$ colunas, antes que as freqüências de todas células sejam determinadas.
- Porém, não podemos ter freqüências negativas ou freqüências tão grandes que a soma de qualquer linha (ou coluna) exceda total das freqüências observadas.
- Em um teste de independência com uma tabela de contingência, a região crítica se localiza apenas na **cauda direita**.

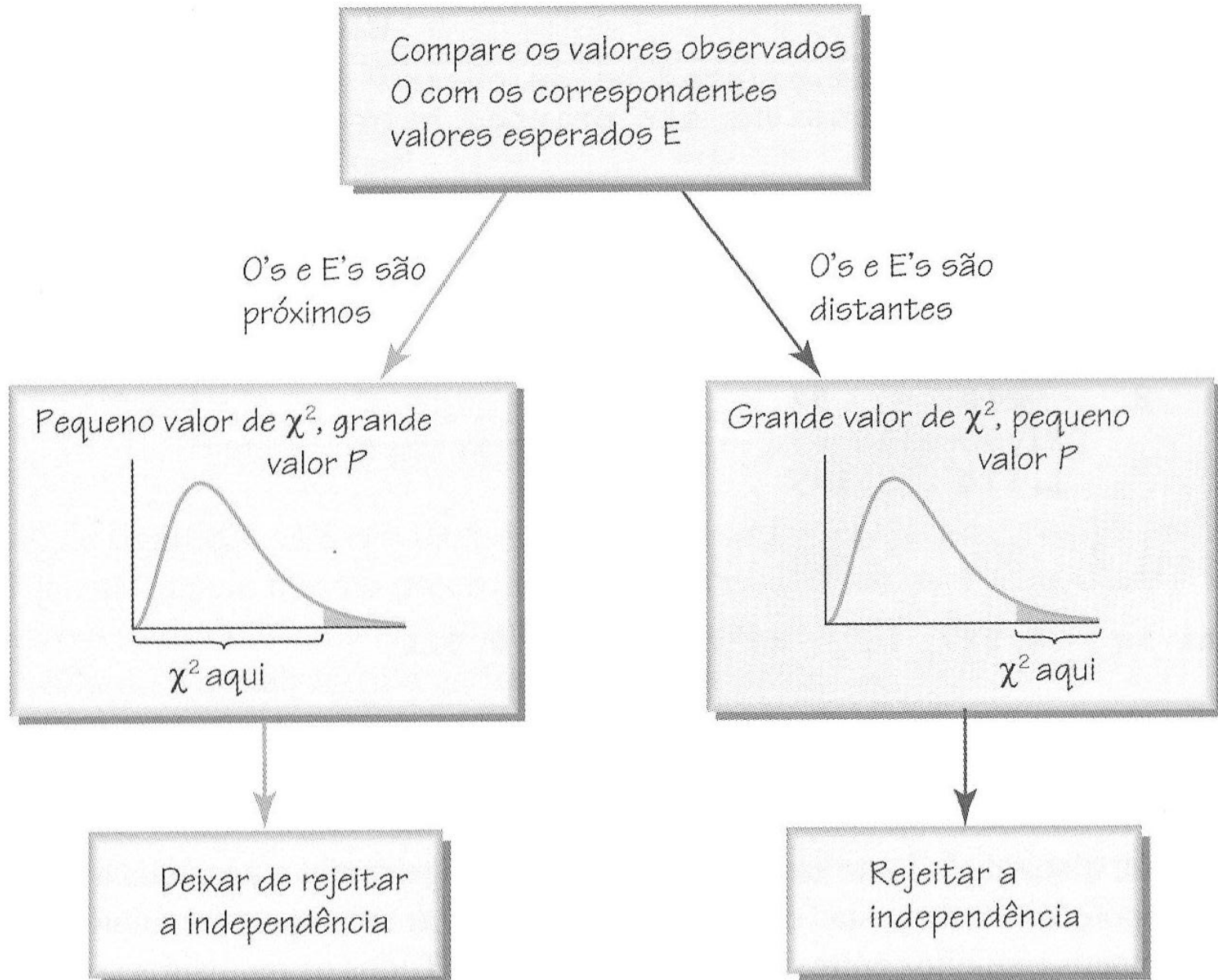
FREQÜÊNCIA ESPERADA PARA UMA CÉLULA

- Freqüência esperada (E) pode ser calculada para cada célula:
 - Multiplicando-se o total das freqüências das linhas pelo total das freqüências das colunas.
 - Dividindo-se o resultado pelo grande total das freqüências.

$$\text{freqüência esperada} = (\text{total geral}) * \frac{(\text{total da linha})}{(\text{total geral})} * \frac{(\text{total da coluna})}{(\text{total geral})}$$

$$E = \text{freqüência esperada} = \frac{(\text{total da linha}) * (\text{total da coluna})}{(\text{total geral})}$$

COMPONENTES-CHAVE NO TESTE DE INDEPENDÊNCIA



TESTE DE HOMOGENEIDADE

- Amostras podem ser extraídas de **populações diferentes** e desejamos determinar se essas populações têm as mesmas proporções da característica em consideração.
- Em um **teste de homogeneidade**, testamos a afirmativa de que populações diferentes têm a mesma proporção de alguma característica.
- Ao realizar um teste de homogeneidade, podemos usar **mesmos** requisitos, estatística de teste, valor crítico e demais procedimentos já apresentados.
- **Exceção** é que em vez de testar a hipótese nula de independência entre as variáveis linha e coluna, testamos a hipótese nula de que as diferentes populações têm as mesmas proporções de alguma característica.

TESTE EXATO DE FISHER

- Para análise de tabelas 2 x 2, incluímos requisito de que toda célula deve ter frequência esperada de, no mínimo, 5.
- Esse requisito é necessário para que a distribuição χ^2 seja uma aproximação adequada para a distribuição exata da estatística de teste.
- O **teste exato de Fisher** é usado para tabelas 2 x 2, porque fornece valor P exato e não exige técnica de aproximação.
- Ou seja, é útil para casos em que frequência esperada em alguma célula é menor do que 5.
- Geralmente é estimado com programas computacionais.

INTENÇÃO DE VOTO PARA PRESIDENTE

– Datafolha (10/10/2010), margem de erro ($\pm 2\%$):

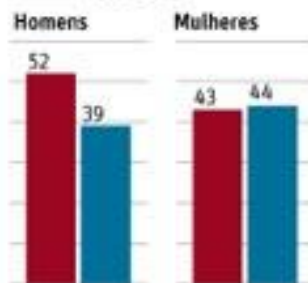
DATAFOLHA ESTRATIFICAÇÃO DO ELEITORADO

Intenção de voto para presidente

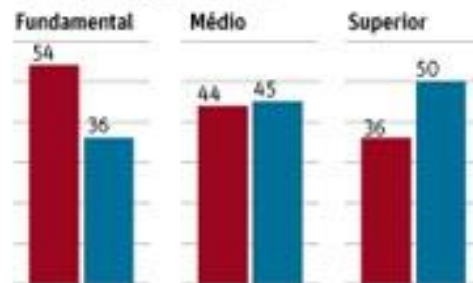
Dilma Serra



POR SEXO



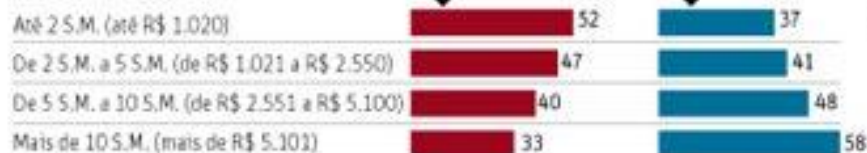
POR ESCOLARIDADE



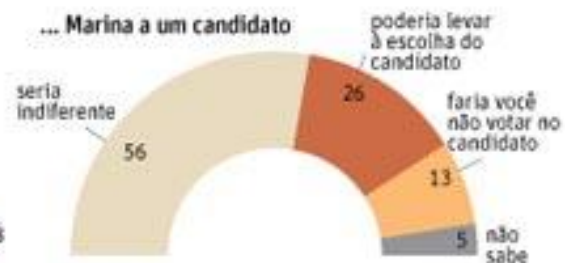
O APOIO DE...
... Lula a um candidato



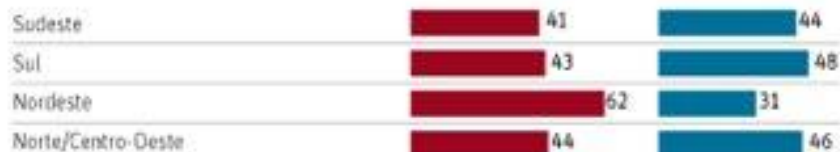
POR RENDA



... Marina a um candidato



POR REGIÃO



QUEM MARINA DEVERIA APOIAR NO 2º TURNO



QUEM MARINA VAI APOIAR NO 2º TURNO



INTENÇÃO DE VOTO PARA PRESIDENTE (%)

– Valores observados:

Sexo	Dilma	Serra	Total
Homem	52	39	91
Mulher	43	44	87
Total	95	83	178

– Valores esperados:

Sexo	Dilma	Serra	Total
Homem	$(95 \cdot 91) / 178 =$ 48,57	$(83 \cdot 91) / 178 =$ 42,43	91
Mulher	$(95 \cdot 87) / 178 =$ 46,43	$(83 \cdot 87) / 178 =$ 40,57	87
Total	95	83	178

$$- \chi^2 = (52-48,57)^2/48,57 + \dots + (44-40,57)^2/40,57 \approx 1,063$$

$$- gl = (r-1)(c-1) = (2-1)(2-1) = 1$$

– Tabela A-4 (pág.621): valor crítico = 3,841; $\alpha = 0,05$.

NO STATA

– Entrando com dados:

```
clear
input mulher dilma votos
0 1 52
1 1 43
0 0 39
1 0 44
end
```

– Teste de qui-quadrado:

```
tab mulher dilma [fweight=votos], chi2
```

mulher	dilma		Total
	0	1	
0	39	52	91
1	44	43	87
Total	83	95	178

Pearson chi2(1) = 1.0645 Pr = 0.302

TESTE DE MCNEMAR PARA DADOS EMPARELHADOS

TESTE DE MCNEMAR

- Procedimentos para a tabela de contingência, apresentados anteriormente, se baseiam em **dados independentes**.
- Para tabelas 2 x 2 que consistem em contagens de freqüência que resultam de **dados emparelhados**, não temos independência.
- **Teste de McNemar** usa contagens de freqüências de dados emparelhados nominais com duas categorias.
- É testada **hipótese nula** de que as freqüências das diferentes categorias ocorrem na mesma proporção.
- Freqüências b e c ocorrem na mesma proporção?

		Tratamento X	
		Curado	Não curado
Tratamento Y	Curado	a	b
	Não curado	c	d

REQUISITOS

- Dados amostrais foram selecionados **aleatoriamente**.
- Temos **dados emparelhados** de contagens de freqüências.
- Dados são do **nível nominal** de mensuração.
- Cada observação pode ser classificada de duas maneiras (**duas variáveis dicotômicas**).
- Freqüências são tais que **$b + c \geq 10$** .
- **Estatística de teste** para hipótese nula de que freqüências b e c ocorrem na mesma proporção:

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

- **Região crítica** se localiza apenas na cauda direita.
- **Graus de liberdade** = 1 (Tabela A-4).

APLICAÇÕES DO TESTE DE MCNEMAR

- Em vez de usarmos resultados “cura/cura” e resultados “não cura/não cura”, usamos apenas os resultados “cura/não cura” e “não cura/cura”.
- Isto é, usamos apenas resultados de categorias diferentes.
- **Pares discordantes** de resultados são provenientes de pares de categorias nas quais as duas categorias são diferentes.
- Além da comparação de tratamentos atribuídos a dados emparelhados, o teste de McNemar é usado para testar a hipótese nula de nenhuma mudança em **experimentos do tipo antes/depois**.