Tipos de variáveis

Variáveis Y e X nominais

Variável Y ordinal e X nominais (ni. fixos)

Variáveis ordinais (n fixo)

Variáveis ordinais (ni. fixos)

Variáveis estratificadoras (Confundimento e modificacora de efeito)

Tabelas de contingência

Extensões

Tabelas de contingência

Extensões

Tipos de variáveis

Variáveis Y e X nominais

Variável Y ordinal e X nominais (ni. fixos)

Variáveis ordinais (n fixo)

Variáveis ordinais (ni. fixos)

Variáveis estratificadoras (Confundimento e modificacora de efeito)

Tipos de variáveis

Variáveis Y e X nominais

Variável Y ordinal e X nominais (ni. fixos)

Variáveis ordinais (n fixo)

Variáveis ordinais (ni. fixos)

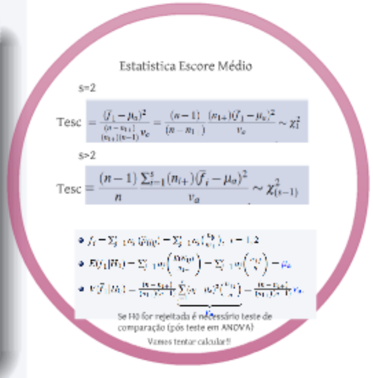
Variáveis estratificadoras (Confundimento e modificacora de efeito)

Variável Y ordinal e X nominal com totais ni. fixos

Avaliação de tratamentos em pacientes com artrite reumatóide.

Tratamentos	Melhora do Paciente: Nenhuma	Alguma	Acentuada	Totais
Ativo	13	7	21	**41**
Placebo	29	7	7	**43**
Totais	42	14	28	84

Fonte: Stokes et al. (2000)



O objetivo é testar H0: p1 = p2 = ... = pr (hipótese de homogeneidade)
Modelo: Produto de multinomiais
Observe que desta forma não utilizamos a natureza ordinal da variável

Solução: atribuir escores (a1,a2,...,ar) para as categorias de Y.
Novo teste H0: F1=F2 são duas comparações se ativo igual passivo.

$$\overline{F}_i = \sum_{j=1}^{r} a_j \left(p_{(i)j}\right) \qquad i = 1, \ldots, s.$$

- $\overline{f}_i = \sum_{j=1}^{r} a_j \left(\widehat{p}_{(i)j}\right) = \sum_{j=1}^{r} a_j \left(\frac{n_{ij}}{n_{i+}}\right), \quad i = 1,2$
- $E(\overline{f}_1 | H_0) = \sum_{j=1}^{r} a_j \left(\frac{E(N_{1j})}{n_{1+}}\right) = \sum_{j=1}^{r} a_j \left(\frac{n_{+j}}{n}\right) = \mu_{\mathbf{a}}$
- $V(\overline{f}_1 | H_0) = \frac{(n-n_{1+})}{(n_{1+})(n-1)} \underbrace{\sum_{j=1}^{r} (a_j - \mu_a)^2 \left(\frac{n_{+j}}{n}\right)}_{\nu_{\mathbf{a}}} = \frac{(n-n_{1+})}{(n_{1+})(n-1)} \nu_{\mathbf{a}}.$

Estatistica Escore Médio

s=2

$$\text{Tesc} = \frac{(\bar{f}_1 - \mu_a)^2}{\frac{(n-n_{1+})}{(n_{1+})(n-1)} v_a} = \frac{(n-1)}{(n-n_{1+})} \frac{(n_{1+})(\bar{f}_1 - \mu_a)^2}{v_a} \sim \chi_1^2$$

s>2

$$\text{Tesc} = \frac{(n-1)}{n} \frac{\sum_{i=1}^{s} (n_{i+})(\bar{f}_i - \mu_a)^2}{v_a} \sim \chi^2_{(s-1)}$$

- $\bar{f}_i = \sum_{j=1}^{r} a_j \left(\widehat{p}_{(i)j}\right) = \sum_{j=1}^{r} a_j \left(\frac{n_{ij}}{n_{i+}}\right), \quad i = 1, 2$
- $E(\bar{f}_1 | H_0) = \sum_{j=1}^{r} a_j \left(\frac{E(N_{1j})}{n_{1+}}\right) = \sum_{j=1}^{r} a_j \left(\frac{n_{+j}}{n}\right) = \mu_{\mathbf{a}}$
- $V(\bar{f}_1 | H_0) = \frac{(n-n_{1+})}{(n_{1+})(n-1)} \underbrace{\sum_{j=1}^{r} (a_j - \mu_a)^2 \left(\frac{n_{+j}}{n}\right)}_{v_{\mathbf{a}}} = \frac{(n-n_{1+})}{(n_{1+})(n-1)} v_{\mathbf{a}}.$

Se H0 for rejeitada é necessário teste de comparação (pós teste em ANOVA)

Vamos tentar calcular!!

Variável Y ordinal e X nominal (ni. fixos)

	Horas de alívio					
Medicamentos	0	1	2	3	4	Totais
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

Fonte: Stokes et al. (2000)

O objetivo é testar H0: p1.=p2.=p3. (teste independencia)
Modelo: Produto de multinomiais
Podemos utilizar escore médio e a H0: F1=F2=F3
Se p-valor associado a Estatistica Escore Medio, apresentar significancia, isto é, existe evidencias de que pelo menos dois medicamentos diferem entre si, necessitamos discriminar quais são os medicamentos

Podemos utilizar teste para comparações dois a dois (método de Bonferroni)

Estimativas f1=1,36 f2=2,64 e f3= 2,41
o nivel de significancia agora é 0,017, isto é , 0,05/3

Sendo assim calculamos a estatistica para :
Placebo x padrão => 11,66 [gl=1 e p=0,0006]
Placebo x novo => 8,59 [gl=1 e p=0,0034]
Padrão x novo => 0,46 [gl=1 e p=0,4951]

Solução placebo difere de novo e padrão.
Obs: Existe freq menor que 5 (Tpearson não é adequado)

ar
os
os

Podemos utilizar teste para comparações dois a dois (método de Bonferroni)

Estimativas f1=1,36 f2=2,64 e f3= 2,41
o nivel de significancia agora é 0,017, isto é , 0,05/3

Sendo assim calculamos a estatistica para :
Placebo x padrão => 11,66 [gl=1 e p=0,0006]
Placebo x novo => 8,59 [gl=1 e p=0,0034]
Padrão x novo => 0,46 [gl=1 e p=0,4951]

Solução placebo difere de novo e padrão.
Obs: Existe freq menor que 5 (Tpearson não é adequado)

Variáveis Y e X ordinais (n fixo)

Consciência do risco	Tabaco: Não usa	Tabaco: Usa	Totais
Mínima	70	33	103
Moderada	202	40	242
Substancial	218	11	229
Totais	490	84	**574**

Fonte: Stokes et al. (2000)

Estudo Transversal
Modelo: multinomial
H0: pij = pi. p.j, para i = 1,2,3 e j =1,2
Ha: pelo menos um par (i,j) difere
i.é pij difere pi. e p.j
Este exemplo temo dicotomica e ordinal

Podemos usar escores para as duas variáveis para o tabaco "a " e para a consciencia "c"
Ex: a=(a1,a2)=(0,1) e c=(c1,c2,c3)=(1,2,3)

Escore medio é $\overline{F} = \sum_{i=1}^{3}\sum_{j=1}^{2} c_i a_j p_{ij}$

Estimativa => $f = \sum_{i=1}^{3}\sum_{j=1}^{2} c_i a_j \hat{p}_{ij} = \sum_{i=1}^{3}\sum_{j=1}^{2} \frac{c_i a_j n_{ij}}{n}$

Enfim a Estatistica de teste é dada por:

$$= \frac{(\bar{f} - \mu_c \mu_a)^2}{V(f)} = \ldots = (n-1)(r_{ac})^2 \sim \chi_1^2$$

$$E(f) = \sum_{i=1}^{3}\sum_{j=1}^{2} \frac{c_i a_j}{n} E(N_{ij}) = \sum_{i=1}^{3} c_i \binom{n_{i+}}{n} \sum_{j=1}^{2} a_j \binom{n_{+j}}{n} = \mu_c \mu_a$$

$$V(\bar{f}) = \left\{ \sum_{i=1}^{3} (c_i - \mu_c)^2 \left(\frac{n_{i+}}{n}\right) \sum_{j=1}^{2} \frac{(a_j \quad \mu_a)^2 (n_{+j}/n)}{(n-1)} \right\}.$$

Neste caso rac é o coeficiente de correlação de Pearson

Vamos Calcular?

Escore medio é

$$\overline{F} = \sum_{i=1}^{3}\sum_{j=1}^{2} c_i\, a_j\, p_{ij}$$

Estimativa =>

$$\bar{f} = \sum_{i=1}^{3}\sum_{j=1}^{2} c_i\, a_j\, \widehat{p}_{ij} = \sum_{i=1}^{3}\sum_{j=1}^{2} \frac{c_i\, a_j\, n_{ij}}{n}$$

Enfim a Estatistica de teste é dada por:

$$= \frac{(\bar{f} - \mu_c\, \mu_a)^2}{V(\bar{f})} = \ldots = (n-1)\,(r_{ac})^2 \sim \chi_1^2$$

$$E(\bar{f}) = \sum_{i=1}^{3}\sum_{j=1}^{2} \frac{c_i\, a_j}{n} E(N_{ij}) = \sum_{i=1}^{3} c_i \left(\frac{n_{i+}}{n}\right) \sum_{j=1}^{2} a_j \left(\frac{n_{+j}}{n}\right) = \mu_c\, \mu_a$$

$$V(\bar{f}) = \left\{ \sum_{i=1}^{3} (c_i - \mu_c)^2 \left(\frac{n_{i+}}{n}\right) \sum_{j=1}^{2} \frac{(a_j - \mu_a)^2 (n_{+j}/n)}{(n-1)} \right\}.$$

Neste caso rac é o coeficiente de correlação de Pearson

Vamos Calcular?

E este exemplo??

Tratamentos	Limpeza Baixa	Média	Alta	Totais
Água	27	14	5	46
Água + dose única trat padrão	10	17	26	53
Água + dose dupla trat padrão	5	12	50	67
Totais	42	43	81	166

Fonte: Stokes et al. (2000)

Podemos assumir escores para ambos (1,2,3)
Calcular a estatistica anterior "Escore correlação"

Resposta:
Tec= 50,6 [gl=1, p=0,0001] com rac=0,554(ié aumenta limpeza e aumenta a dosagem)

ou a Estatistica escore Tscore
Tscore= 52,77 [gl=2 p<0,01]
e para esta estatistica as comparações
H0: F1=F2 Tscore=21,71 [p<0,001 gl=1]
H0: F1=F3 Tscore=49,06 [p<0,001 gl=1]
H0: F2=F3 Tscore=8,02 [p<0,001 gl=1]
Portanto existe diferença entre os tipos
f1=1,52 <f2=2,3 <f3 =2,67

Água + dose única trat padrão	[illegible]	[illegible]	[illegible]	[illegible]
Água + dose dupla trat padrão	5	12	50	67
Totais	42	43	81	166

Fonte: Stokes et al. (2000)

Podemos assumir escores para ambos (1,2,3)
Calcular a estatistica anterior "Escore correlação"

Resposta:
Tec= 50,6 [gl=1, p=0,0001] com rac=0,554(ié aumenta limpeza e aumenta a dosagem)

ou a Estatistica escore Tscore
Tscore= 52,77 [gl=2 p<0,01]
e para esta estatistica as comparações
H0: F1=F2 Tscore=21,71 [p<0,001 gl=1]
H0: F1=F3 Tscore=49,06 [p<0,001 gl=1]
H0: F2=F3 Tscore=8,02 [p<0,001 gl=1]
Portanto existe diferença entre os tipos
f1=1,52 <f2=2,3 <f3 =2,67

Resumo

Variáveis ordinais onde pode-se assumir escores

Se total n fixo => estatística da correlação

$$\frac{(n-1)}{(n-n_{1+})}\frac{(n_{1+})(\bar{f}_1-\mu_a)^2}{v_a}\sim\chi^2_1$$

$$=\frac{(n-1)}{n}\frac{\sum_{i=1}^{s}(n_{i+})(\bar{f}_i-\mu_a)^2}{v_a}\sim\chi^2_{(s-1)}$$

Se total ni. fixos => estatistica escore e/ou da correlação

$$=\frac{(\bar{f}-\mu_c\,\mu_a)^2}{V(\bar{f})}=\ldots=(n-1)\,(r_{ac})^2\sim\chi^2_1$$

Análise Estratificada

Justificativa: Existe a presença de variaveis que acabam confundindo ou modificando algum efeito

- Variáveis de confundimento , distorcem a associação entre X e Y -alteram o sentido.
- Variáveis que modificam o efeito, mostram o efeito de X sobre Y variando de acordo com as categorias de outra variável Z, indicando que presença de interação entre X e Z

Ideia é minimizar / controlar o efeito destas variáveis, e a solução é estratificação pelas variáveis.
A origem da estratificação pode ser devido ao delineamento adotado ou por necessidade pós coleta

Exemplos

Fumo Passivo	Fumo Voluntário	Câncer de Pulmão Sim	Câncer de Pulmão Não	Totais
Sim	Sim			
Sim	Não			
Totais				
Não	Sim			
Não	Não			
Totais				

Resposta : Cancer pulmão
explicativa : Fumo voluntario
modificadora de efeito: Fumo passivo

Idade (anos)	Consumo de ferro na dieta	Anemia Sim	Anemia Não	Totais
< 2	Sim			
< 2	Não			
Totais				
Entre 2 e 6	Sim			
Entre 2 e 6	Não			
Totais				

Resposta : Anemia
explicativa: Consumo Fe
modif. efeito : Idade cça

Justificativa: Existe a presença de va
modificando algum efeito

- Variáveis de confundimento , distorcem a associação entre X e Y -alteram o sentido.

- Variáveis que modificam o efeito, mostram o efeito de X sobre Y variando de acordo com as categorias de outra variável Z, indicando que presença de interação entre X e Z

Ex

iaveis que acabam confundindo ou

Ideia é minimizar / controlar o efeito destas variáveis, e a solução é estratificação pelas variáveis.

A origem da estratificação pode ser devido ao delineamento adotado ou por necessidade pós coleta

emplos

Exemplios

Fumo Passivo	Fumo Voluntário	Câncer de Pulmão Sim	Câncer de Pulmão Não	Totais
Sim	Sim			
Sim	Não			
Totais				
Não	Sim			
Não	Não			
Totais				

Resposta : Cancer pulmão
explicativa : Fumo voluntario
modificadora de efeito: Fumo passivo

Idade (anos)	Consumo de ferro na dieta	Anemia Sim	Anemia Não	Totais
< 2	Sim			
< 2	Não			
Totais				
Entre 2 e 6	Sim			
Entre 2 e 6	Não			
Totais				

Resposta : Anemia
explicativa: Consumo Fe
modif. efeito : Idade cça

Como discriminar as variáveis??

Estratificação			Tabelas 2 x 2 x 2
sem	com		
OR	OR_1	OR_2	
3,0	3,0	3,0	$OR \approx OR_1 \approx OR_2 \Rightarrow$ ausência de ambos
3,0	1,9	2,0	$OR \neq OR_1$ e OR_2, mas $OR_1 \approx OR_2 \Rightarrow$ confundimento
3,0	4,2	4,3	$OR \neq OR_1$ e OR_2, mas $OR_1 \approx OR_2 \Rightarrow$ confundimento
3,0	1,0	6,2	$OR \neq OR_1$ e OR_2, mas $OR_1 < OR_2 \Rightarrow$ interação
3,0	3,8	0,5	$OR \neq OR_1$ e OR_2, mas $OR_1 > OR_2 \Rightarrow$ interação

SE MOFICADORA DE EFEITO (INTERAÇÃO SIGNIFICATIVA) APRESENTA RR OU OR ASSOCIADAS A CADA ESTRATO

SE CONFUNDIMENTO APRESENTAR MEDIDA GLOBAL PELO CONFUNDIMENTO (MANTEL -HAENSZEL)

Estatistica de Mantel Haenszel(1959)

Supor q tabelas 2 por 2, se q=2 temos 2 tabelas do tipo 2x2

Tratamentos	Resposta $j=1$	$j=2$	Totais
$i=1$	n_{h11}	n_{h12}	n_{h1+}
$i=2$	n_{h21}	n_{h22}	n_{h2+}
Totais	n_{h+1}	n_{h+2}	n_h

h é o número da referida tabela

Totais marginais na linha nhi. são fixos

Temos interesse em testar H0: ph(1)1=ph(2)1

Condicional a H0, Nh11 tem distribuição Hipergeométrica, de onde a estatistica de MH sob H0 e para somatoria n nas q tabelas é suficientemente grande.

$$Q_{MH} = \frac{\left(\sum_{h=1}^{q} n_{h11} - \sum_{h=1}^{q} e_{h11}\right)^2}{\sum_{h=1}^{q} v_{h11}} \sim \chi^2_{(1)}$$

$$OR_{MHestimado} = \frac{\sum_{h=1}^{q} \frac{n_{h11} n_{h22}}{n_h}}{\sum_{h=1}^{q} \frac{n_{h12} n_{h21}}{n_h}}$$

Exemplo

Centros	Medicamentos	Efeito Favorável	Não favorável	Totais
1	Novo	29	16	**45**
1	Padrão	14	31	**45**
Totais		43	47	90
2	Novo	37	8	**45**
2	Padrão	24	21	**45**
Totais		61	29	90

Resposta : Efeito
Explicativa: Medicamento
Estratificadora: Centros

OR=3,76

OR1=4,013 e OR2=4,04

TMH= 18,41 (p<0,001)

TBreslow-Day p=0,99

OR_{MH} = 4,02

A chance de melhora dos pac que tomaram Novo foi de 4 vezes em relaçao aos que tomaram placebo

- Indica que existe associação entre as variáveis medicamento e resposta do paciente, controlando a variável centro.
- Veja que o medicamento novo apresentou proporçao maior que o padrao em relação ao efeito favoravel, ph(1)1>ph(2)1

Estatística de Mantel-Haenszel para outros casos

Y ordinal e X nominal (ni. fixos)	Estatistica Escore de Mantel-Haenszel
Y e X ordinais (ni. fixos)	Estatistica Escore de Mantel-Haenszel ou Estatistica Escore Correlação de Mantel-Haenszel
Y e X ordinais (n fixos)	Estatistica Escore Correlação de Mantel-Haenszel

Vamos Calcular??

Centros	Medicamentos	Efeito Favorável	Não favorável	Totais
1	Novo	29	16	**45**
1	Padrão	14	31	**45**
Totais		43	47	90
2	Novo	37	8	**45**
2	Padrão	24	21	**45**
Totais		61	29	90

```
dados=array(c(29,14,16,31,37,24,8,21),dim=c(2,2,2))
mantelhaen.test(dados,correct=F)

source(breslowday.test.R.txt)

breslowday.test(dados)
```

Supor q tabelas 2 por 2, se q=2 temos 2 tabelas do tipo 2x2

Tratamentos	Resposta $j=1$	$j=2$	Totais
$i=1$	n_{h11}	n_{h12}	n_{h1+}
$i=2$	n_{h21}	n_{h22}	n_{h2+}
Totais	n_{h+1}	n_{h+2}	n_h

h é o número da referida tabela

Totais marginais na linha nhi. são fixos

Temos interesse em testar H0: ph(1)1=ph(2)1

Condicional a H0, Nh11 tem distribuição Hipergeométrica, de onde a estatistica de MH sob H0 e para somatoria n nas q tabelas é suficientemente grande.

$$MH = \frac{\left(\sum_{h=1}^{q} n_{h11} - \sum_{h=1}^{q} e_{h11}\right)^2}{\sum_{h=1}^{q} v_{h11}} \sim \chi^2_{(1)}$$

$$OR_{MHestimado} = \frac{\sum_{h=1}^{q} \frac{n_{h11} n_{h22}}{n_h}}{\sum_{h=1}^{q} \frac{n_{h12} n_{h21}}{n_h}}$$

Exemplo

Centros	Medicamentos	Efeito		Totais
		Favorável	Não favorável	
1	Novo	29	16	**45**
1	Padrão	14	31	**45**
Totais		43	47	90
2	Novo	37	8	**45**
2	Padrão	24	21	**45**
Totais		61	29	90

Resposta : Efeito
Explicativa: Medicamento
Estratificadora: Centros

OR=3,76

OR1=4,013 e OR2=4,04

TMH= 18,41 (p<0,001)

TBreslow-Day p=0,99

OR_{MH} = 4,02

A chance de melhora dos pac que tomaram Novo foi de 4 vezes em relaçao aos que tomaram placebo

- Indica que existe associação entre as variáveis medicamento e resposta do paciente, controlando a variável centro.
- Veja que o medicamento novo apresentou proporçao maior que o padrao em relação ao efeito favoravel, ph(1)1>ph(2)1

Estatística de Mantel-Haenszel para outros casos

Y ordinal e X nominal (ni. fixos)	Estatistica Escore de Mantel-Haenszel
Y e X ordinais (ni. fixos)	Estatistica Escore de Mantel-Haenszel ou Estatistica Escore Correlação de Mantel-Haenszel
Y e X ordinais (n fixos)	Estatistica Escore Correlação de Mantel-Haenszel

Vamos Calcular??

		Efeito		
Centros	Medicamentos	Favorável	Não favorável	Totais
1	Novo	29	16	**45**
1	Padrão	14	31	**45**
Totais		43	47	90
2	Novo	37	8	**45**
2	Padrão	24	21	**45**
Totais		61	29	90

```
dados=array(c(29,14,16,31,37,24,8,21),dim=c(2,2,2))

mantelhaen.test(dados,correct=F)

    source(breslowday.test.R.txt)

    breslowday.test(dados)
```