Tabela 2x2
Estudos Epidemiológicos

Hipóteses

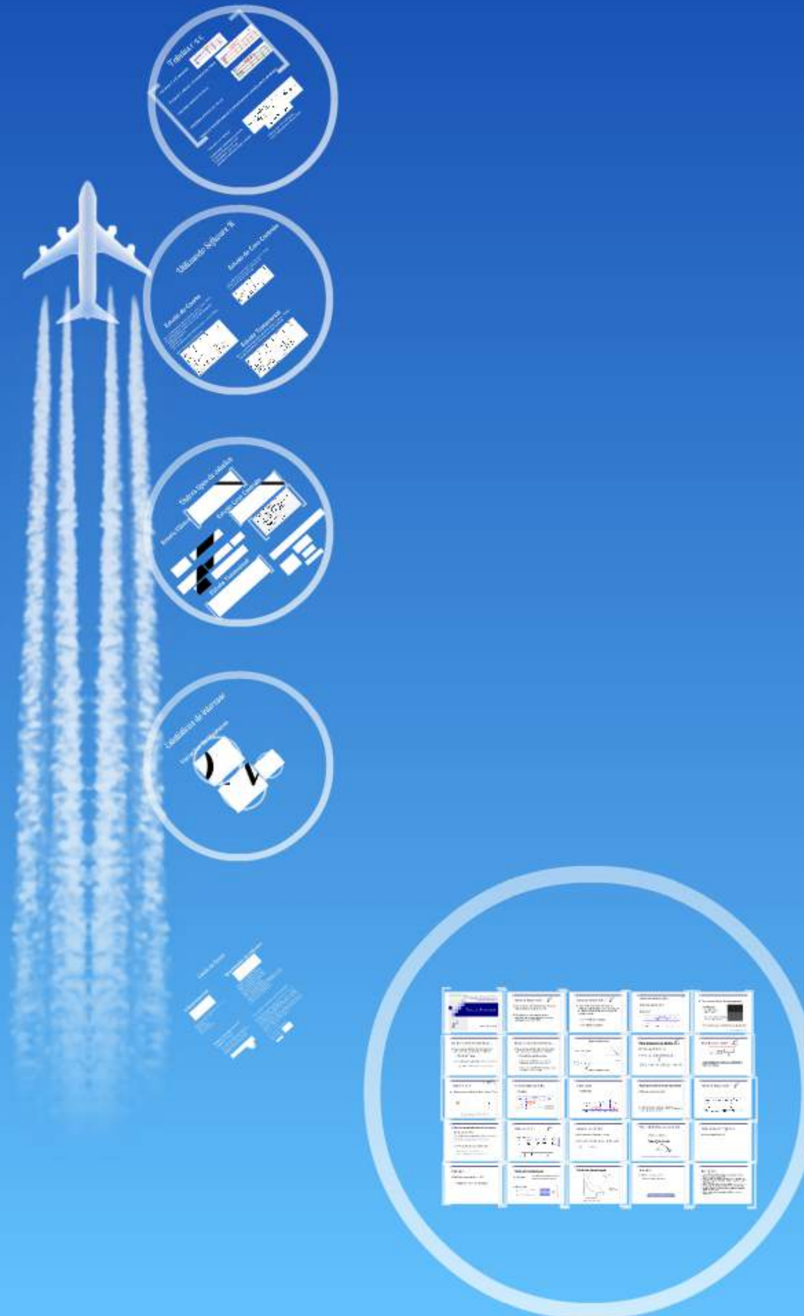
Testes

Medidas de associação

Prezi

Tabela 2x2
Estudos Epidemiológicos

Hipóteses
Testes
Medidas de associação



Estudo de Coorte

O Delineamento

	Categoria da Variável Y		
Categoria da Variável X	j (1)	j (2)	Total
i (1)	$n11\ a$	$n12\ b$	$n1.\ a+b$
i (2)	$n21\ b$	$n22\ d$	$n2.\ c+d$
Total	$n.1\ a+c$	$n.2\ b+d$	$n\ a+b+c+d$

Frequencia - nij - numero de individuos na categoria i de X e categoria j de Y, onde i e j =1,2
Totais marginais
linha - frequencia ni.
coluna - frequencia n.j
Total geral - n soma dos ni

Quantidades de interesse

	Categoria da Variável Y		
Categoria da Variável X	D (1)	ND (2)	Total
E (1)	$p_{(1)1}$	$p_{(1)2}$	1
NE (2)	$p_{(2)1}$	$p_{(2)2}$	1
Total	p_1	p_2	1

- pij P(X=i,Y=j) probabilidade conjunta
- pi(j) P(X=i|Y=j) probabilidade condicional
- p(i)j P(Y=j|X=i) probabilidade condicional
- pi. P(X=i) probabilidades marginais linha
- p.j P(X=j) Probabilidades marginais coluna
- p(1)1=incidencia dos expostos
- p(2)1=incidencia não dos expostos
- N11 e variável aleatoria Binomial com parâmetros n1. e p(1)1
- N21 e variável aleatoria Binomial com parametros n2. e p(2)1
- N11 e o número de indivíduos D e E
- N21 e o número de indivíduos D e NE

Modelos para dados binarios
Tabela unica 2x2

- N11 e o número de indivıduos doentes expostos
- N21 e o número de indivíduos doentes não expostos

O modelo probabilístico é o produto de binomiais (Probabilidade conjunta de acontecer N11 e N21).

$P(N_{11}=n_{11}, N_{21}=n_{21}) = \binom{n_{1\cdot}}{n_{11}}\binom{n_{2\cdot}}{n_{21}}$ [illegible]

As probabilidades de p_{ij} são estimadas por:

$p_{(1)1}$ = incidência dos expostos [illegible]

$p_{(2)1}$ = incidência não dos expostos [illegible]

[illegible]

Vamos lembrar do teste Qui-quadrado de Pearson!!!

Karl Pearson, foi matemático britânico, nasceu a 27 de Março de 1857, em Londres. Formou-se na Universidade de Cambridge em Matemática, 1879.
Criou o teste do "qui-quadrado" , em 1900 para verificar a possibilidade de um ajustamento. O teste do qui-quadrado constitui a base da Estatística das pequenas amostras de populações normais, servindo para medir a confiança de resultados estatísticos, testar hipóteses, etc. Inventou o termo "desvio-padrão" (1893).

Como os dados experimentais podem variar de amostra para amostra, uma maneira sensata de avaliar quão grandes ou quão pequenas são as diferenças é utilizar o quadrado dos desvios e dividi-los por um valor estável, isto é, um valor que se mantenha constante em qualquer amostra. Esse valor, em geral, é dado pela H0.

Outras estatísticas utilizadas

O Delineamento

	Categoria da Varáivel Y		
Categoria da Variável X	j (1)	j (2)	Total
i (1)	$n11$ a	$n12$ b	$n1.$ $a+b$
i (2)	$n21$ b	$n22$ d	$n2.$ $c+d$
Total	$n.1$ $a+c$	$n.2$ $b+d$	n $a+b+c+d$

Frequencia - nij - numero de indivıduos na categoria i de X e categoria j de Y, onde i e j =1,2

Totais marginais

linha - frequencia ni.

coluna - frequencia n.j

Total geral - n soma dos ni

Quantidades de interesse

	Categoria da Variável Y		
Categoria da Variável X	D (1)	ND (2)	Total
E (1)	$p_{(1)1}$	$p_{(1)2}$	1
NE (2)	$p_{(2)1}$	$p_{(2)2}$	1
Total	$p_{.1}$	$p_{.2}$	1

- pij P(X=i,Y=j) probabilidade conjunta
- pi(j) P(X=i|Y=j) probabilidade condicional
- p(i)j P(Y=j|X=i) probabilidade condicional
- pi. P(X=i) probabilidades marginais linha
- p.j P(X=j) Probabilidades marginais coluna
- p(1)1=incidencia dos expostos
- p(2)1=incidencia não dos expostos
- N11 e variável aleatoria Binomial com parâmetros n1. e p(1)1
- N21 e variável aleatoria Binomial com parametros n2. e p(2)1
- N11 e o número de indivíduos D e E
- N21 e o número de indivíduos D e NE

Modelos para dados binarios Tabela unica 2x2

- N11 e o número de indivıduos doentes expostos
- N21 e o número de individuos doentes não expostos

O modelo probabilístico é o produto de binomiais (Probabilidade conjunta de acontecer N11 e N21).

$$P(N_{11}=n_{11},N_{21}=n_{21})=\binom{n_{1.}}{n_{11}}\binom{n_{2.}}{n_{21}}p_{(1)1}^{n_{1.}}(1-p_{(1)1})^{n_{1.}-n_{11}}p_{(2)1}^{n_{2.}}(1-p_{21)1})^{n_{2.}-n_{21}}$$

As probabilidades de p_{ij} são estimadas por:

$p_{(1)1}$ = incidência dos expostos $\frac{a}{a+c}$ ou $\hat{p}_{(i)j}=\frac{n_{ij}}{n_{i.}}$.

$p_{(2)1}$ = incidência não dos expostos $\frac{c}{a+c}$ ou $\hat{p}_{(2)1}=\frac{n_{21}}{n_{2.}}$.

$$\begin{cases} H_0 : p_{(1)1} = p_{(2)1} = p_{.1}; \\ H_a : p_{(1)1} \neq p_{(2)1}. \end{cases}$$

$$Tq=\sum_{i=1}^{2}\sum_{j=1}^{2}\frac{(n_{ij}-e_{ij})^2}{e_{ij}}\sim\chi^2_{(1)}$$

$$E[N_{i1}]=n_{i.}p_{.1}$$

$$E[N_{i2}]=n_{i.}p_{.2}$$

$$e_{ij}=\frac{n_{i.}n_{.j}}{n}$$

- N21 é o número de indivíduos doentes não expo

- O modelo probabilístico é o produto de binomiais (Probabilidade conjunta de acontecer N11 e N21).

$$P(N_{11}=n_{11},N_{21}=n_{21})=\binom{n_{1.}}{n_{11}}\binom{n_{2.}}{n_{21}}p_{(1)1}^{n_{1.}}(1-p_{(1)1})^{n_{1.}-n_{11}}p_{(2)1}^{n_{2.}}(1-p_{21)1})^{n_{2.}-n_{21}}$$

As probabilidades de p_{ij} são estimadas por:

$p_{(1)1}$ = incidência dos expostos $\frac{a}{a+c}$ ou $\hat{p}_{(i)j}=\frac{n_{ij}}{n_{i.}}$.

$p_{(2)1}$ = incidência não dos expostos $\frac{c}{a+c}$ ou $\hat{p}_{(2)1}=\frac{n_{21}}{n_{2.}}$.

$$\begin{cases} H_0: p_{(1)1}=p_{(2)1}=p_{.1}; \\ H_a: p_{(1)1}\neq p_{(2)1}. \end{cases}$$

$$Tq=\sum_{i=1}^{2}\sum_{j=1}^{2}\frac{(n_{ij}-e_{ij})^2}{e_{ij}}\sim\chi^2_{(1)}$$

$$E[N_{i1}]=n_{i.}p_{.1}$$

$$E[N_{i2}]=n_{i.}p_{.2}$$

$$e_{ij}=\frac{n_{i.}n_{.j}}{n}$$

- N21 e o número de indivíduos D e NE

Vamos lembrar do teste Qui-quadrado de Pearson!!!

Karl Pearson, foi matemático britânico, nasceu a 27 de Março de 1857, em Londres. Formou-se na Universidade de Cambridge em Matemática, 1879.
Criou o teste do "qui-quadrado" , em 1900 para verificar a possibilidade de um ajustamento. O teste do qui-quadrado constitui a base da Estatística das pequenas amostras de populações normais, servindo para medir a confiança de resultados estatísticos, testar hipóteses, etc. Inventou o termo "desvio-padrão" (1893).

Como os dados experimentais podem variar de amostra para amostra, uma maneira sensata de avaliar quão grandes ou quão pequenas são as diferenças é utilizar o quadrado dos desvios e dividi-los por um valor estável, isto é, um valor que se mantenha constante em qualquer amostra. Esse valor, em geral, é dado pela H0.

Outras estatısticas utilizadas

Estatistica da Razão de Verossimilhança

Estatística de Neyman

$$Trv = -2\sum_{i=1}^{2}\sum_{j=1}^{2} n_{ij} log(\frac{e_{ij}}{n_{ij}} \sim \chi^2_{(1)}$$

qualquer amostra. Esse valor, em geral, é dado p

Outras estatísticas utilizadas

Estatistica da Razão de Verossimilhança

Estatística de Neyman

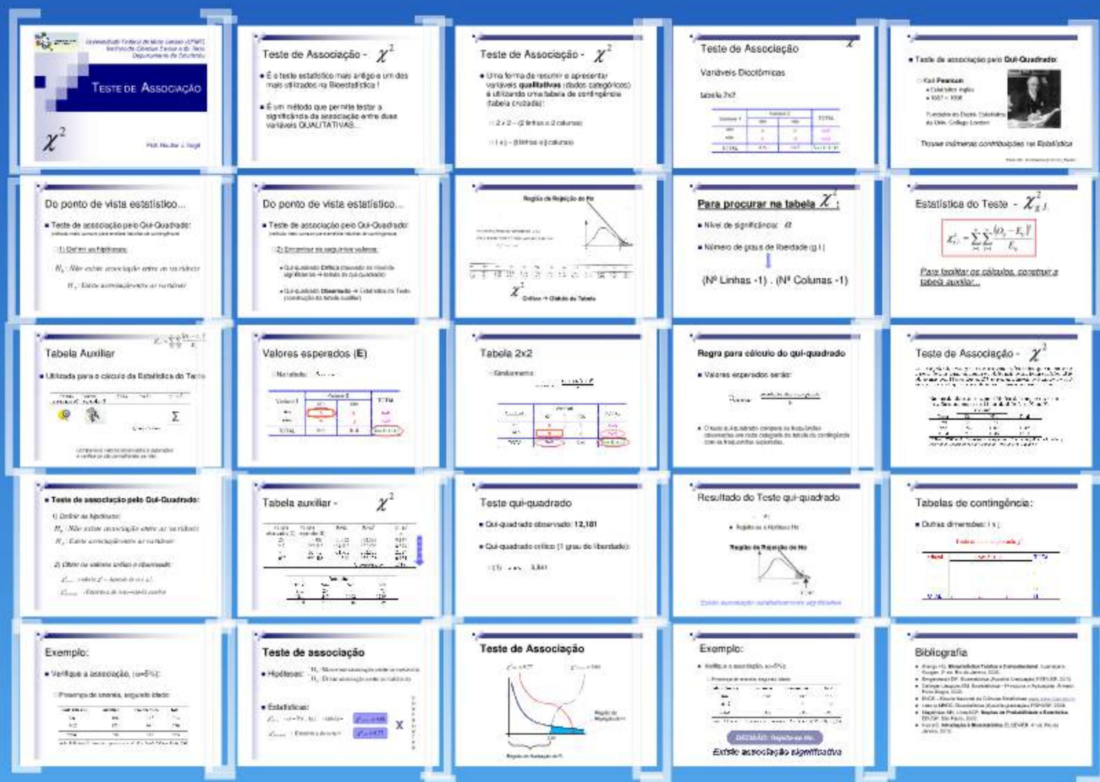
Teste exato de Fisher.

$$T_{Fisher} = \frac{n_{1.}!n_{2.}!n_{.1}!n_{.2}!}{n_{11}!n_{12}!n_{21}!n_{22}!}$$

$$Trv = -2\sum_{i=1}^{2}\sum_{j=1}^{2} n_{ij}log(\frac{e_{ij}}{n_{ij}} \sim \chi^2_{(1)}$$

$$Tney = \sum_{i=1}^{2}\sum_{j=1}^{2}\frac{(n_{ij}-e_{ij})^2}{n_{ij}} \sim \chi^2_{(1)}$$

Prezi


Teste de Associação
Teste de Associação - χ^2
Teste de Associação - χ^2
Teste de Associação
Do ponto de vista estatístico...
Do ponto de vista estatístico...
Para procurar na tabela χ^2
Estatística do Teste - $\chi^2_{g.l.}$
Tabela Auxiliar
Valores esperados (E)
Tabela 2x2
Regra para cálculo do qui-quadrado
Teste de Associação - χ^2
Tabela auxiliar - χ^2
Teste qui-quadrado
Resultado do Teste qui-quadrado
Tabelas de contingência:
Exemplo:
Teste de associação
Teste de Associação
Exemplo:
Bibliografia

Prezi

Universidade Federal de Mato Grosso
Universidade Federal de Mato Grosso (UFMT)
Instituto de Ciências Exatas e da Terra
Departamento de Estatística

TESTE DE ASSOCIAÇÃO

$$\chi^2$$

Prof. Neuber J. Segri

Teste de Associação - χ^2

- É o teste estatístico mais antigo e um dos mais utilizados na Bioestatística !

- É um método que permite testar a significância da associação entre duas variáveis QUALITATIVAS...

Teste de Associação - χ^2

- Uma forma de resumir e apresentar variáveis **qualitativas** (dados categóricos) é utilizando uma tabela de contingência (tabela cruzada):
  - 2 x 2 – (2 linhas e 2 colunas)
  - i x j – (**i** linhas e **j** colunas)

χ^2

Teste de Associação

Variáveis Dicotômicas

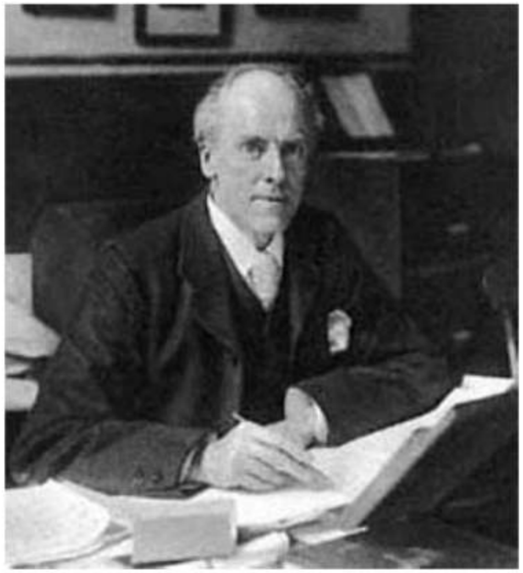
tabela 2x2

Variável 1	Variável 2		TOTAL
	sim	não	
sim	a	b	a+b
não	c	d	c+d
TOTAL	a+c	b+d	N=a+b+c+d

- Teste de associação pelo **Qui-Quadrado**:

  - Karl **Pearson**
    - Estatístico Inglês
    - 1857 – 1936

    Fundador do Depto. Estatística da Univ. College London



Trouxe inúmeras contribuições na Estatística

Fonte: *http://pt.wikipedia.org/wiki/Karl_Pearson*

Do ponto de vista estatístico...

- Teste de associação pelo Qui-Quadrado:
  (método mais comum para analisar tabelas de contingência)
  - 1) Definir as hipóteses:

$$H_0 : Não\ existe\ associação\ entre\ as\ variáveis$$

$$H_A : Existe\ associação\ entre\ as\ variáveis$$

$p_{21)1})^{n_{2.}-n_{21}}$

$$\begin{cases} H_0 : p_{(1)1} = p_{(2)1} = p_{.1}; \\ H_a : p_{(1)1} \neq p_{(2)1}. \end{cases}$$

$$Tq = \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{(n_{ij} - e_{ij})^2}{e_{ij}} \sim \chi^2_{(1)}$$

$$E[N_{i1}] = n_{i.}p_{.1}$$

$$E[N_{i2}] = n_{i.}p_{.2}$$

$$e_{ij} = \frac{n_{i.}n_{.j}}{n}$$

Prezi

Do ponto de vista estatístico...

- Teste de associação pelo Qui-Quadrado:

  (método mais comum para analisar tabelas de contingência)

  - 2) Encontrar os seguintes valores:

    - Qui-quadrado **Crítico** (baseado no nível de significância → tabela do qui-quadrado)

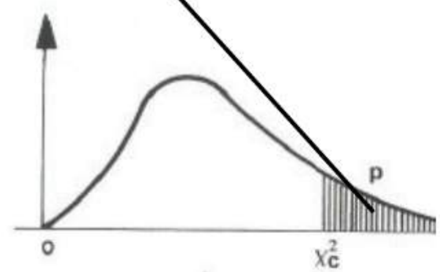
    - Qui-quadrado **Observado** → Estatística do Teste (construção da tabela auxiliar)

Região de Rejeição de Ho

DISTRIBUIÇÃO DE QUIQUADRADO: $\chi^2(n)$

VALORES CRÍTICOS DE QUIQUADRADO TAIS QUE

$$P(\chi^2 > \chi^2_c) = p$$



95%	90%	80%	70%	50%	30%	20%	10%	5%	4%	2,5%	2%	1%
,004	0,016	0,064	0,148	0,455	1,074	1,642	2,706	3,841	4,218	5,024	5,412	6,635
,103	0,211	0,446	0,713	1,386	2,408	3,219	4,605	5,991	6,438	7,378	7,824	9,210

χ^2 **Crítico → Obtido da Tabela**

Para procurar na tabela χ^2:

- Nível de significância: α

- Número de graus de liberdade (g.l.)

(Nº Linhas -1) . (Nº Colunas -1)

Fonte: *http://pt.wikipedia.org/wiki/Karl_Pearson*

Estatística do Teste - $\chi^2_{g.l.}$

$$\boxed{\chi^2_{g.l.} = \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\left(O_{ij}-E_{ij}\right)^2}{E_{ij}}}$$

Para facilitar os cálculos, construir a tabela auxiliar...

$$\chi^2_{g.l.} = \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{(O_{ij}-E_{ij})^2}{E_{ij}}$$

Tabela Auxiliar

- Utilizada para o cálculo da Estatística do Teste

Valores observados (O)	Valores esperados (E)	(O-E)	$(O\text{-}E)^2$	$\frac{(O-E)^2}{E}$
				$\sum$
			Qui-quadrado=	

Compara os valores observados e esperados e verifica se são semelhantes ou não...

Valores esperados (**E**)

- Na tabela: $a_{esperado}$

Variável 1	Variável 2		TOTAL
	sim	não	
sim	a	b	a+b
não	c	d	c+d
TOTAL	a+c	b+d	N=a+b+c+d

Tabela 2x2

- Similarmente:

$$c_{esperado} = \frac{(a+c).(c+d)}{N}$$

Variável 1	Variável 2		TOTAL
	sim	não	
sim	a	b	a+b
não	c	d	c+d
TOTAL	a+c	b+d	N=a+b+c+d

Regra para cálculo do qui-quadrado

- Valores esperados serão:

$$\square_{esperado} = \frac{produtos\ das\ marginais}{N}$$

- O teste qui-quadrado compara as frequências observadas em cada categoria da tabela de contingência com as frequencias esperadas.

Teste de Associação - χ^2

Com o objetivo de investigar a associação entre história de bronquite na infância e presença de tosse diurna ou noturna em idades mais velhas, foram estudados 1.319 adolescentes com 14 anos. Destes, 273 apresentaram história de bronquite até os 5 anos de idade sendo que 26 apresentaram tosse diurna ou noturna aos 14 anos.

Número de adolescentes segundo história de bronquite aos 5 anos e tosse diurna ou noturna aos 14 anos de idade. Local X, ano Y.

	Bronquite		
Tosse	Sim	Não	Total
Sim	26	44	70
Não	247	1002	1249
Total	273	1046	1319

Holland, WW et al.. Long-term consequences of respiratory disease in infancy. *Journal of Epidemiology and Community Health* 1978; 32: 256-9.

Tabelas de contingência:

Teste de associação pelo Qui-Quadrado:

1) Definir as hipóteses:

$$H_0: \textit{Não existe associação entre as variáveis}$$

$$H_A: \textit{Existe associação entre as variáveis}$$

2) Obter os valores crítico e observado:

$$\chi^2_{crítico} \rightarrow tabela\ \chi^2 - depende\ de\ \alpha\ e\ g.l.$$

$$\chi^2_{observado} \rightarrow Estatística\ do\ teste = tabela\ auxiliar$$

Tabela auxiliar - χ^2

Valores observados (O)	Valores esperados (E)	(O-E)	$(O-E)^2$	$\frac{(O-E)^2}{E}$
26	14,488	11,512	132,526	9,147
247	258,512	-11,512	132,526	0,513
44	55,512	-11,512	132,526	2,387
1002	990,488	11,512	132,526	0,134
			Qui-quadrado=	12,181

soma

	Bronquite		
Tosse	Sim	Não	Total
Sim	26	44	70
Não	247	1002	1249
Total	273	1046	1319

Teste qui-quadrado

- Qui-quadrado observado: **12,181**

- Qui-quadrado crítico (1 grau de liberdade):

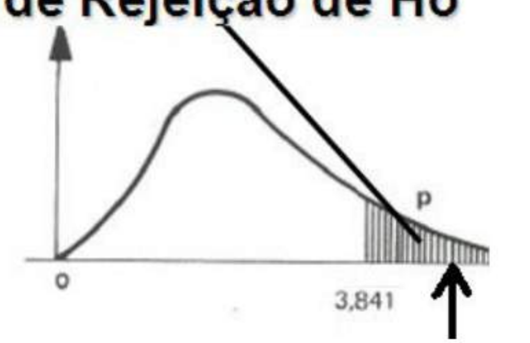
  - (1) $\alpha = 5\%$ 3,841

Resultado do Teste qui-quadrado

$\alpha = 5\%$

- Rejeita-se a hipótese Ho

Região de Rejeição de Ho



Existe associação estatisticamente significativa

Tabelas de contingência:

- Outras dimensões: i x j

Teste de associação pelo χ^2

variável 1	variável 2			TOTAL
	1	...	j	
1 ... i				n_1 ... n_i
TOTAL	m_1	....	m_j	N

Exemplo:

- Verifique a associação, (α=5%):
  - Presença de anemia, segundo idade:

idade (meses)	anêmico	não anêmico	total
0-6	166	147	313
6-12	172	98	270
total	338	245	583

Fonte: Uchimura T "Anemia e peso ao nascer". *Rev. Saúde Pública* 37(4): 2003.

Teste de associação

- Hipóteses: $\begin{cases} H_0 : \text{Não existe associação (entre as variáveis)} \\ H_A : \text{Existe associação (entre as variáveis)} \end{cases}$

- Estatísticas:

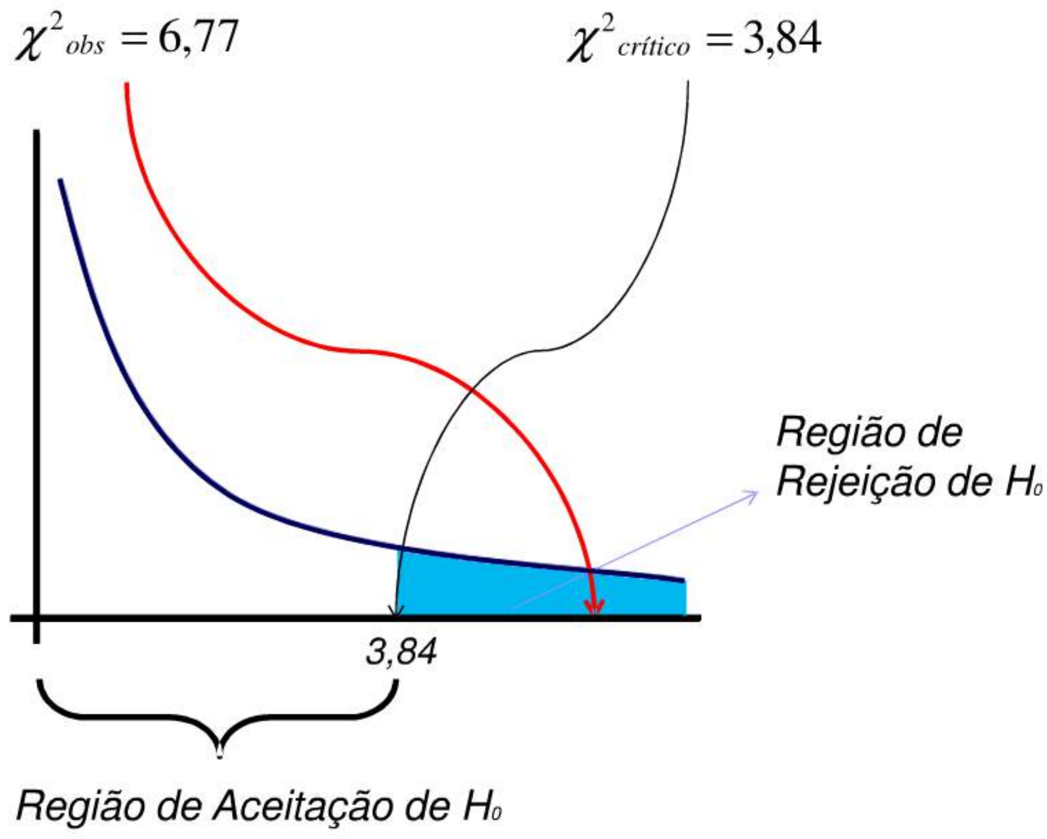
$$\chi^2_{crítico} \rightarrow \alpha = 5\% \; ; 1 g.l. \rightarrow tabela = \boxed{\chi^2_{crítico} = 3{,}84}$$

$$\chi^2_{observado} \rightarrow Estatística \; do \; teste = \boxed{\chi^2_{obs} = 6{,}77}$$

X

C
O
N
F
R
O
N
T
A
R

Teste de Associação



Exemplo:

- Verifique a associação, (α=5%):
  - Presença de anemia, segundo idade:

idade (meses)	anêmico	não anêmico	total
0-6	166	147	313
6-12	172	98	270
total	338	245	583

Fonte: Uchimura T "Anemia e peso ao nascer". *Rev. Saúde Pública* 37(4): 2003.

DECISÃO: Rejeita-se Ho.

Existe associação significativa

Bibliografia

- Arango HG. **Bioestatística Teórica e Computacional**. Guanabara Koogan. 2ª ed. Rio de Janeiro, 2005.
- Bergamaschi DP. Bioestatística (Apostila Graduação) FSP/USP, 2010.
- Callegari-Jacques SM. Bioestatística – Princípios e Aplicações. Artmed. Porto Alegre, 2003.
- ENCE – Escola Nacional de Ciências Estatísticas www.ence.ibge.gov.br
- Latorre MRDO. Bioestatística (Apostila graduação) FSP/USP, 2009.
- Magalhães MN; Lima ACP. **Noções de Probabilidade e Estatística**. EDUSP. São Paulo, 2002.
- Vieira S. **Introdução à Bioestatística.** ELSEVIER. 4ª ed. Rio de Janeiro, 2010.

Prezi

Estatísticas de interesse

Variações da Hipóteses

- $H_0: p_{1(1)} = p_{2(1)} = p_1$
- $H_0: p_{1(1)} - p_{2(1)} = 0$
  (diferença entre incidências) $\longrightarrow$ risco atribuível
- $$H_0: \frac{p_{1(1)}}{p_{2(1)}}$$
  (razão de incidências) $\longrightarrow$ risco relativo
- $$H_0: \frac{p_{1(1)}/(1-p_{1(1)})}{p_{2(1)}/(1-p_{2(1)})} = \frac{p_{11}p_{22}}{p_{12}p_{21}} = 1$$
  (razão dos produtos cruzados $\longrightarrow$ odds ratio ou razão de chances)

- Se $RR = 1$ a probabilidade de resposta positiva não difere entre os indivíduos expostos e não expostos
- Se $RR > 1$ a probabilidade de resposta positiva é maior entre os indivíduos expostos
- Se $RR < 1$ a probabilidade de resposta positiva é maior entre os indivíduos não-expostos
- Se $OR = 1$ a chance de resposta positiva não difere entre os indivíduos expostos e não expostos
- Se $OR > 1$ a chance de resposta positiva é maior entre os indivíduos expostos
- Se $OR < 1$ a chance de resposta positiva é maior entre os indivíduos não-expostos

[illegible]

Prezi

Variações da Hipóteses

- H_0: $p_{(1)1} = p_{(2)1} = p_{.1}$
- H_0: $p_{(1)1} - p_{(2)1} = 0$
  (diferença entre incidências) $\Longrightarrow$ risco atribuível
- $$H_0 : \frac{p_{(1)1}}{p_{(2)1}}$$
  (razão de incidências) $\Longrightarrow$ risco relativo
- $$H_0 : \frac{p_{(1)1}/(1 - p_{(1)1})}{p_{(2)1}/(1 - p_{(2)1})} = \frac{p_{(1)1}p_{(2)2}}{p_{(1)2}p_{(2)1}} = 1$$
  (razão dos produtos cruzados $\Longrightarrow$ odds ratio ou razão de chances)

- Se $RR = 1$ a probabilidade de resposta positiva não difere entre os indivíduos expostos e não-expostos
- Se $RR > 1$ a probabilidade de resposta positiva é maior entre os indivíduos expostos
- Se RR¡1 a probabilidade de resposta positiva é maior entre os indivíduos não-expostos
- Se $OR = 1$ a chance de resposta positiva não difere entre os indivíduos expostos e não-expostos
- Se $OR > 1$ a chance de resposta positiva é maior entre os indivíduos expostos
- Se OR¡1 a chance de resposta positiva é maior entre os indivíduos não-expostos

$$IC(RR) = exp(log(\widehat{RR}) \pm z_{\alpha/2}\sqrt{Var(\widehat{log(RR)})})$$

$$Var(\widehat{log(RR)}) = \frac{1 - p_{(1)1}}{n_1 p_{(1)1}} + \frac{1 - p_{(2)1}}{n_2 p_{(2)1}}$$

$$IC(OR) = exp(log(\widehat{OR}) \pm z_{\alpha/2}\sqrt{Var(\widehat{log(OR)})})$$

$$Var(\widehat{log(OR)}) = \frac{1}{n11} + \frac{1}{n12} + \frac{1}{n21} + \frac{1}{n22}$$

Variaçoes da Hipoteses

- H_0: $p_{(1)1} = p_{(2)1} = p_{.1}$
- H_0: $p_{(1)1} - p_{(2)1} = 0$
  (diferença entre incidências) $\Longrightarrow$ risco atribuível
- $$H_0 : \frac{p_{(1)1}}{p_{(2)1}}$$
  (razão de incidências) $\Longrightarrow$ risco relativo
- $$H_0 : \frac{p_{(1)1}/(1 - p_{(1)1})}{p_{(2)1}/(1 - p_{(2)1})} = \frac{p_{(1)1}p_{(2)2}}{p_{(1)2}p_{(2)1}} = 1$$
  (razão dos produtos cruzados $\Longrightarrow$ odds ratio ou razão de chances)

- Se $RR = 1$ a probabilidade de resposta positiva não difere entre os indivíduos expostos e não-expostos
- Se $RR > 1$ a probabilidade de resposta positiva é maior entre os indivíduos expostos
- Se RR¡1 a probabilidade de resposta positiva é maior entre os indivíduos não-expostos
- Se $OR = 1$ a chance de resposta positiva não difere entre os indivíduos expostos e não-expostos
- Se $OR > 1$ a chance de resposta positiva é maior entre os indivíduos expostos
- Se OR¡1 a chance de resposta positiva é maior entre os indivíduos não-expostos

$IC(RR) = exp$

$Var(\widehat{log(R}$

$IC(OR) = exp($

$Var(\widehat{log(OR}$

re os

os in-

s in-

duos

duos

$$IC(RR) = exp(log(\widehat{RR}) \pm z_{\alpha/2}\sqrt{Var(\widehat{log(RR)}))}$$

$$Var(\widehat{log(RR)}) = \frac{1-p_{(1)1}}{n_{1.}p_{(1)1}} + \frac{1-p_{(2)1}}{n_{2.}p_{(2)1}}$$

$$IC(OR) = exp(log(\widehat{OR}) \pm z_{\alpha/2}\sqrt{Var(\widehat{log(OR)}))}$$

$$Var(\widehat{log(OR)}) = \frac{1}{n11} + \frac{1}{n12} + \frac{1}{n21} + \frac{1}{n22}$$

Outros tipos de estudos

Ensaio Clínico

	Categoria da Variável Y		
Categoria da Variável X	D (1)	ND (2)	Total
TratA (1)	$p_{(1)1}$	$p_{(1)2}$	1
TratB(2)	$p_{(2)1}$	$p_{(2)2}$	1
Total	$p_{.1}$	$p_{.2}$	1

$N_{1j} \sim Bin(n_{.1}, p_{1(j)})$

$N_{2j} \sim Bin(n_{.2}, p_{2(j)})$

$$P(N_{1j} = n_{1j}, N_{2j} = n_{2j}) = \prod_{j=1}^{2}\left[n_{.j}!\prod_{i=1}^{2}\frac{p_{i(j)}^{n_{ij}}}{n_{ij}!}\right]$$

$$\begin{cases} H_0 : p_{1(1)} = p_{1(2)} = p_{1.} \\ H_a : p_{1(1)} \neq p_{1(2)}. \end{cases}$$

$$T_q = \sum_{i=1}^{2}\sum_{j=1}^{2}\frac{(n_{ij} - e_{ij})^2}{e_{ij}} \sim \chi^2_{(1)}$$

$$e_{ij} = \frac{n_{i.}n_{.j}}{n}$$

$$\hat{p}_{i(j)} = \frac{n_{ij}}{n_{.j}}$$

$$E[N_{1j}] = n_{.j}p_{1.} e E[N_{2j}] = n_{.j}p_{2.}$$

Estudo Caso Controle

	Categoria da Variável Y		
Categoria da Variável X	Caso (1)	Controle (2)	Total
E (1)	$p_{1(1)}$	$p_{1(2)}$	$p_{1.}$
NE (2)	$p_{2(1)}$	$p_{2(2)}$	$p_{2.}$
Total	1	1	1

→ D = doença, $\overline{D}$ = não doença, $P(\overline{D}) = 1 - P(D)$

⇒ E – exposição e $\overline{E}$ – não-exposição

$$\begin{aligned} RR &= \frac{p_{(1)1}}{p_{(2)1}} = \frac{P(D|E)}{P(D|\overline{E})} = \frac{P(D)P(E|D)/[P(D)P(E|D) + P(\overline{D})P(E|\overline{D})]}{P(D)P(\overline{E}|D)/[P(D)P(\overline{E}|D) + P(\overline{D})P(\overline{E}|\overline{D})]} \\ &= \frac{P(E|D)\{P(\overline{E}|\overline{D}) + P(D)[P(\overline{E}|D) - P(\overline{E}|\overline{D})]\}}{P(\overline{E}|D)\{P(E|\overline{D}) + P(D)[P(E|D) - P(E|\overline{D})]\}} \\ & \underbrace{\text{Se doença rara, P(D)} \to 0} \\ RR &\approx \frac{P(E|D)P(\overline{E}|\overline{D})}{P(\overline{E}|D)P(E|\overline{D})} = \frac{p_{1(1)}p_{2(2)}}{p_{2(1)}p_{1(2)}} = OR \end{aligned}$$

Estudo Transversal

	Categoria da Variável Y		
Categoria da Variável X	D (1)	ND (2)	Total
E (1)	p_{11}	p_{12}	
NE (2)	p_{21}	p_{22}	
Total			1

$$P(N = n) = P(N_{11} = n_{11}, N_{12} = n_{12}, N_{21} = n_{21}, N_{22} = n_{22}) = n!\prod_{i=1}^{2}\prod_{j=1}^{2}\frac{p_{(ij)}^{n_{ij}}}{n_{ij}!}$$

$$\begin{cases} H_0 : p_{ij} = p_{i.}p_{.j} \\ H_a : p_{ij} \neq p_{i.}p_{.j} \end{cases}$$

$$\widehat{OR} = \frac{n_{11}n_{22}}{n_{12}n_{21}}$$

$$\widehat{RP} = \frac{n_{11}/n_1}{n_{21}/n_2}$$

$$\hat{p}_{(ij)} = \frac{n_{ij}}{n}$$

$$e_{ij} = n\left(\frac{n_{i.}}{n}\right)\left(\frac{n_{.j}}{n}\right)$$

Outros tipos de estud

	Categoria da Variável Y		
Categoria da Variável X	D (1)	ND (2)	Total
TratA (1)	$p_{(1)1}$	$p_{(1)2}$	1
TratB(2)	$p_{(2)1}$	$p_{(2)2}$	1
Total	$p_{.1}$	$p_{.2}$	1

Estudo Caso Con

	Categ

Prezi

$p_{.1}$	$p_{.2}$	1

udo Caso Controle

	Categoria da Variável Y		
Categoria da Variável X	Caso (1)	Controle (2)	Total
E (1)	$p_{1(1)}$	$p_{1(2)}$	$p_{1.}$
NE (2)	$p_{2(1)}$	$p_{2(2)}$	$p_{2.}$
Total	1	1	1

$\Rightarrow D =$ doença, $\overline{D} =$ não doença, $P(\overline{D}) = 1 - P(D)$

$\Rightarrow E =$ exposição e $\overline{E} =$ não-exposição

$$RR = \frac{p_{(1)1}}{p_{(2)1}} = \frac{P(D|E)}{P(D|\overline{E})} = \frac{P(D)P(E|D)/[P(D)P(E|D)+P(\overline{D})P(E|\overline{D})]}{P(D)P(\overline{E}|D)/[P(D)P(\overline{E}|D)+P(\overline{D})P(\overline{E}|\overline{D})]}$$

$$= \frac{P(E|D)\{P(\overline{E}|\overline{D})+P(D)[P(\overline{E}|D)-P(\overline{E}|\overline{D})]\}}{P(\overline{E}|D)\{P(E|\overline{D})+P(D)[P(E|D)-P(E|\overline{D})]\}}$$

Prezi

Estu

$$N_{11} \sim Bin(n_{.1}, p_{1(1)})$$

$$N_{12} \sim Bin(n_{.2}, p_{1(2)})$$

$$P(N_{1j} = n_{1j}, N_{2j} = n_{2j}) = \prod_{j=1}^{2} \left[n_{.j}! \prod_{i=1}^{2} \frac{p_{i(j)}^{n_{ij}}}{n_{ij}!} \right]$$

$$\begin{cases} H_0 : p_{1(1)} = p_{1(2)} = p_{1.} \\ H_a : p_{1(1)} \neq p_{1(2)}. \end{cases}$$

$$Tq = \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{(n_{ij} - e_{ij})^2}{e_{ij}} \sim \chi^2_{(1)}$$

$$e_{ij} = \frac{n_{i.} n_{.j}}{n}$$

$$\hat{p}_{i(j)} = \frac{n_{ij}}{n_{.j}}$$

$$E[N_{1j}] = n_{.j} p_{1.} e E[N_{2j}] = n_{.j} p_{2.}$$

Estudo Transversal

Tabela 6.5: Tabela 2x2

	Categoria da Variável Y

Total	1	1	1

$\Rightarrow D =$ doença, $\overline{D} =$ não doença, $P(\overline{D}) = 1 - P(D)$
$\Rightarrow E =$ exposição e $\overline{E} =$ não-exposição

$$
\begin{aligned}
RR &= \frac{p_{(1)1}}{p_{(2)1}} = \frac{P(D|E)}{P(D|\overline{E})} = \frac{P(D)P(E|D)/[P(D)P(E|D)+P(\overline{D})P(E|\overline{D})]}{P(D)P(\overline{E}|D)/[P(D)P(\overline{E}|D)+P(\overline{D})P(\overline{E}|\overline{D})]} \\
&= \frac{P(E|D)\{P(\overline{E}|\overline{D})+P(D)[P(\overline{E}|D)-P(\overline{E}|\overline{D})]\}}{P(\overline{E}|D)\{P(E|\overline{D})+P(D)[P(E|D)-P(E|\overline{D})]\}}
\end{aligned}
$$

$$
\overbrace{\text{Se doença rara, P(D)} \to 0}
$$

$$
RR \approx \frac{P(E|D)P(\overline{E}|\overline{D})}{P(\overline{E}|D)P(E|\overline{D})} = \frac{p_{1(1)}p_{2(2)}}{p_{2(1)}p_{1(2)}} = OR
$$

$$p_{i(j)} = \frac{n_{ij}}{n_{.j}}$$

$$E[N_{1j}] = n_{.j}p_{1.} e E[N_{2j}] = n_{.j}p_{2.}$$

Estudo Transversal

	Categoria da Variável Y		
Categoria da Variável X	D (1)	ND (2)	Total
E (1)	p_{11}	p_{12}	
NE (2)	p_{21}	p_{22}	
Total			1

$$P(N = n) = P(N_{11} = n_{11}, N_{12} = n_{12}, N_{21} = n_{21}, N_{22} = n_{22}) = n! \prod_{i=1}^{2} \prod_{j=1}^{2} \frac{p_{i(j)}^{n_{ij}}}{n_{ij}!}$$

$$\begin{cases} H_0 : p_{ij} = p_{i.}p_{.j} \\ H_a : p_{ij} \neq p_{i.}p_{.j} \end{cases}$$

$$\widehat{OR} = \frac{n_{11}n_{22}}{n_{12}n_{21}}$$

$$\widehat{RP} = \frac{n_{11}/n_{1.}}{n_{21}/n_{2.}}$$

$$\hat{p}_{i(j)} = \frac{n_{ij}}{n}$$

$$e_{ij} = n \left(\frac{n_{i.}}{n}\right) \left(\frac{n_{.j}}{n}\right)$$

Prezi

Utilizando Software R

Estudo de Coorte

```
dat <- as.table(matrix(c(180,120,125,175), nrow = 2, byrow = TRUE))
epi.2by2(dat,method="cohort.count",conf.level=0.95)
epi.2by2(dat,method="cohort.count",conf.level=0.95,verbose=T)
chisq.test(dat,correct=F)
fisher.test(dat)
dat <- as.table(matrix(c(50,450,5,495), nrow = 2, byrow = TRUE))
```

pg215 Fundamentos da epidemiologia

```
             Disease +    Disease -       Total          Inc risk *
Exposed +          180          120         300                60.0
Exposed -          125          175         300                41.7
Total              305          295         600                50.8
              Odds
Exposed +    1.500
Exposed -    0.714
Total        1.034

Point estimates and 95 % CIs:
-------------------------------------------------------------------
Inc risk ratio                             1.44 (1.22, 1.69)
Odds ratio                                 2.1 (1.5, 2.95)
Attrib risk *                              18.33 (10.47, 26.2)
Attrib risk in population *                9.17 (2.3, 16.03)
Attrib fraction in exposed (%)             30.56 (18.29, 40.98)
Attrib fraction in population (%)          18.03 (9.68, 25.62)
-------------------------------------------------------------------
 * Cases per 100 population units
```

Estudo de Caso Controle

```
dat <- as.table(matrix(c(100,30,400,470), nrow = 2, byrow = TRUE))
epi.2by2(dat,method="case.control",conf.level=0.95)
```

pg235 Fundamentos da epidemiologia

```
             Disease +    Disease -       Total        Prevalence *       Odds
Exposed +          100           30         130                76.9      3.333
Exposed -          400          470         870                46.0      0.851
Total              500          500        1000                50.0      1.000

Point estimates and 95 % CIs:
-------------------------------------------------------------------
Odds ratio                                 3.91 (2.52, 6.23)
Attrib prevalence *                        30.95 (22.98, 38.91)
Attrib prevalence in population *          4.02 (-0.51, 8.56)
Attrib fraction (est) in exposed  (%)      74.43 (60.26, 83.95)
Attrib fraction (est) in population (%)    14.82 (10.61, 18.97)
-------------------------------------------------------------------
 * Cases per 100 population units
```

Estudo Transversal

```
dat <- as.table(matrix(c(100,30,400,470), nrow = 2, byrow = TRUE))
> epi.2by2(dat,method="cross.sectional",conf.level=0.95)
```

```
             Disease +    Disease -       Total        Prevalence *       Odds
Exposed +          100           30         130                76.9      3.333
Exposed -          400          470         870                46.0      0.851
Total              500          500        1000                50.0      1.000

Point estimates and 95 % CIs:
-------------------------------------------------------------------
Prevalence ratio                           1.67 (1.49, 1.88)
Odds ratio                                 3.91 (2.52, 6.23)
Attrib prevalence *                        30.95 (22.98, 38.91)
Attrib prevalence in population *          4.02 (-0.51, 8.56)
Attrib fraction in exposed (%)             40.23 (32.71, 46.91)
Attrib fraction in population (%)          8.05 (5.53, 10.49)
-------------------------------------------------------------------
 * Cases per 100 population units
```

Estudo de Coorte

```
dat <- as.table(matrix(c(180,120,125,175), nrow = 2, byrow = TRUE))
epi.2by2(dat,method="cohort.count",conf.level=0.95)
epi.2by2(dat,method="cohort.count",conf.level=0.95,verbose=T)
chisq.test(dat,correct=F)
fisher.test(dat)
dat <- as.table(matrix(c(50,450,5,495), nrow = 2, byrow = TRUE))
```

pg215 Fundamentos da epidemiologia

```
             Disease +    Disease -      Total        Inc risk *
Exposed +          180          120        300              60.0
Exposed -          125          175        300              41.7
Total              305          295        600              50.8
                  Odds
Exposed +        1.500
Exposed -        0.714
Total            1.034

Point estimates and 95 % CIs:
---------------------------------------------------------
Inc risk ratio                          1.44 (1.22, 1.69)
Odds ratio                              2.1 (1.5, 2.95)
Attrib risk *                           18.33 (10.47, 26.2)
Attrib risk in population *             9.17 (2.3, 16.03)
Attrib fraction in exposed (%)          30.56 (18.29, 40.98)
Attrib fraction in population (%)       18.03 (9.68, 25.62)
---------------------------------------------------------
 * Cases per 100 population units
```

Estudo de Caso Controle

dat <- as.table(matrix(c(100,30,400,470), nrow = 2, byrow = TRUE))
epi.2by2(dat,method="case.control",conf.level=0.95)
pg235 Fundamentos da epidemiologia

```
             Disease +    Disease -      Total        Prevalence *        Odds
Exposed +          100           30        130                76.9       3.333
Exposed -          400          470        870                46.0       0.851
Total              500          500       1000                50.0       1.000

Point estimates and 95 % CIs:
---------------------------------------------------------
Odds ratio                                  3.91 (2.52, 6.23)
Attrib prevalence *                         30.95 (22.98, 38.91)
Attrib prevalence in population *           4.02 (-0.51, 8.56)
Attrib fraction (est) in exposed  (%)       74.43 (60.26, 83.95)
Attrib fraction (est) in population (%)     14.89 (10.61, 18.97)
---------------------------------------------------------
 * Cases per 100 population units
```

Prezi

```
8, 25.62)
--------
```

Estudo Transversal

dat <- as.table(matrix(c(100,30,400,470), nrow = 2, byrow = TRUE))

> epi.2by2(dat,method="cross.sectional",conf.level=0.95)

```
             Disease +    Disease -      Total        Prevalence *        Odds
Exposed +          100           30        130                76.9       3.333
Exposed -          400          470        870                46.0       0.851
Total              500          500       1000                50.0       1.000

Point estimates and 95 % CIs:
---------------------------------------------------------
Prevalence ratio                             1.67 (1.49, 1.88)
Odds ratio                                   3.91 (2.52, 6.23)
Attrib prevalence *                          30.95 (22.98, 38.91)
Attrib prevalence in population *            4.02 (-0.51, 8.56)
Attrib fraction in exposed (%)               40.23 (32.71, 46.91)
Attrib fraction in population (%)            8.05 (5.53, 10.49)
---------------------------------------------------------
 * Cases per 100 population units
```

Tabelas r x c

- Variáveis Y e X nominais

	Bebê com baixo peso		
Tabagismo mãe	Sim	Não	
sim	50	450	500
Não	5	495	500
Totais	55	945	1000

- Variável Y ordinal e X nominal (ni. fixos)

	Horas de promoção de alivio da dor					Totais
Medicamentos	0	1	2	3	4	
Placebo	6	9	6	3	1	25
Padrão	1	4	6	6	8	25
Novo	2	5	6	8	6	27
Totais	9	18	18	17	15	77

- Variáveis ordinais (n fixo)

Produto	Classificação da limpeza			
	Baixa	Média	Alta	
Agua	27	14	5	46
Agua +prod1	10	17	26	53
Agua +prod2	5	12	50	67
	42	43	81	166

- Variáveis ordinais (ni. fixos)

- Variáveis estratificadoras (confundimento e modificadora de efeito)

- Variáveis Y e X nominais

Estabelecidas as hipóteses adequadas
homogeneidade para ni. fixos
independencia para n fixo
multiplicatividade para totais aleatórios

$$Tq = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(n_{ij} - e_{ij})^2}{e_{ij}} \sim \chi^2_{(r-1)(c-1)}$$

$$e_{ij} = n \left(\frac{n_{i.}}{n}\right) \left(\frac{n_{.j}}{n}\right)$$

Pode-se usar TRV e TNeyman
E alternativamente Exato de Fisher

Tabelas r x c

- **Variáveis Y e X nominais**

	Bebê com baixo peso		
Tabagismo mãe	Sim	Não	
sim	50	450	500
Não	5	495	500
Totais	55	945	1000

- **Variável Y ordinal e X nominal (ni. fixos)**

	Horas de promoção de alivio da dor					Totais
Medicamentos	0	1	2	3	4	
Placebo	6	9	6	3	1	25
Padrão	1	4	6	6	8	25
Novo	2	5	6	8	6	27
Totais	9	18	18	17	15	77

- **Variáveis ordinais (n fixo)**

Produto	Classificação da limpeza			
	Baixa	Média	Alta	
Agua	27	14	5	46
Agua +prod1	10	17	26	53
Agua +prod2	5	12	50	67
	42	43	81	166

- **Variáveis ordinais (ni. fixos)**

- **Variáveis estratificadoras (confundimento e modificadora de efeito)**