o Cat

GEOMETRIA DAS TABELAS DE CONTINGÊNCIA 2 X 2

Este exemplar corresponde a redação final da tese devidamente corrigida e defendida pela Sra. Silvia Helena Venturoli Perri e aprovada pela Comissão Julgadora.

Campinas, 07 de dezembro de 1990

Prof. Dr. Euclydes Custódio de Lima Filho

Dissertação apresentada ao Instituto de Matemática, Estatística e Ciência da Computação, UNICAMP, como requisito parcial para obtenção do Título de Mestre em Estatística.

Aos meus pais

Ao Fernando, Rodrigo e Tiago

Aos meus mestres

AGRADECIMENTOS

Aos meus pais que me deram a vida e a minha formação.

Ao Fernando, Rodrigo e Tiago que foram os meus grandes incentivadores, dos quais muitas vezes estive longe durante o preparo desta tese.

Ao meu orientador Prof. Dr. Euclydes Custódio de Lima Filho, na pessoa de quem homenageio todos os meus mestres, pela clareza com que delineou o meu trabalho de pesquisa, pelo apoio e confiança demonstrados ao longo de toda a nossa convivência.

Aos Profs. Drs. Nemre Adas Saliba e Orlando Saliba e demais membros do Departamento de Odontologia Social da Faculdade de Odontologia do "Campus" de Araçatuba - UNESP, pelo que me incentivaram e toleraram com tanta paciência a minha ausência durante a realização deste trabalho.

A minha família pelo carinho e colaboração e a todos aqueles que de alguma forma contribuiram para a realização deste trabalho.

I N D I C E

INTRODUÇÃO

A literatura sobre a interpretação de tabelas de contingência 2x2 é volumosa e numerosas funções têm sido propostas para medir o grau de associação entre duas características.

O uso do odds ratio para medir o grau de associação tem sido empregado por diversos autores, sendo Fisher (1935) o primeiro a utilizá-lo . Muitas tentativas têm sido feitas para responder as questões propostas por Fisher (1935).

Edwards (1963) mostrou que uma medida de associação em uma tabela 2X2 deve ser alguma função do odds ratio.

Em Gart e Zweifel (1967) é analisado o vício de diversos estimadores do logito, apresentando uma comparação numérica do vício desses estimadores e das condições sobre as quais um ou outro deve ser preferido.

Mosteller (1968) avalia um conjunto de idéias sobre estimação na análise de tabelas de contingência, notando também que o odds ratio é invariante sob ambas as escalas linha e coluna.

Fienberg e Gilbert (1970) discutem idéias sobre tabelas de contingência 2X2 em termos da geometria do simplex tridimensional.

Edwards (1971) introduziu a projeção estereográfica da hiperesfera k-dimensional usando a transformação angular no espaço paramétrico da distribuição multinomial, mostrando que esta transformação fornece uma medida de distância no espaço Euclidiano.

Thompson (1972) prova que se a função de verossimilhança para proporções deduzida de dados multinomiais é plotada no espaçö Euclidiano k-dimensional, as curvas de nível do logaritmo da verossimilhança formam hiperesferas centradas no ponto de máxima veroximilhança.

Kalbfleisch e Sprott (1970) introduziram e discutiram diversos métodos de eliminar um conjunto de parâmetros nuisance de uma função de verossimilhança de modo que inferências possam ser feitas sobre o parâmetro de interesse considerado.

O principal propósito do artigo de Kalbfleisch e Sprott (1973) é dar exemplos das aplicações das funções de verossimilhança marginal e condicional para vários problemas que têm ocorrido na literatura.

Plackett (1977) considera inferências sobre o odds ratio quando a função de verossimilhança é deduzida dos totais marginais e generaliza os resultados obtidos por Kalbfleisch e Sprott (1973).

Berkson (1978) contesta a base Fisheriana do teste exato, na qual os totais marginais são "estatísticas ancilares" e portanto não produzem informações a respeito da configuração do corpo da tabela, e em lugar de usar o teste exato de Fisher ele prefere usar o teste estatístico normal.

Corsten e Kroon (1979) rejeitam a preferência expressada por Berkson (1978) de usar o teste normal ao invés do teste exato condicional de Fisher para testar igualdade de duas probabilidades binomiais.

Kempthorne (1979) salienta a importância do artigo de Berkson (1978) registrando as reações de alguns autores sobre a questão de testes de significância para tabelas 2X2 levantada por Berkson (1978).

Barnarb (1979) também comenta o artigo de Berkson (1978) relativamente a análise de dados na forma de tabelas 2X2.

Yates (1984) faz a colocação de que o teste exato de Fisher tem sido repetidamente contestado nos últimos 40 anos e recentemente com o suporte do uso de computadores para simulações. Neste seu artigo Yates discute com ajuda de exemplos simples, que estas criticas ao teste exato de Fisher são incompreendidas devido principalmente à aceitação não crítica do uso de níveis nominais para os testes de significância de Neyman-Pearson e na recusa de aceitar os argumentos condicionais sobre as marginais.

Good e Mittal (1987) discute a questão do que pode acontecer com uma medida de associação quando se faz a junção de tabelas 2X2 por adição e as condições para a homogeneidade de duas subpopulações com respeito a várias medidas de associação. Apresenta também propriedades que o odds ratio satisfaz, e algunas interpretações geométricas para tabelas 2X2.

Neste trabalho apresentaremos na primeira parte do capítulo 1 a definição do odds ratio para quatro pontos sobre um segmento de reta e suas propriedades,Beskin (1975). Na segunda parte do capitulo 1 a definição do odds ratio para uma tabela 2X2 e suas propriedades, fazendo uma analogia entre o odds ratio para tabelas 2X2 e para quatro pontos sobre um segmento de reta. Apresentaremos também as principais propriedades do odds ratio importantes para a sua escolha como medida do grau de associação.

No capítulo 2 iremos considerar o modelo multinomial para gerar tabelas de contingência 2x2 e apresentaremos algumas transformações no espaço dos parâmetros com o objetivo de encontrar estatísticas que fatorem a distribuição multinomial em partes de maneira que cada uma dependa de um parâmetro de interesse, Cox (1970).

No capítulo 3 discutiremos algumas idéias sobre tabelas de contingência 2x2 em termos da geometria do simplex tridimensional, mostrando o local dos pontos correspondentes a certas classes de tabelas 2x2, Fienberg e Gilbert (1970).

No capítulo 4 apresentaremos dois exemplos para chamar a atenção da importância de se encontrar uma transformação que fatore a multinomial em partes de modo que uma estatística dependa apenas do parâmetro de interesse e as outras independam dele, ilustrando que a transformação proposta por Fisher não é a procurada, Yates (1984).

No apêndice A colocaremos os cálculos dos jacobianos e das inversas das transformações introduzidas no capitulo 2 e também o cálculo da matriz de informação de Fisher para a 6ª transformação.

No apêndice B faremos os cálculo das frequências relativas das tabelas utilizadas nos exemplos do capítulo 4.

CAPÍTULO 1. ODDS RATIO

1.1. ODDS RATIO PARA QUATRO PONTOS SOBRE UM SEGMENTO DE RETA

Seja um segmento AB sobre uma reta e dois pontos de divisão C e D. (os pontos devem ser ordenados por ex. C é o 1º ponto e D é o 2º). Então temos duas razões simples:

- o ponto C divide o segmento AB na razão $\lambda = (ABC) = \frac{AC}{CB}$
- o ponto D divide o segmento AB na razão $\mu = (ABD) = \frac{AD}{DB}$

A razão destas razões é chamada odds ratio e denotado por:

$$\theta = (ABCD) = \frac{\lambda}{\mu} = \frac{(ABC)}{(ABD)} \qquad (1)$$

Temos a definição de odds ratio por:

$$\theta = (ABCD) = \frac{AC}{CB} : \frac{AD}{DB} = \frac{AC \cdot DB}{CB \cdot AD} \qquad (2)$$

onde $\begin{cases} \text{A e B são respectivamente pontos inicial e final} \\ \text{C e D são respectivamente o 1º e o 2º pontos de divisão} \end{cases}$

Observação: Identificando os pontos com números reais na reta orientada, podemos escrever o odds ratio como:

$$\theta = \frac{(C-A).(B-D)}{(B-C).(D-A)}$$

Os pontos de divisão podem ser arranjados com respeito ao segmento AB de 4 maneiras:

(a) os dois pontos estão dentro do segmento	A C D B	
(b) os dois pontos estão fora do segmento	C A B D	(3)
(c) o 1º ponto está dentro e o 2º está fora	A C B D	
(d) o 1º ponto está fora e o 2º está dentro .	C A D B	

Agora, pela observação anterior, determinamos o sinal do odds ratio nos casos (3):

(a) $\lambda > 0$, $\mu > 0$; $\theta > 0$

(b) $\lambda < 0$, $\mu < 0$; $\theta > 0$

(c) $\lambda > 0$, $\mu < 0$; $\theta < 0$

(d) $\lambda < 0$, $\mu > 0$; $\theta < 0$

Assim, o odds ratio é positivo se os dois pontos de divisão são "igualmente" arranjados com respeito ao segmento e negativo se estes pontos são arranjados de uma maneira diferente. Isto pode ser formulado como se segue:

- se os pares de pontos não estão separados um do outro o odds ratio é positivo, e se eles estão separados ele é negativo.

É interessante saber o que acontece ao odds ratio se qualquer dois pontos coincidem.

Vamos nos limitar aos casos quando o quarto ponto coincide com um dos outros pontos. Substituindo D por A, B e C (um depois o outro) na fórmula (2) obtemos:

$$D = A \rightarrow \theta = (ABCA) = \frac{AC \, . \, AB}{CB \, . \, AA} = \infty$$

$$D = B \rightarrow \theta = (ABCB) = \frac{AC \, . \, BB}{CB \, . \, AB} = 0$$

$$D = C \rightarrow \theta = (ABCC) = \frac{AC \, . \, CB}{CB \, . \, AC} = 1$$

PROPRIEDADE INVARIANTE DE UM ODDS RATIO COM RESPEITO AS PROJEÇÕES PARALELA E CENTRAL.

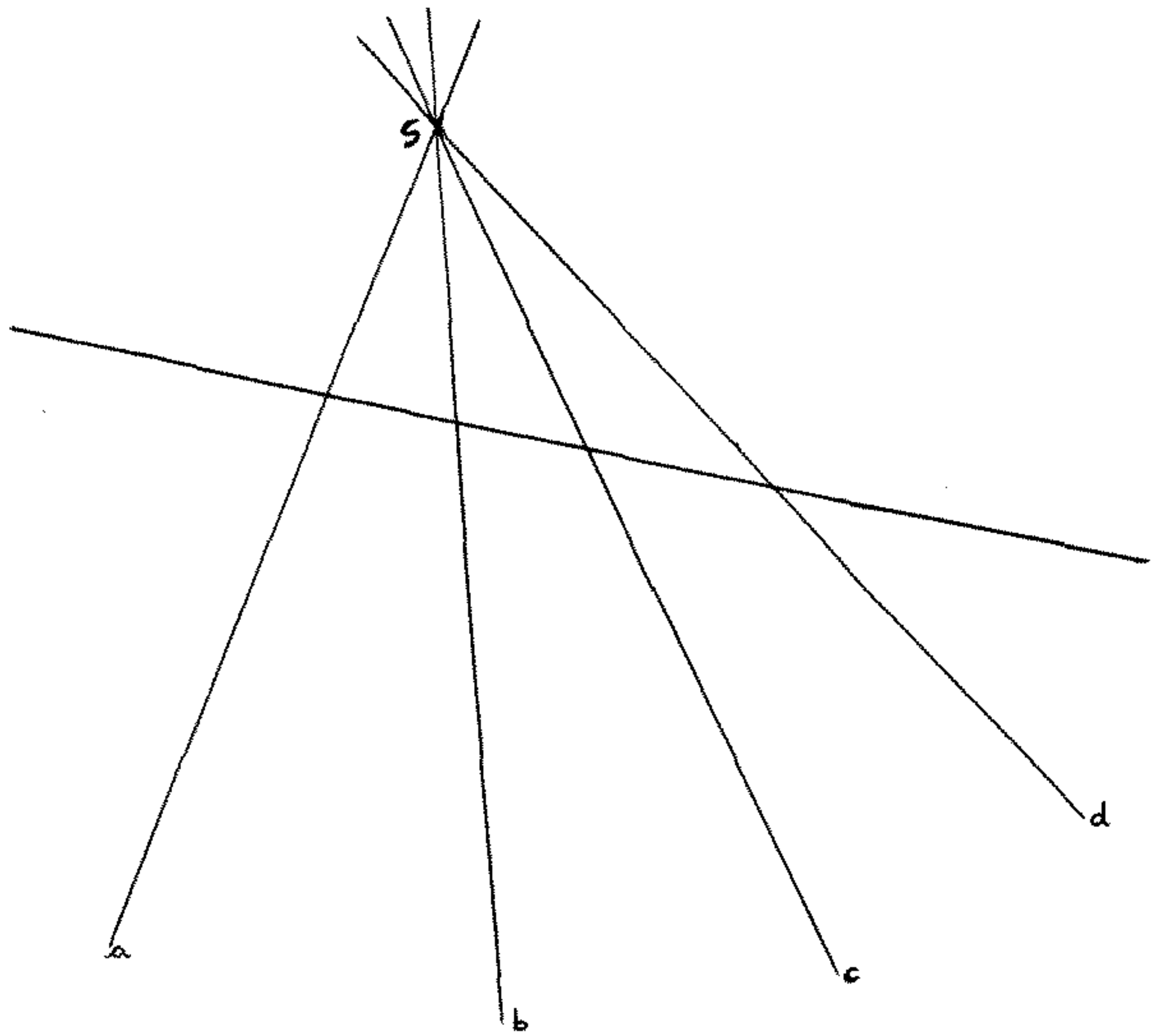
Na projeção central as retas projetadas passam através de um mesmo ponto chamado o centro da projeção.

O odds ratio não muda com a projeção central e também não muda com a paralela.

Se quatro pontos sobre uma reta são projetados sobre outra reta, seu odds ratio permanece inalterado.

O odds ratio é independente da reta a qual pertence para o mesmo feixe de retas a, b, c, d, no plano. Assim podemos introduzir a noção de odds ratio de 4 retas de uma projeção central.

O odds ratio de um feixe de retas ordenadas de uma projeção central é o odds ratio dos 4 pontos, os quais são cortados por estas retas sobre qualquer outra reta. (naturalmente, se esta reta não passar através do centro de projeção).



O odds ratio de 4 retas é denotado por (abcd).

O odds ratio de 4 retas pode ser expresso em termos dos ângulos formados por estas retas, i.e. sem o recurso da reta cortante.

$$(abcd) = \frac{\text{sen}(a,c)}{\text{sen}(c,b)} : \frac{\text{sen}(a,d)}{\text{sen}(d,b)}$$

Permutação de elementos no odds ratio

Há 24 maneiras de permutarmos 4 elementos distintos. Considerando os 4 elementos do odds ratio (ABCD), onde A e B são pontos inicial e final e C e D são 1º e 2º pontos de divisão. O odds ratio é formado por 2 pares: AB e CD, havendo uma correspondência entre os elementos dos pares, C corresponde a A e D corresponde a B.

Regra 1. O odds ratio não muda quando os 2 pares são trocados:

$$(ABCD) = (CDAB)$$

Regra 2. O odds ratio não muda se os elementos em cada par são trocados simultaneamente:

$$(ABCD) = (BADC)$$

Observação: O odds ratio não muda se duas permutações são realizadas uma depois da outra. Assim:

$$\theta = (ABCD) = (CDAB) = (BADC) = (DCBA) \qquad (4)$$

Observamos em (4) que o odds ratio não muda quando os elementos são escritos na ordem inversa, assim não podemos tratar como diferentes "os pontos inicial e final" e "os pontos de divisão", visto que os pares são equivalentes.

Podemos dizer que o odds ratio é formado por 2 pares AB e CD não ordenados mas há uma correspondência entre os elementos destes pares, no odds ratio (ABCD) o elemento C do 2º par corresponde ao elemento A do 1º par e o elemento D do 2º par corresponde ao elemento B do 1º par, ou seja :

- se a sequência do 1º par é AB → a sequência do 2º par deve ser CD.

- se a sequência do 2º par é BA → a sequência do 2º par deve ser DC.

Regra 3. Se os elementos são trocados apenas em um par, então o valor do odds ratio é trocado para seu inverso:

$$\theta = (ABCD) = \frac{AC \cdot DB}{CB \cdot AD}$$

Então:

$$(BACD) = \frac{BC \cdot DA}{CA \cdot BD} = \frac{-CB \cdot -AD}{-AC \cdot -DB} = \frac{1}{\theta}$$

Aplicando as regras 1 e 2 em $\frac{1}{\theta}$, temos:

$$\frac{1}{\theta} = (BACD) = (CDBA) = (ABDC) = (DCAB)$$

Regra 4. Se os elementos não correspondentes de diferentes pares são trocados, então o valor do odds ratio é sua diferença da unidade:

$$\theta = (ABCD) = \frac{AC \cdot DB}{CB \cdot AD}$$

Então, dividindo o segmento AB pelo ponto C e o segmento DC por B na 2ª igualdade abaixo, temos:

$$(ACBD) = \frac{AB \cdot DC}{BC \cdot AD} = \frac{-(AC+CB).(DB+BC)}{CB \cdot AD} = -\frac{ACDB + ACBC + CBDB + CBBC}{CB \cdot AD} =$$

$$= -\frac{ACDB + CB\ (\ DB + BC + CA\ )}{CB \cdot AD} = -\frac{AC \cdot DB + CB \cdot DA}{CB \cdot AD}$$

$$= -\frac{AC \cdot DB}{CB \cdot AD} + 1 = 1 - \theta$$

Aplicando as regras 1 e 2 em $1-\theta$, temos:

$$1-\theta = (ACBD) = (BDAC) = (CADB) = (DBCA)$$

Aplicando a regra 3 para $1-\theta$ e depois as regras 1 e 2, temos:

$$\frac{1}{1-\theta} = (CABD) = (BDCA) = (ACDB) = (DBAC)$$

Aplicando a regra 4 para $\frac{1}{\theta}$ e depois as regras 1 e 2, temos:

$$\frac{\theta - 1}{\theta} = (BCAD) = (ADBC) = (CBDA) = (DACB)$$

Aplicando a regra 3 no odds ratio $\frac{\theta - 1}{\theta}$ e depois as regras 1 e 2, temos:

$$\frac{\theta}{\theta - 1} = (CBAD) = (ADCB) = (DABC) = (BCDA)$$

Sabemos que 4 elementos podem ser permutados em 24 maneiras, mas somente 6 odds ratio diferentes correspondem a eles. Resumindo os resultados abaixo temos:

$$\begin{aligned}
\theta &= (ABCD) = (CDAB) = (BADC) = (DCBA) \\
\frac{1}{\theta} &= (BACD) = (CDBA) = (ABDC) = (DCAB) \\
1-\theta &= (ACBD) = (BDAC) = (CADB) = (DBCA) \\
\frac{1}{1-\theta} &= (CABD) = (BDCA) = (ACDB) = (DBAC) \\
\frac{\theta - 1}{\theta} &= (BCAD) = (ADBC) = (CBDA) = (DACB) \\
\frac{\theta}{\theta - 1} &= (CBAD) = (ADCB) = (BCDA) = (DABC)
\end{aligned} \tag{5}$$

Há alguns arranjos especiais de 4 pontos diferentes sobre uma reta, de maneira que a estes arranjos correspondem menos de 6 odds diferentes, por exemplo:

a) $\theta = \frac{1}{\theta} \Rightarrow \theta^2 = 1 \Rightarrow \begin{cases} \theta = 1 \\ \theta = -1 \end{cases}$

O valor θ=1 deve ser rejeitado pois ele corresponde a uma quadra degenerada (ABCC), (θ=1 $\Leftrightarrow$ D=C).

$$\text{b)}\ \theta = 1 - \theta \quad \Rightarrow \quad \theta = \frac{1}{2}$$

$$\text{c)}\ \theta = \frac{1}{1-\theta} \Rightarrow \theta^2 - \theta + 1 = 0 \quad \Rightarrow \text{ raízes imaginárias}$$

$$\text{d)}\ \theta = \frac{\theta-1}{\theta} \quad \Rightarrow \theta^2 - \theta + 1 = 0 \quad \Rightarrow \text{ raízes imaginárias}$$

$$\text{e)}\ \theta = \frac{\theta}{\theta-1} \quad \Rightarrow \quad \theta^2 - 2\theta = 0 \quad \Rightarrow \begin{cases} \theta = 0 \\ \theta = 2 \end{cases}$$

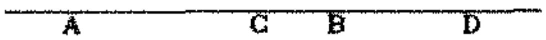
O valor θ = 0 deve ser rejeitado pois corresponde a uma quadra degenerada (ABCB), (θ = 0 $\Leftrightarrow$ D = B).

Para cada um dos valores de θ encontrados nós listamos os seus 6 correspondentes:

θ	$\frac{1}{\theta}$	$1-\theta$	$\frac{1}{1-\theta}$	$\frac{\theta-1}{\theta}$	$\frac{\theta}{\theta-1}$
-1	-1	2	$\frac{1}{2}$	2	$\frac{1}{2}$
$\frac{1}{2}$	2	$\frac{1}{2}$	2	-1	-1
2	$\frac{1}{2}$	-1	-1	$\frac{1}{2}$	2

Como em todas as linhas temos os mesmos números, então os três valores encontrados para θ representam a mesma solução, ou seja, correspondem a mesma quadra de pontos ordenados de maneira diferentes. É a chamada quadra harmônica, Beskin (1975).

Quando $\theta = -1$, temos um segmento AB, um ponto interior C que o divide em uma certa razão λ, e um ponto exterior D que divide o mesmo segmento na razão μ ($\mu = \lambda$ em valor absoluto). Esses quatro pontos são chamados quadra harmônica.



(6)

$$\lambda = \frac{AC}{CB} = \frac{3}{1} = 3$$

$$\mu = \frac{AD}{DB} = \frac{6}{-2} = -3$$

Numa quadra harmônica os pares são equivalentes, se C e D são os pontos extremos então A e B são os pontos de divisão.

Por ex., em (6) se A divide o segmento CD na razão $-\frac{1}{2}$, então B divide o segmento CD na razão $\frac{1}{2}$.

As expressões do lado esquerdo das fórmulas (5) possuem uma propriedade de grupo. Se qualquer dessas expressões é substituída em qualquer ordem por θ nós obtemos uma delas novamente.

Usando a notação: $\begin{cases} a_1 = \theta, & a_2 = \dfrac{1}{\theta}, & a_3 = 1-\theta, \\ a_4 = \dfrac{1}{1-\theta}, & a_5 = \dfrac{\theta-1}{\theta}, & a_6 = \dfrac{\theta}{\theta-1} \end{cases}$

onde multiplicar a_i por a_j ($a_i \circ a_j$) significa substituir a_j em lugar de θ no a_i, por exemplo:

$$a_3 \circ a_4 = 1 - \frac{1}{1-\theta} = \frac{-\theta}{1-\theta} = \frac{\theta}{\theta-1} = a_6$$

Ou seja $a_i \odot a_j = f_i[f_j(\theta)] = a_k$, temos o seguinte quadro:

$\odot$	a_1	a_2	a_3	a_4	a_5	a_6
a_1	a_1	a_2	a_3	a_4	a_5	a_6
a_2	a_2	a_1	a_4	a_3	a_6	a_5
a_3	a_3	a_5	a_1	a_6	a_2	a_4
a_4	a_4	a_6	a_2	a_5	a_1	a_3
a_5	a_5	a_3	a_6	a_1	a_4	a_2
a_6	a_6	a_4	a_5	a_2	a_3	a_1

1.2. ODDS RATIO PARA UMA TABELA DE CONTINGÊNCIA 2 X 2

Definiremos agora a medida de associação θ (odds ratio) para uma tabela de contingência 2 x 2 .

Considere uma população de indivíduos, que pode ser classificada segundo duas características A e B (A^c e B^c são os complementares de A e B), gerando a tabela:

	A	A^c
B	p_1	p_2
B^c	p_3	p_4

onde p_1 ,p_2 ,p_3 e p_4 são os parâmetros populacionais correspondentes às proporções das categorias AB , A^cB , AB^c e A^cB^c respectivamente.

O interesse estatístico é apresentar as propriedades do odds ratio, importantes para a sua escolha como medida do grau de associação.

Por definição o odds ratio é dado por:

$$\theta = \frac{p_1 \, p_4}{p_2 \, p_3}$$

Fazendo uma analogia entre a definição de θ para quatro pontos sobre um segmento de reta e para uma tabela 2x2, podemos escrever:

$$\theta \cong (A\ A^cB\ B^c) \cong \frac{A\ B\ .\ B^cA^c}{B\ A^c.\ A\ B^c}$$

onde agora as distâncias AB , BA^c , AB^c e B^cA^c são respectivamente as probabilidades das categorias AB , BA^c , AB^c e B^cA^c as quais são dadas por p_1 , p_2, p_3 e p_4 nesta ordem. Assim:

$$\theta \cong (A\ A^cB\ B^c) = \frac{p_1\ p_4}{p_2\ p_3}$$

Regra 1. O odds ratio não se altera quando trocamos as linhas por colunas, ou seja quando os dois pares são trocados.(usamos a matriz transposta). É uma medida simétrica.

	A	A^c
B	p_1	p_2
B^c	p_3	p_4

→

	B	B^c
A	p_1	p_3
A^c	p_2	p_4

$$\theta \cong (\ A\ A^cB\ B^c) \cong (\ B\ B^cA\ A^c) = \frac{p_1\ p_4}{p_2\ p_3}$$

Regra 2. O odds ratio não muda quando trocamos os elementos em ambos os pares simultaneamente.

	A	A^c
B	p_1	p_2
B^c	p_3	p_4

→

	A^c	A
B^c	p_4	p_3
B	p_2	p_1

$$\theta \cong (\,A\ A^cB\ B^c) \cong (\,A^cA\ B^cB\,) = \frac{p_1\ p_4}{p_2\ p_3}$$

O odds ratio não se altera se duas permutações são realizadas uma após a outra:

	A	A^c
B	p_1	p_2
B^c	p_3	p_4

→

	A^c	A
B^c	p_4	p_3
B	p_2	p_1

→

	B	B^c
A	p_1	p_3
A^c	p_2	p_4

→

	B^c	B
A^c	p_4	p_2
A	p_3	p_1

$$\theta \cong (\,A\ A^cB\ B^c) \cong (\,A^cA\ B^cB\,) \cong (\,B\ B^cA\ A^c) \cong (\,B^cB\ A^cA\,) = \frac{p_1\ p_4}{p_2\ p_3}$$

Regra 3. Se os elementos são trocados apenas em um par, então o valor do odds ratio é trocado pelo seu inverso:

	A	A^c
B	p_1	p_2
B^c	p_3	p_4

⇒

	A^c	A
B	p_2	p_1
B^c	p_4	p_3

$$(\,A^cA\ B\ B^c) = \frac{p_2\ p_3}{p_1\ p_4} = \frac{1}{\theta}$$

$$\Rightarrow \quad \begin{array}{c|c|c|} & A & A^c \\ \hline B^c & p_3 & p_4 \\ \hline B & p_1 & p_2 \\ \hline \end{array} \qquad (\ A\ \ A^cB^cB\ ) = \frac{p_3\ p_2}{p_1\ p_4} = \frac{1}{\theta}$$

Aplicando as regras 1 e 2 em $\frac{1}{\theta}$, temos

$$\frac{1}{\theta} \cong (\ A^cA\ \ B\ B^c) \cong (\ A\ \ A^cB^cB\ ) \cong (\ B\ \ B^cA^cA\ ) \cong (\ B^cB\ \ A\ A^c) = \frac{p_2\ p_3}{p_1\ p_4}$$

Regra 4. Se os elementos não correspondentes de diferentes pares são trocados, então o valor do odds ratio é sua diferença da unidade.

$$\begin{array}{c|c|c|} & A & A^c \\ \hline B & p_1 & p_2 \\ \hline B^c & p_3 & p_4 \\ \hline \end{array} \ \Rightarrow \ \begin{array}{c|c|c|} & A & B \\ \hline A^c & \quad & \quad \\ \hline B^c & & \\ \hline \end{array} \ \Rightarrow \ \begin{array}{c|c|c|} & A^c & B^c \\ \hline A & \quad & \quad \\ \hline B & & \\ \hline \end{array}$$

$$\theta \ \Rightarrow \ 1-\theta \qquad \Rightarrow \ \begin{array}{c|c|c|} & B^c & A^c \\ \hline B & \quad & \quad \\ \hline A & & \\ \hline \end{array} \ \Rightarrow \ \begin{array}{c|c|c|} & B & A \\ \hline B^c & \quad & \quad \\ \hline A^c & & \\ \hline \end{array}$$

Aplicando a Regra 3 em $1-\theta$, e depois as regras 1 e 2, temos :

$$\begin{array}{c|c|c|} & A & A^c \\ \hline B & p_1 & p_2 \\ \hline B^c & p_3 & p_4 \\ \hline \end{array} \ \Rightarrow \ \begin{array}{c|c|c|} & B & A \\ \hline A^c & \quad & \quad \\ \hline B^c & & \\ \hline \end{array} \ \Rightarrow \ \begin{array}{c|c|c|} & A^c & B^c \\ \hline B & \quad & \quad \\ \hline A & & \\ \hline \end{array}$$

$\theta \Rightarrow \dfrac{1}{1-\theta}$

$\Rightarrow$

	B^c	A^c
A		
B		

$\Rightarrow$

	A	B
B^c		
A^c		

Aplicando a Regra 4 em $\dfrac{1}{\theta}$, e depois as regras 1 e 2, temos:

	A	A^c
B	p_1	p_2
B^c	p_3	p_4

$\Rightarrow$

	A^c	B
A		
B^c		

$\Rightarrow$

	A	B^c
A^c		
B		

$\Rightarrow$

$\theta \Rightarrow \dfrac{\theta-1}{\theta}$

$\Rightarrow$

	B^c	A
B		
A^c		

$\Rightarrow$

	B	A^c
B^c		
A		

Aplicando a Regra 3 em $\dfrac{\theta-1}{\theta}$, e depois as regras 1 e 2, temos:

	A	A^c
B	p_1	p_2
B^c	p_3	p_4

$\Rightarrow$

	B	A^c
A		
B^c		

$\Rightarrow$

	A	B^c
B		
A^c		

$\Rightarrow$

$\theta \Rightarrow \dfrac{\theta}{\theta-1}$

$\Rightarrow$

	B^c	A
A^c		
B		

$\Rightarrow$

	A^c	B
B^c		
A		

A aplicação da Regra 4 e as consequentes aplicações sob seu resultado não tem sentido estatistico, por causa das definições das caracteristicas AB , A^cB , AB^c, A^cB^c.

Em termos de tabelas 2x2, as propriedades do odds ratio importantes para a sua escolha como medida do grau de associação são as seguintes:

1. O odds ratio não se altera quando trocamos linhas por colunas, ou seja, usamos a matriz transposta. É uma medida simétrica. (Regra 1)

	A	A^c
B	p_1	p_2
B^c	p_3	p_4

$\Rightarrow$

	B	B^c
A	p_1	p_3
A^c	p_2	p_4

$$\theta = \frac{p_1\ p_4}{p_2\ p_3}$$

2. O odds ratio não muda quando trocamos os elementos em ambos os pares simultaneamente. (Regra 2)

	A	A^c
B	p_1	p_2
B^c	p_3	p_4

$\Rightarrow$

	A^c	A
B^c	p_4	p_3
B	p_2	p_1

$$\theta = \frac{p_1\ p_4}{p_2\ p_3}$$

3. Se os elementos são trocados apenas em um par, então o valor do odds ratio é trocado pelo seu inverso. (Regra 3)

	A	A^c
B	p_1	p_2
B^c	p_3	p_4

$\Rightarrow$

	A^c	A
B	p_2	p_1
B^c	p_4	p_3

$$\theta' = \frac{p_2 \; p_3}{p_1 \; p_4} = \frac{1}{\theta}$$

$\Rightarrow$

	A	A^c
B^c	p_3	p_4
B	p_1	p_2

$$\theta' = \frac{p_3 \; p_2}{p_1 \; p_4} = \frac{1}{\theta}$$

4. Se multiplicarmos uma ou duas linhas (e ou colunas) por uma constante positiva, o odds ratio não se altera , ou seja, o odds ratio é invariante sob ambas as escalas linha e coluna.

	A	A^c
B	p_1	p_2
B^c	p_3	p_4

$\Rightarrow$

		c_1	c_2
		A	A^c
r_1	B	$r_1 c_1 p_1$	$r_1 c_2 p_2$
r_2	B^c	$r_2 c_1 p_3$	$r_2 c_2 p_4$

$$\theta = \frac{p_1 \; p_4}{p_2 \; p_3}$$

Assim temos que o odds ratio é uma importante medida do grau de associação por satisfazer as propriedades anteriores e muito mais importante do que isto é o fato de que qualquer medida de associação deve ser alguma função monotônica do odds ratio, para ter as suas mesmas propriedades ótimas de 1 a 4. [ver referências Good e Mittal (1987) e Edwards (1963)].

CAPITULO 2. CARACTERIZAÇÃO DA MULTINOMIAL

Iremos considerar o modelo multinomial para gerar tabelas de contingência 2x2, Cox (1970).

Gostaríamos de encontrar uma transformação (bijeção de classe C^1 com inversa de classe C^1 no espaço dos parâmetros, ou seja difeomorfismo de classe C^1) de maneira que as estatísticas fatorassem a distribuição multinomial em partes, cada uma das quais dependendo apenas de um parâmetro de interesse. Isso seria o ideal, mas como estamos interessados na medida de associação θ (odds ratio), gostaríamos de encontrar uma transformação de maneira que as estatísticas fatorassem a multinomial em partes, uma dependendo apenas do parâmetro de interesse θ e as outras dependendo dos outros parâmetros, ou seja uma estatística que dependa apenas de θ mas nuisance para os outros parâmetros e as outras independendo de θ.

Como a distribuição multinomial depende de três parâmetros, precisamos de pelo menos três estatísticas para podermos fatorá-la.

Encontrar uma transformação de maneira que existam estatísticas tais que uma estatística dependa apenas de θ (parâmetro de interesse) sendo nuisance para os outros parâmetros e as outras não dependam de θ, não é uma tarefa fácil.

Queremos que a distribuição dessa estatística que só depende de θ não mude se os outros parâmetros mudarem, que ela só se altere com a mudança do valor de θ (identificabilidade de θ).

Será que existe estatisticas que fatorem a multinomial em partes de maneira que cada estatistica dependa apenas de um parâmetro de interesse ?

Como estamos interessados na medida de associação θ (odds ratio), a pergunta poderia ser reformulada da seguinte maneira:

Será que existe uma transformação tal que as estatisticas fatorem a multinomial em partes, de maneira que uma estatistica dependa apenas do parâmetro de interesse θ e as outras estatisticas dependam dos outros parâmetros, ou seja uma estatistica dependente de θ mas muisance para os outros parâmetros e as outras independentes de θ ?

Como não temos um teorema que garanta a existência de tal transformação, investigaremos 6 transformações cada uma das quais importantes para cada caso particular com o objetivo de encontrá-la.

Considere uma população de individuos que pode ser classificada segundo duas caracteristicas A e B (A^c e B^c são os complementos de A e B), gerando a tabela:

	A	A^c
B	p_1	p_2
B^c	p_3	p_4

onde p_1 , p_2 e p_3 são os parâmetros populacionais desconhecidos correspondentes as proporções das categorias AB , A^cB e AB^c.

Considere o seguinte experimento: retira-se uma amostra de tamanho n da população, com reposição, classificando-a nas categorias. Os dados para este experimento são (x_1,x_2,x_3) e seguem o modelo multinomial (n ,p_1 ,p_2 ,p_3).

Tem-se a seguinte tabela 2x2 :

	A	A^c	
B	x_1	x_2	n_1
B^c	x_3	x_4	
	m_1		n

A distribuição conjunta das variáveis X_1,X_2,X_3 é dada por:

$$P(x_1,x_2,x_3 \;/\; n,p_1,p_2,p_3) = \frac{n!}{x_1!x_2!x_3!(n-x_1-x_2-x_3)!}\, p_1^{x_1}\, p_2^{x_2}\, p_3^{x_3}\, p^{n-x}$$

onde $p_i>0$, $\sum_{i=1}^{4} p_i = 1$ $x = x_1+ x_2+ x_3$ e $p = p_1+ p_2+ p_3$

$1^{\underline{a}}$ TRANSFORMAÇÃO

A decomposição a seguir foi proposta por Fisher (1935) para testar a independência entre as características A e B, ou seja, se o fato de o indivíduo ter a característica A ou A^c não influencia a probabilidade de ter a característica B e vice-versa.

A hipótese a ser testada é:

$$H_0) \; P(B/A) = P(B/A^c) \quad \text{ou} \quad P(B^c/A) = P(B^c/A^c)$$

segundo a estrutura paramétrica da tabela temos:

$$H_0) \quad \frac{p_1}{p_1+p_3} = \frac{p_2}{p_2+p_4} \qquad \text{ou} \qquad \frac{p_3}{p_1+p_3} = \frac{p_4}{p_2+p_4}$$

Das relações acima segue que a hipótese nula pode ser equivalentemente escrita como: $H_0)\ \theta = 1$. Assim θ é o parâmetro de interesse e os outros (ϕ,ψ) serão os parâmetros nuisance.

Trabalharemos com a estatística (x_1, n_1, m_1) em lugar (x_1, x_2, x_3). O espaço paramétrico do modelo é o conjunto:

$$\Omega = \left\{ (p_1,p_2,p_3) \,/\, p_1+ p_2+ p_3 \le 1 \;,\; p_1> 0 \;,\; p_2> 0 \;,\; p_3> 0 \right\}$$

Consideraremos a seguinte reparametrização do modelo, usada por Fisher (1935):

$$\theta = \frac{p_1\ p_4}{p_2\ p_3} \quad , \quad \phi = \frac{p_3}{1-p_1-p_2} \quad , \psi = p_1 + p_2$$

A condição de bijeção entre (θ,ϕ,ψ) e (p_1,p_2,p_3) está satisfeita, (ver Apêndice A.1). O novo espaço paramétrico é :

$$\Omega = \left\{ (\ \theta,\ \phi,\ \psi\ ) \text{ tal que } \theta > 0\ ,\ 0 < \phi < 1\ ,\ 0 < \psi < 1 \right\}$$

e neste caso temos que θ, ϕ e ψ são parâmetros de variação independente.

A tabela acima pode ser escrita como:

	A	A^c	
B	$\frac{\theta\ \phi\ \psi}{1-\phi\ +\ \theta\phi}$	$\frac{(1-\phi)\ \psi}{1-\phi\ +\ \theta\phi}$	ψ
B^c	$\phi\ (1-\psi)$	$(1-\phi)(1-\psi)$	$1-\psi$

A transformação:

$$\psi = p_1 + p_2 \quad \rightarrow \quad p_2 = \psi - p_1 \qquad (1)$$

$$\phi = \frac{p_3}{1-p_1-p_2} \quad \rightarrow \quad p_3 = \phi(1-\psi) \qquad (2)$$

$$\theta = \frac{p_1\ p_4}{p_2\ p_3} \qquad (3)$$

Mas $p_4 = 1-p_1-p_2-p_3$, então substituindo (1) e (2) temos:

$$p_4 = 1-\psi - \phi(1-\psi) \;\Rightarrow\; p_4 = (1-\psi)(1-\phi) \qquad (4)$$

Substituindo (1) , (2) e (4) em (3) temos:

$$\theta = \frac{p_1(1-\psi)(1-\phi)}{(\psi-p_1)\phi(1-\psi)} \;\Rightarrow\; \theta\,\psi\,\phi - \theta\, p_1\phi = p_1 - p_1\phi \;\Rightarrow$$

$$\Rightarrow\; p_1(1- \phi + \phi\,\theta) = \theta\,\phi\,\psi$$

Então $$p_1 = \frac{\theta\,\phi\,\psi}{1- \phi + \phi\,\theta} \qquad (5)$$

Substituindo (5) em (1) temos:

$$p_2 = \frac{\psi\,(1-\phi)}{1-\phi + \phi\,\theta}$$

A distribuição de (x_1,x_2,x_3) de acordo com a reparametrização é dada por:

$P\,(\,x_1,x_2,x_3 \,/\, n,p_1,p_2,p_3\,)$ que pode ser escrita como:

$$\frac{n!(x_1+x_2)!(n-x_1-x_2)!}{(x_1+x_2)!(n-x_1-x_2)!x_1!x_2!x_3!(n-x_1-x_2-x_3)!} \cdot \left(\frac{\theta\,\phi\,\psi}{1-\phi + \theta\,\phi}\right)^{x_1} \cdot$$

$$\cdot\left(\frac{(1- \phi)\,\psi}{1- \phi + \theta\,\phi}\right)^{x_2} \cdot \left[\phi\,(1- \psi\,)\right]^{x_3} \cdot \left[(1- \psi)(1-\phi)\right]^{n-x_1-x_2-x_3} =$$

$$= \binom{n}{x_1+x_2}\binom{x_1+x_2}{x_1}\binom{n-x_1-x_2}{x_3}\,\psi^{x_1+x_2}\,(\,1- \psi\,)^{n-x_1-x_2} \;.$$

$$.\; \phi^{x_1+x_3}\,(\,1- \phi\,)^{n-x_1-x_3} \cdot \frac{\theta^{x_1}}{(1- \phi + \theta\,\phi\,)^{x_1+x_2}} =$$

$$= \binom{n}{x_1+x_2} \psi^{x_1+x_2} (1- \psi)^{n-x_1-x_2} \cdot \frac{\binom{x_1+x_2}{x_1} \binom{n-x_1-x_2}{x_3} \theta^{x_1}}{\displaystyle\sum_{k=0}^{x_1+x_3} \binom{x_1+x_2}{k} \binom{n-x_1-x_2}{x_1+x_3-k} \theta^{k}} \cdot$$

$$\frac{\phi^{x_1+x_3} (1- \phi)^{n-x_1-x_3}}{(1- \phi + \theta \phi)^{x_1+x_2}} \cdot \sum_{k=0}^{x_1+x_3} \binom{x_1+x_2}{k} \binom{n-x_1-x_2}{x_1+x_3-k} \theta^{k}$$

$$= P(x_1+x_2 / \psi) . P(x_1+x_3 / x_1+x_2 , \phi , \theta) . P(x_1 / x_1+x_3 , x_1+x_2 ,\theta)$$

$$= P(n_1 / \psi).P(m_1 / n_1,\phi,\theta).P(x_1 / m_1,n_1,\theta)$$

$$= P(x_1 ,n_1,m_1 / \theta,\phi,\psi)$$

Esta transformação não é a que procuramos pois a distribuição marginal de x_1+x_3 condicionada a x_1+x_2 envolve o parâmetro θ e portanto deve conter informação sobre ele.

2ª TRANSFORMAÇÃO

A transformação a seguir vai nos permitir escrever a multinomial como produto de três binomiais, cada uma das quais dependendo apenas de um parâmetro, só que o parâmetro de interesse θ não está explicitado. Assim essa transformação não é a que estamos procurando pois quando escrevemos esta transformação em função de θ não temos mais cada estatística dependendo apenas de um parâmetro de interesse.

Seja a transformação :

$$\gamma_1 = p_1 \quad , \gamma_2 = \frac{p_2}{1- p_1} \quad , \quad \gamma_3 = \frac{p_3}{1- p_1 - p_2}$$

é 1-1 e sobre. $\Omega = (0,1)^3$, (ver Apêndice A.2)

Temos: $\gamma_1 = p_1 \quad \Rightarrow \quad p_1 = \gamma_1$

$$\gamma_2 = \frac{p_2}{1- p_1} \quad \Rightarrow \quad p_2 = \gamma_2(1-\gamma_1)$$

$$\gamma_3 = \frac{p_3}{1- p_1 - p_2} \quad \Rightarrow \quad p_3 = \gamma_3 \,[\, 1- \gamma_1 - \gamma_2(1- \gamma_1) \,]$$

$$\Rightarrow \quad p_3 = \gamma_3(1- \gamma_2)\,(1- \gamma_1)$$

Assim podemos escrever, $P(x_1,x_2,x_3 \,/\, p_1,p_2,p_3,n)$ como:

$$\frac{n!}{x_1!x_2!x_3!(n-x_1-x_2-x_3)!}\;\gamma_1^{x_1}\,\gamma_2^{x_2}\,(1-\gamma_1)^{x_2}\,\gamma_3^{x_3}\,(1-\gamma_1)^{x_3}\,(1-\gamma_2)^{x_3}.$$

$$.\left[\,1-\gamma_1-\gamma_2(1-\gamma_1)-\gamma_3(1-\gamma_1)(1-\gamma_2)\,\right]^{n-x_1-x_2-x_3} =$$

$$= \frac{n!}{x_1!x_2!x_3!(n-x_1-x_2-x_3)!}\;\gamma_1^{x_1}\,\gamma_2^{x_2}\,(1-\gamma_1)^{x_2}\,\gamma_3^{x_3}\,(1-\gamma_1)^{x_3}\,(1-\gamma_2)^{x_3}.$$

$$.\left[\,(1-\gamma_1)(1-\gamma_2)(1-\gamma_3)\,\right]^{n-x_1-x_2-x_3} =$$

$$= \frac{n!(n-x_1)!(n-x_1-x_2)!}{x_1!(n-x_1)!x_2!(n-x_1-x_2)!x_3!(n-x_1-x_2-x_3)!}\,.\gamma_1^{x_1}\,(1-\gamma_1)^{n-x_1}\,.$$

$$.\;\gamma_2^{x_2}\,(1-\gamma_2)^{n-x_1-x_2}\;.\;\gamma_3^{x_3}\,(1-\gamma_3)^{n-x_1-x_2-x_3} =$$

$$= \binom{n}{x_1}\binom{n-x_1}{x_2}\binom{n-x_1-x_2}{x_3}\,.\gamma_1^{x_1}\,(1-\gamma_1)^{n-x_1}.\;\gamma_2^{x_2}\,(1-\gamma_2)^{n-x_1-x_2}\,.$$

$$.\;\gamma_3^{x_3}\,(1-\gamma_3)^{n-x_1-x_2-x_3} =$$

$$= B(x_1 \,/\, n,\gamma_1)\;.\;B(x_2 \,/\, n-x_1,\gamma_2)\;.\;B(x_3 \,/\, n-x_1-x_2,\gamma_3)$$

3º TRANSFORMAÇÃO

Com esta transformação também conseguimos fatorar a multinomial em produto de binomiais, onde cada estatística depende apenas de um parâmetro mas não conseguimos explicitar o parâmetro de interesse θ e ainda manter essa independência.

Seja a transformação:

$$\psi = p_1+p_2 \quad , \delta = \frac{p_1}{p_1+p_2} \quad , \quad \phi = \frac{p_3}{1-p_1-p_2}$$

é 1-1 e sobre. $\Omega = (0,1)^3$, (ver Apêndice A.3).

Temos:

$$\psi = p_1+p_2 \Rightarrow p_2 = \psi - \psi\,\delta \Rightarrow p_2 = \psi\,(1-\delta)$$

$$\delta = \frac{p_1}{p_1+p_2} \Rightarrow p_1 = \psi\,\delta$$

$$\phi = \frac{p_3}{1-p_1-p_2} \Rightarrow p_3 = \phi\,[\,1-\psi\,\delta-\psi\,(1-\delta)\,] \Rightarrow p_3 = \phi\,(\,1-\psi\,)$$

Mas: $p_4 = 1-p_1-p_2-p_3 \Rightarrow p_4 = 1-\psi\,\delta-\psi\,(1-\delta)-\phi\,(1-\psi)$

$$\Rightarrow p_4 = 1-\psi\,\delta-\psi+\psi\,\delta-\phi(1-\psi)$$

$$\Rightarrow p_4 = (1-\psi)(1-\phi)$$

A probabilidade multinomial $P(x_1,x_2,x_3/\ n,p_1,p_2,p_3)$

pode ser escrita como:

$$\frac{n!}{x_1!x_2!x_3!(n-x_1-x_2-x_3)!}(\psi\,\delta)^{x_1}\left[\psi(1-\delta)\right]^{x_2}\left[\phi(1-\psi)\right]^{x_3}.$$

$$.\left[(1-\psi)(1-\phi)\right]^{n-x_1-x_2-x_3} =$$

$$= \frac{n!(x_1+x_2)!(n-x_1-x_2)!}{(x_1+x_2)!(n-x_1-x_2)!x_1!x_2!x_3!(n-x_1-x_2-x_3)!}\psi^{x_1}\ \delta^{x_1}\ \psi^{x_2}\ (1-\delta)^{x_2}.$$

$$.\phi^{x_3}\ (1-\psi)^{x_3}\ (1-\psi)^{n-x_1-x_2-x_3}\ (1-\phi)^{n-x_1-x_2-x_3} =$$

$$= \binom{n}{x_1+x_2}\binom{x_1+x_2}{x_1}\binom{n-x_1-x_2}{x_3}\psi^{x_1+x_2}\ (1-\psi)^{n-x_1-x_2}\ .\ \delta^{x_1}\ (1-\delta)^{x_2}.$$

$$.\phi^{x_3}\ (1-\phi)^{n-x_1-x_2-x_3} =$$

$$= B\left(x_1+x_2\ /\ n\ ,\ \psi\right)\ .\ B\left(x_1\ /\ x_1+x_2\ ,\ \delta\right)\ .\ B\left(x_3\ /\ n-x_1-x_2\ ,\ \phi\right)$$

4º TRANSFORMAÇÃO

Essa transformação é importante pois ela fatora a multinomial em produto de binomiais sendo que a distribuição de x_1+x_2, depende de ψ e as outras estatísticas estão em função do logito, mas não é a que procuramos, Cox (1970).

Seja a transformação :

$$\psi = p_1 + p_2 \quad , \alpha = \ln \frac{p_3}{p_4} \;\Rightarrow\; \alpha = \ln \phi \quad , \lambda = \ln \frac{p_1 \; p_4}{p_2 \; p_3} \;\Rightarrow \lambda = \ln \theta$$

é 1-1 e sobre. $\Omega = \left\{ (\psi,\alpha,\lambda) \text{ tal que } 0 < \psi < 1 \; , -\infty < \alpha,\lambda < \infty \right\}$, (ver Apêndice A.4).

Temos:

$$\psi = p_1 + p_2 \quad \Rightarrow p_1 = \psi - p_2 \qquad (1)$$

$$e^{\alpha} = \frac{p_3}{p_4} \quad \Rightarrow p_3 = e^{\alpha} \; p_4 \qquad (2)$$

$$e^{\lambda} = \frac{p_1 \; p_4}{p_2 \; p_3} \quad \Rightarrow p_2 = \frac{p_1 \; p_4}{e^{\lambda} \; p_3} \qquad (3)$$

Substituindo (1) e (2) em (3) temos:

$$p_2 = \frac{(\psi - p_2) \; p_4}{e^{\lambda} \; e^{\alpha} \; p_4} \quad \Rightarrow p_2 \; e^{\lambda} \; e^{\alpha} = \psi - p_2 \quad \Rightarrow p_2 (1 + e^{\lambda+\alpha}) = \psi$$

$$p_2 = \frac{\psi}{1 + e^{\lambda+\alpha}} \qquad (4)$$

Substituindo (4) em (1) temos:

$$p_1 = \psi - p_2 \quad \rightarrow \quad p_1 = \psi - \frac{\psi}{1 + e^{\lambda+\alpha}} \quad \rightarrow p_1 = \frac{\psi\, e^{\lambda+\alpha}}{1 + e^{\lambda+\alpha}}$$

Mas $\quad p_4 = 1 - p_1 - p_2 - p_3 \qquad (5)$

Substituindo (1) e (2) em (5) temos:

$$p_4 = 1 - \psi - e^{\alpha} p_4$$

$$p_4 (1 + e^{\alpha}) = 1 - \psi \quad \rightarrow \quad p_4 = \frac{1 - \psi}{1 + e^{\alpha}}$$

Substituindo o valor de p_4 em (2) temos :

$$p_3 = \frac{e^{\alpha} (1 - \psi)}{1 + e^{\alpha}}$$

Assim :

$$P(x_1, x_2, x_3 \,/\, n, p_1, p_2, p_3) = \frac{n!}{x_1!x_2!x_3!(n-x_1-x_2-x_3)!}\, p_1^{x_1}\, p_2^{x_2}\, p_3^{x_3} \,.$$

$$.\ (1-p_1-p_2-p_3)^{n-x_1-x_2-x_3} \qquad =$$

$$= \frac{n!\,(x_1+x_2)!\,(n-x_1-x_2)!}{(x_1+x_2)!\,(n-x_1-x_2)!\,x_1!\,x_2!\,x_3!\,(n-x_1-x_2-x_3)!} \left(\frac{\psi\, e^{\lambda+\alpha}}{1 + e^{\lambda+\alpha}} \right)^{x_1} \cdot$$

$$\cdot \left(\frac{\psi}{1 + e^{\lambda+\alpha}} \right)^{x_2} \cdot \left(\frac{e^{\alpha}\,(1 - \psi)}{1 + e^{\alpha}} \right)^{x_3} \cdot \left(\frac{1 - \psi}{1 + e^{\alpha}} \right)^{n-x_1-x_2-x_3} =$$

$$= \binom{n}{x_1+x_2} \binom{x_1+x_2}{x_1} \binom{n-x_1-x_2}{x_3} \psi^{x_1+x_2} (1-\psi)^{n-x_1-x_2} \cdot$$

$$\cdot \left(\frac{e^{\lambda+\alpha}}{1 + e^{\lambda+\alpha}} \right)^{x_1} \cdot \left(\frac{1}{1 + e^{\lambda+\alpha}} \right)^{x_2} \cdot \left(\frac{e^{\alpha}}{1 + e^{\alpha}} \right)^{x_3} \cdot \left(\frac{1}{1 + e^{\alpha}} \right)^{n-x_1-x_2-x_3}$$

$$= B\left(x_1+x_2 \,/\, n \,,\, \psi\right) \cdot B\left(x_1 \,/\, x_1+x_2 \,,\, \frac{e^{\lambda+\alpha}}{1+e^{\lambda+\alpha}}\right) \cdot B\left(x_3 \,/\, n-x_1-x_2 \,,\, \frac{e^{\alpha}}{1+e^{\alpha}}\right)$$

$5^{\underline{a}}$ TRANSFORMAÇÃO

Essa transformação fatora a multinomial em produto de binomiais e nosso parâmetro de interesse θ está explicitado, mas não conseguimos a fatoração em que cada estatística depende apenas de um parâmetro de interesse como queremos. Esta transformação relacionando apenas parâmetros internos nos dá idéia da estrutura da tabela 2x2 com a qual estamos trabalhando.

Seja a transformação:

$$\theta = \frac{p_1\ p_4}{p_2\ p_3} \quad , \quad \beta = \frac{p_2}{p_4} \quad , \quad \gamma = \frac{p_3}{p_4}$$

é 1-1 e sobre , (ver Apêndice A.5). O novo espaço paramétrico é:

$$\Omega = \left\{ (\theta,\beta,\gamma) \text{ tal que } \theta > 0 \ , \ \beta > 0 \ , \ \gamma > 0 \right\}$$

Assim temos:

$$p_2 = \beta\ p_4 \qquad (1)$$

$$p_3 = \gamma\ p_4 \qquad (2)$$

$$p_1 = \frac{\theta\ p_2\ p_3}{p_4} \qquad (3)$$

Substituindo (1) e (2) em (3), temos:

$$p_1 = \frac{\theta\ \beta\ p_4\ \gamma\ p_4}{p_4} \quad \rightarrow \quad p_1 = \theta\ \beta\ \gamma\ p_4 \qquad (4)$$

Mas: $p_4 = 1 - p_1 - p_2 - p_3$ (5)

Então substituindo (4), (1) e (2) em (5), temos:

$$p_4 = 1 - (\theta\ \beta\ \gamma + \beta + \gamma)p_4 \quad \rightarrow$$

$$p_4(1 + \theta\ \beta\ \gamma + \beta + \gamma) = 1$$

$$\rightarrow \quad p_4 = \frac{1}{1 + \beta + \gamma + \theta\ \beta\ \gamma} \qquad (6)$$

Substituindo (6) em (4), (1) e (2) respectivamente, temos:

$$p_1 = \frac{\theta\ \beta\ \gamma}{1 + \beta + \gamma + \theta\ \beta\ \gamma}$$

$$p_2 = \frac{\beta}{1 + \beta + \gamma + \theta\ \beta\ \gamma}$$

$$p_3 = \frac{\gamma}{1 + \beta + \gamma + \theta\ \beta\ \gamma}$$

Essa transformação é 1 a 1 e podemos escrever a multinomial como:

$$P(x_1, x_2, x_3 \;/\; n, p_1, p_2, p_3) = \frac{n!}{x_1! x_2! x_3! (n-x_1-x_2-x_3)!} \, p_1^{x_1} \, p_2^{x_2} \, p_3^{x_3} \cdot$$

$$\cdot \, (1-p_1-p_2-p_3)^{n-x_1-x_2-x_3} \qquad =$$

$$= \frac{n!}{x_1! x_2! x_3! (n-x_1-x_2-x_3)!} \left(\frac{\theta\beta\gamma}{1+\beta+\gamma+\theta\beta\gamma} \right)^{x_1} \left(\frac{\beta}{1+\beta+\gamma+\theta\beta\gamma} \right)^{x_2} \cdot$$

$$\cdot \left(\frac{\gamma}{1+\beta+\gamma+\theta\beta\gamma} \right)^{x_3} \left(\frac{1}{1+\beta+\gamma+\theta\beta\gamma} \right)^{n-x_1-x_2-x_3} \qquad =$$

$$= \frac{n!\,(x_1+x_3)!\,(n-x_1-x_3)!}{(x_1+x_3)!\,(n-x_1-x_3)!\,x_1!\,x_3!\,x_2!\,(n-x_1-x_2-x_3)!} \cdot \left(\frac{\theta\beta\gamma + \gamma}{1+\beta+\gamma+\theta\beta\gamma} \right)^{x_1+x_3} \cdot$$

$$\cdot \left(\frac{1+\beta}{1+\beta+\gamma+\theta\beta\gamma} \right)^{n-x_1-x_3} \cdot \left(\frac{\beta}{1+\beta} \right)^{x_2} \cdot \left(\frac{1}{1+\beta} \right)^{n-x_1-x_2-x_3} \cdot$$

$$\cdot \left(\frac{\theta\beta}{1+\theta\beta} \right)^{x_1} \cdot \left(\frac{1}{1+\theta\beta} \right)^{x_3} \qquad =$$

$$= \binom{n}{x_1+x_3} \binom{x_1+x_3}{x_1} \binom{n-x_1-x_3}{x_2} \cdot \left(\frac{\theta\beta\gamma + \gamma}{1+\beta+\gamma+\theta\beta\gamma} \right)^{x_1+x_3} \cdot$$

$$\cdot \left(\frac{1 + \beta}{1+\beta+\gamma+\theta\beta\gamma} \right)^{n-x_1-x_3} \cdot \left(\frac{\beta}{1+\beta} \right)^{x_2} \cdot \left(\frac{1}{1+\beta} \right)^{n-x_1-x_2-x_3} \cdot$$

$$\cdot \left(\frac{\theta\beta}{1+\theta\beta} \right)^{x_1} \cdot \left(\frac{1}{1+\theta\beta} \right)^{x_3} =$$

$$= B\left(x_1+x_3 \;/\; n, \frac{\theta\beta\gamma + \gamma}{1+\beta+\gamma+\theta\beta\gamma} \right) \cdot B\left(x_2 \;/ n- x_1-x_3, \frac{\beta}{1+\beta} \right) \cdot$$

$$\cdot B\left(x_1 \;/\; x_1+x_3, \frac{\theta\beta}{1+\theta\beta} \right)$$

$6^{\underline{a}}$ TRANSFORMAÇÃO

Essa transformação fatora a multinomial em produto de binoniais mas não com cada estatística dependendo apenas de um parâmetro de interesse, mas ela é importante pois com esta transformação a matriz de infomação de Fisher é uma matriz diagonal. (ver Apêndice A.7).

Seja a transformação:

$$\omega_1 = \frac{p_1}{p_4} = \frac{p_1}{1-p_1-p_2-p_3}$$

$$\omega_2 = \frac{p_2}{p_1+p_4} = \frac{p_2}{1-p_2-p_3}$$

$$\omega_3 = \frac{p_3}{p_1+p_2+p_4} = \frac{p_3}{1-p_3}$$

é 1-1 e sobre. $\Omega = (0,1)^3$. (ver Apêndice A.6).

Temos:

$$p_1 = \omega_1 p_4 \qquad (1)$$

$$p_2 = \omega_2(p_1+p_4) \qquad (2)$$

Substituindo (2) em (1) temos:

$$p_2 = \omega_2(\omega_1 p_4+p_4)$$

$$p_2 = p_4\omega_2(\omega_1+1) \qquad (3)$$

$$p_3 = \omega_3(p_1+p_2+p_4) \qquad (4)$$

Substituindo (1) e (3) em (4) temos:

$$p_3 = \omega_3\left[p_4\omega_1 + p_4\omega_2(\omega_1+1) + p_4\right]$$

$$p_3 = \omega_3 p_4\left[\omega_1 + \omega_2(\omega_1+1) + 1\right]$$

$$p_3 = \omega_3 p_4(\omega_1+1)(\omega_2+1) \qquad (5)$$

Mas, $p_4 = 1-p_1-p_2-p_3$ (6)

Substituindo (1), (3) e (5) em (6) temos:

$$p_4 = 1 - p_4\left[\ \omega_1 + \omega_2(\omega_1+1) + \omega_3(\omega_2+1)(\omega_1+1)\ \right]$$

$$p_4\left[(1+\omega_1) + \omega_2(\omega_1+1) + \omega_3(\omega_2+1)(\omega_1+1)\ \right] = 1$$

$$p_4 = \frac{1}{(\omega_1+1)(\omega_2+1)(\omega_3+1)} \qquad (7)$$

Substituindo (7) em (1), (3) e (5) respectivamente, temos:

$$p_1 = \frac{\omega_1}{(\omega_1+1)(\omega_2+1)(\omega_3+1)}$$

$$p_2 = \frac{\omega_2(\omega_1+1)}{(\omega_1+1)(\omega_2+1)(\omega_3+1)} = \frac{\omega_2}{(\omega_2+1)(\omega_3+1)}$$

$$p_3 = \frac{\omega_3(\omega_2+1)(\omega_1+1)}{(\omega_1+1)(\omega_2+1)(\omega_3+1)} = \frac{\omega_3}{(\omega_3+1)}$$

Podemos escrever a multinomial como:

$$P(x_1, x_2, x_3 \,/\, n, p_1, p_2, p_3) = \frac{n!}{x_1!x_2!x_3!(n-x_1-x_2-x_3)!}\, p_1^{x_1}\, p_2^{x_2}\, p_3^{x_3} \cdot$$

$$\cdot\, (1-p_1-p_2-p_3)^{n-x_1-x_2-x_3} \qquad =$$

$$= \frac{n!(x_1+x_3)!(n-x_1-x_3)!}{(x_1+x_3)!(n-x_1-x_3)!x_1!x_3!x_2!(n-x_1-x_2-x_3)!} \left[\frac{\omega_1}{(\omega_1+1)(\omega_2+1)(\omega_3+1)}\right]^{x_1} \cdot$$

$$\cdot \left[\frac{\omega_2(\omega_1+1)}{(\omega_1+1)(\omega_2+1)(\omega_3+1)}\right]^{x_2} \cdot \left[\frac{\omega_3(\omega_2+1)(\omega_1+1)}{(\omega_1+1)(\omega_2+1)(\omega_3+1)}\right]^{x_3} \cdot$$

$$\cdot \left[\frac{1}{(\omega_1+1)(\omega_2+1)(\omega_3+1)}\right]^{n-x_1-x_2-x_3} \qquad =$$

$$= \binom{n}{x_1+x_3}\binom{x_1+x_3}{x_1}\binom{n-x_1-x_3}{x_2} \cdot \left[\frac{\omega_1+\omega_3(\omega_2+1)(\omega_1+1)}{(\omega_1+1)(\omega_2+1)(\omega_3+1)}\right]^{x_1+x_3} \cdot$$

$$\cdot \left[\frac{1+\omega_2(\omega_1+1)}{(\omega_1+1)(\omega_2+1)(\omega_3+1)}\right]^{n-x_1-x_3} \cdot \left[\frac{\omega_2(\omega_1+1)}{1+\omega_2(\omega_1+1)}\right]^{x_2} \cdot$$

$$\cdot \left[\frac{1}{1+\omega_2(\omega_1+1)}\right]^{n-x_1-x_2-x_3} \cdot \left[\frac{\omega_1}{\omega_1+\omega_3(\omega_2+1)(\omega_1+1)}\right]^{x_1} \cdot$$

$$\cdot \left[\frac{\omega_3(\omega_2+1)(\omega_1+1)}{\omega_1+\omega_3(\omega_2+1)(\omega_1+1)}\right]^{x_3} =$$

$$= B\left(x_1+x_3 \;/\; n,\; \frac{\omega_1+\omega_3(\omega_2+1)(\omega_1+1)}{(\omega_1+1)(\omega_2+1)(\omega_3+1)}\right) . B\left(x_2 \;/\; n-x_1-x_3,\; \frac{\omega_2(\omega_1+1)}{1+\omega(\omega+1)}\right) .$$

$$. B\left(x_1 \;/\; x_1+x_3\;,\; \frac{\omega_1}{\omega_1+\omega_3(\omega_2+1)(\omega_1+1)}\right)$$

Fizemos um total de seis transformações mas não encontramos a transformação ideal que procuramos, ou seja aquela que fatore a multinomial em partes de modo que uma estatística dependa apenas do parâmetro de interesse θ e as outras sejam independentes do mesmo.

Continuaremos a nossa pesquisa e esperamos que este trabalho sirva para chamar a atenção para o problema e que mais pessoas se interessem em pesquisar nesta direção.

CAPÍTULO 3. GEOMETRIA DE UMA TABELA DE CONTINGÊNCIA 2 X 2

Qualquer tabela de contingência pode ser normalizada para ter entradas que somam 1. Assim há uma correspondência natural entre o conjunto de tabelas 2x2 cujas entradas somam 1 e os pontos de um simplex tridimensional (tetraedro), Fienberg e Gilbert (1970).

Discutiremos algumas idéias estatísticas sobre tabelas 2x2 em termos da geometria do tetraedro. Em particular, para:

(a) todos os pontos correspondendo a tabelas cujas linhas e colunas são independentes,

(b) todos os pontos correspondendo a tabelas com um dado grau de associação, e

(c) todos os pontos correspondentes a tabelas com um conjunto fixado de totais marginais.

Usaremos a estrutura do simplex tridimensional, e faremos a correspondência um a um entre o conjunto de tabelas 2x2 normalizadas e os pontos de um tetraedro através das coordenadas baricêntricas normalizadas. Escolhemos o tetraedro de referência, onde $A_1 = (1,0,0,0)$, $A_2 = (0,1,0,0)$, $A_3 = (0,0,1,0)$ e $A_4 = (0,0,0,1)$ correspondem respectivamente as tabelas 2x2 abaixo:

1	0
0	0

0	1
0	0

0	0
1	0

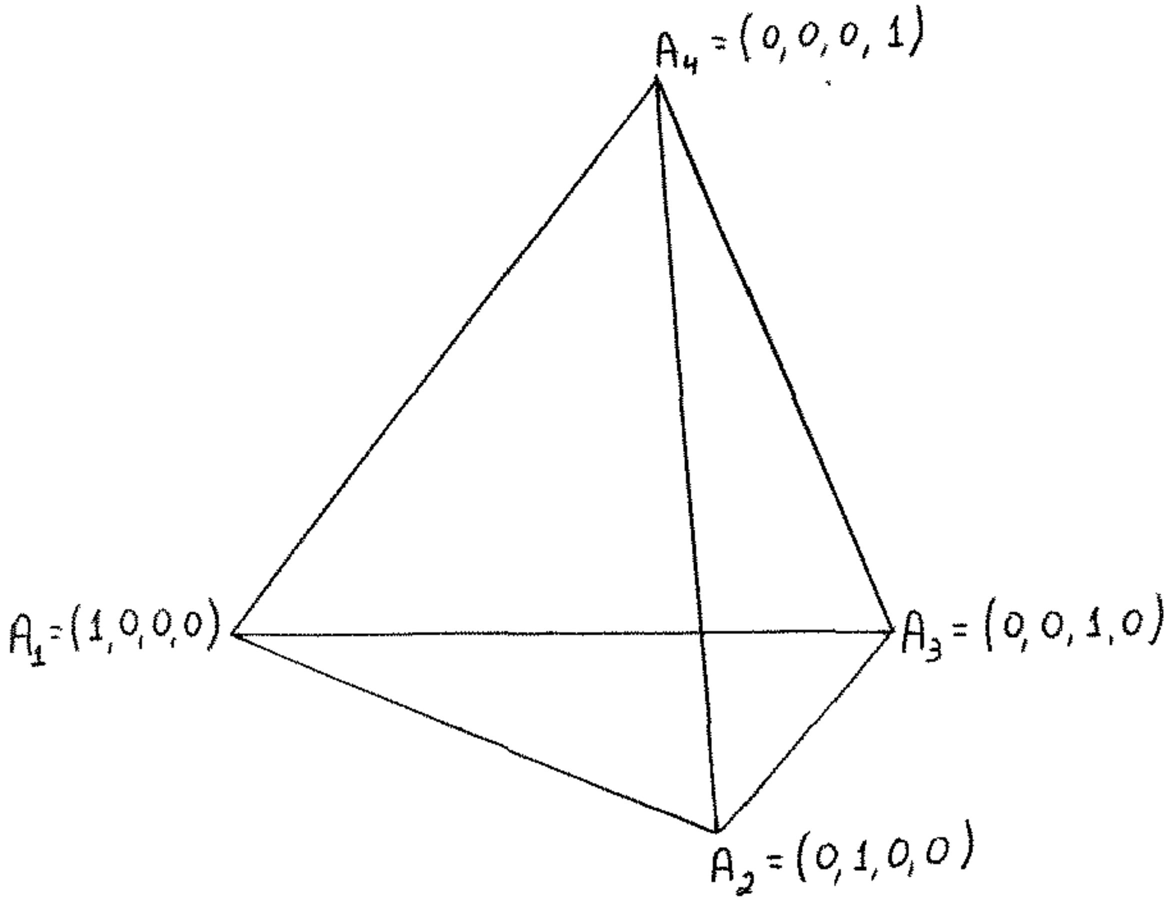
0	0
0	1

O ponto P = (p_1,p_2,p_3,p_4) corresponde a tabela 2x2:

p_1	p_2
p_3	p_4

Há uma correspondência 1a1 entre os pontos no tetraedro e as tabelas 2x2.

TETRAEDRO DE REFERÊNCIA



SUPERFÍCIE DE INDEPENDÊNCIA

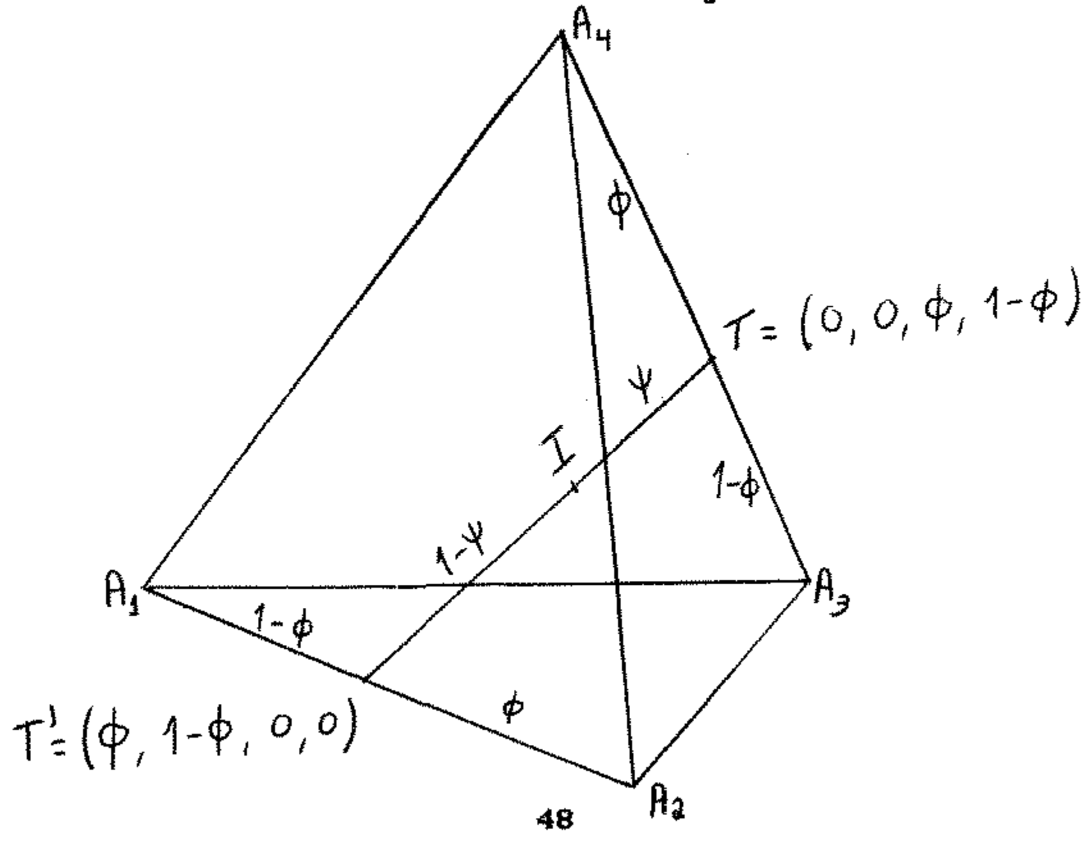
Qualquer ponto T' sobre a linha A_1A_2 é determinado por um número ϕ tal que $0 \le \phi \le 1$ e

$$\frac{\phi}{1-\phi} = \frac{\overline{A_2T'}}{\overline{T'A_1}}$$

onde $\overline{A_2T'}$ e $\overline{T'A_1}$ representam as distâncias entre A_2 e T', e T' e A_1, respectivamente. Assim T' tem coordenadas $(\phi,1-\phi,0,0)$.

Podemos escolher também um outro ponto T sobre A_3A_4 tal que suas coordenadas sejam $(0,0,\phi,1-\phi)$ e

$$\frac{\phi}{1-\phi} = \frac{\overline{A_4T}}{\overline{T\ A_3}}$$



Agora por definição, qualquer ponto I sobre a linha T'T dentro do tetraedro corresponde um número ψ, talque $0 \le \psi \le 1$ e

$$\frac{\psi}{1-\psi} = \frac{\overline{T\ I}}{\overline{I\ T'}}$$

Assim I tem coordenadas $\left(\psi\phi, \psi(1-\phi), (1-\psi)\phi, (1-\psi)(1-\phi)\right)$ e corresponde a tabela abaixo:

$\psi\ \phi$	$\psi(1-\phi)$	ψ
$(1-\psi)\phi$	$(1-\psi)(1-\phi)$	$1-\psi$
ϕ	$(1-\phi)$	

cujas linhas e colunas são independentes.

Permitindo que ϕ e ψ assumam todos os possíveis valores entre 0 e 1, podemos encontrar todos os pontos correspondentes a tabelas cujas linhas e colunas sejam independentes.

As linhas T'T definidas para diferentes valores de ϕ $(0 \le \phi \le 1)$ estão na superfície de independência.

Alternativamente, poderíamos ter escolhido os pontos S' sobre A_1A_3 , S sobre A_2A_4 e I' sobre S'S tais que:

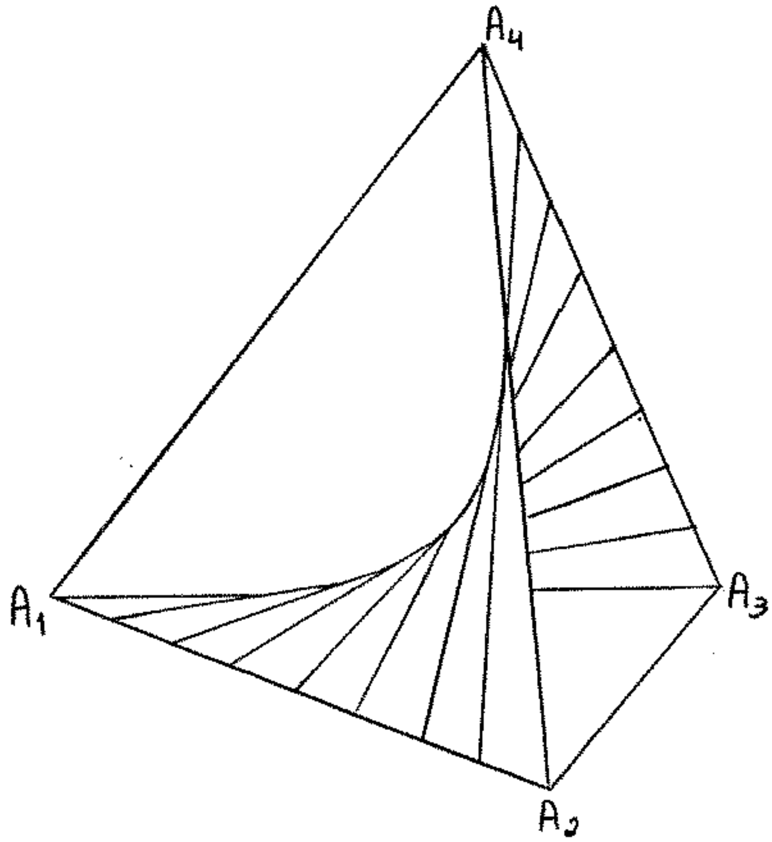
$$\frac{\overline{A_3S'}}{\overline{S'A_1}} = \frac{\psi}{1-\psi} \quad , \quad \frac{\overline{A_4S}}{\overline{S\ A_2}} = \frac{\psi}{1-\psi} \quad \text{e} \quad \frac{\overline{S\ I'}}{\overline{I'S'}} = \frac{\phi}{1-\phi}$$

Assim I' tem as coordenadas :

$$\left[\phi\ \psi\ ,\ \psi(1-\phi)\ ,\ (1-\psi)\phi\ ,\ (1-\phi)(1-\psi)\right]$$

Observamos que: $I = I'$.

SUPERFÍCIE DE INDEPENDÊNCIA



Assim as linhas S'S definidas para diferentes valores de ψ $(0\leq\psi\leq1)$ também estão sobre a superfície de independência. Esta superfície está completamente determinada por qualquer uma das famílias de linhas (T'T ou S'S).

De fato a superfície de independência é um parabolóide hiperbólico e seu ponto de sela é o centro do tetraedro, $C = \left(\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4}\right)$, Fienberg e Gilbert (1970).

O parabolóide hiperbólico é uma superfície duplamente regrada visto que sua superfície contém duas famílias de linhas retas.

As tabelas correspondendo aos pontos sobre qualquer uma das linhas T'T têm os mesmos totais marginais colunas, enquanto as tabelas correspondendo aos pontos sobre qualquer uma das linhas S'S têm os mesmos totais marginais linhas.

SUPERFÍCIE DE ASSOCIAÇÃO CONSTANTE

Para a tabela 2x2, abaixo:

p_1	p_2
p_3	p_4

onde $p_i \geq 0$ e $\sum_{i=1}^{4} p_i = 1$.

Podemos definir uma medida do grau de associação pelo coeficiente θ (odds ratio):

$$\theta = \frac{p_1 \; p_4}{p_2 \; p_3} \qquad , \quad 0 \leq \theta \leq oo$$

Note que os totais marginais linha e coluna de uma tabela são independentes se e somente se $\theta = 1$.

Uma associação "positiva" ($\theta > 1$) por exemplo $\theta = 4$, pode ser considerada como igual , mas com afastamento oposto da independência quando comparado com uma associação "negativa" ($\theta < 1$) por exemplo $\theta = \frac{1}{4}$.

Definimos uma superfície de associação constante θ, do mesmo modo que definimos a superfície de independência, escolhendo T e T^* tais que:

$$\frac{\overline{A_2 T^*}}{\overline{T^* A_1}} = \theta \cdot \frac{\overline{A_4 T}}{\overline{T A_3}}$$

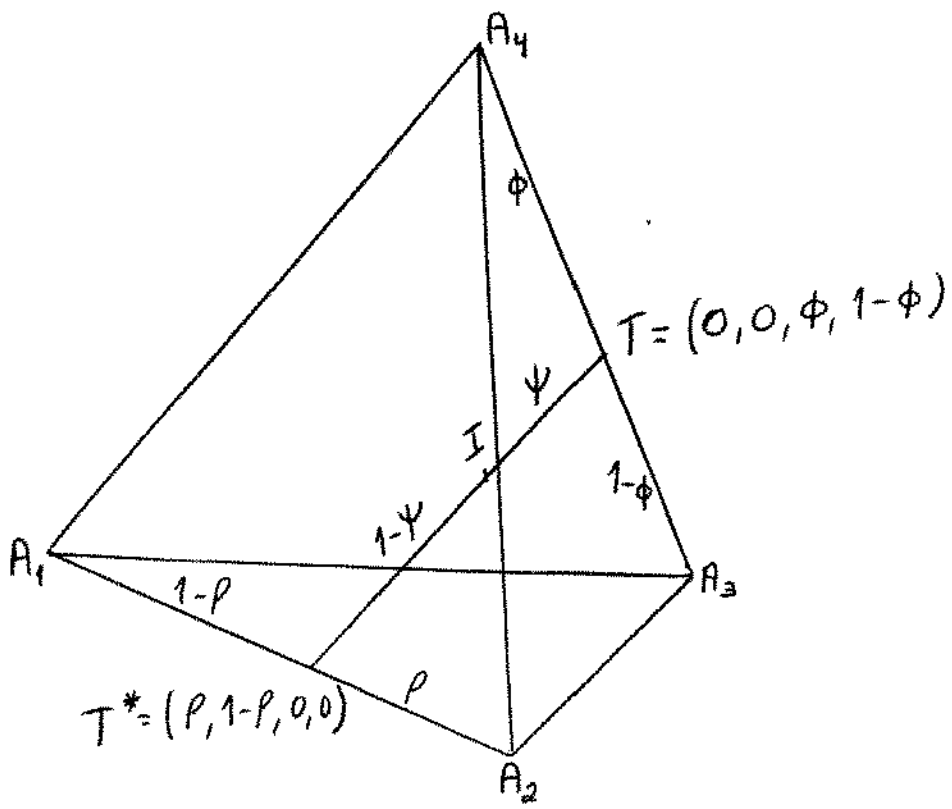
ou seja:

$$\frac{\rho}{1-\rho} = \theta \cdot \frac{\phi}{1-\phi}$$

então temos:

$$\theta = \frac{\rho \,/\, (1-\rho)}{\phi \,/\, (1-\phi)}$$

onde: $0 \leq \rho \leq 1$, $0 \leq \phi \leq 1$ e $0 \leq \theta \leq \infty$.



Assim I tem coordenadas $\left[\psi\ \rho, \psi(1-\rho), (1-\psi)\phi, (1-\psi)(1-\phi)\right]$.

Mas sabemos que:

$$\theta = \frac{\rho \,/\, (1-\rho)}{\phi \,/\, (1-\phi)} \rightarrow \rho = \frac{\theta\phi}{1 - \phi + \theta\phi}$$

Então o ponto I tem as seguintes coordenadas em termos de θ, ψ e ϕ :

$$I = \left[\frac{\theta\ \psi\ \phi}{1 - \phi + \theta\phi} , \frac{\psi\ (1-\phi)}{1 - \phi + \theta\phi} , (1-\psi)\phi , (1-\psi)(1-\phi) \right]$$

e corresponde a tabela 2x2 abaixo:

$\frac{\theta\ \psi\ \phi}{1- \phi + \theta\phi}$	$\frac{\psi\ (1-\phi)}{1 - \phi + \theta\phi}$
$(1-\psi)\phi$	$(1-\psi)(1-\phi)$

Alternativamente, poderíamos ter escolhido os pontos S e S^* tais que:

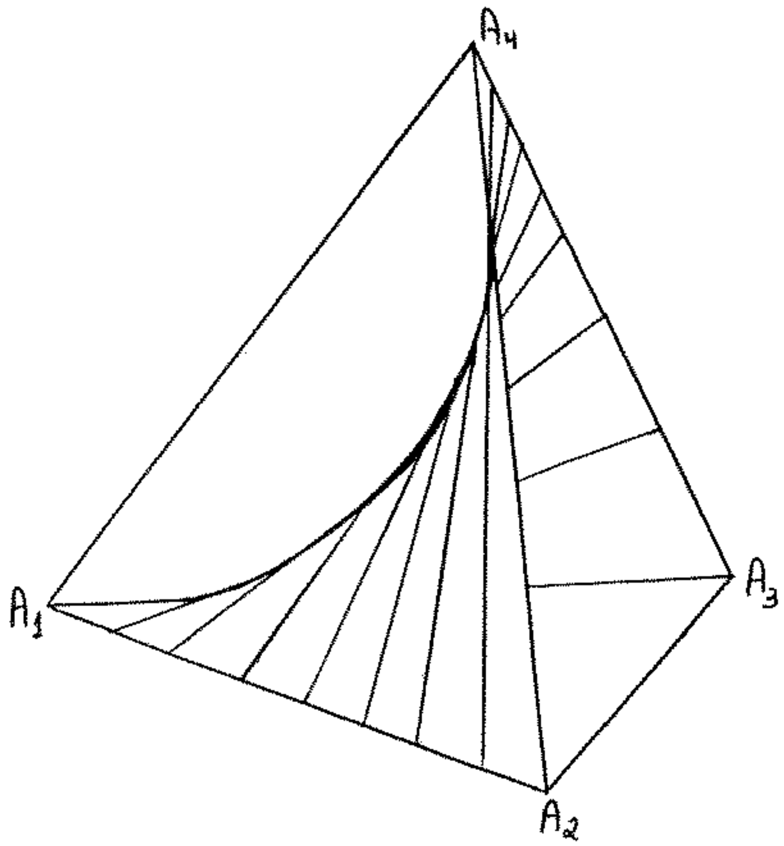
$$\frac{\overline{A_3 S^*}}{\overline{S^* A_1}} = \theta \cdot \frac{\overline{A_4 S}}{\overline{S A_2}}$$

ou seja:

$$\theta = \frac{\gamma \ / \ (1-\gamma)}{\psi \ / \ (1-\psi)} \quad \text{onde} \quad \begin{cases} 0 \leq \gamma \leq 1 \\ 0 \leq \psi \leq 1 \\ 0 \leq \theta \leq \infty \end{cases}$$

Assim as linhas TT^* definidas para diferentes valores de ϕ e ρ $(0 \leq \phi,\rho \leq 1)$ estão sobre a superfície de constante θ.

SUPERFÍCIE DE CONSTANTE θ ($\theta = 3$)



A superfície de associação constante θ está completamente determinada por uma dessas famílias de linhas (TT^* ou SS^*) e estas superfícies interceptam a superfície de independência ao longo das 4 arestas de tetraedro: A_1A_2 , A_1A_3 , A_2A_4 , A_3A_4.

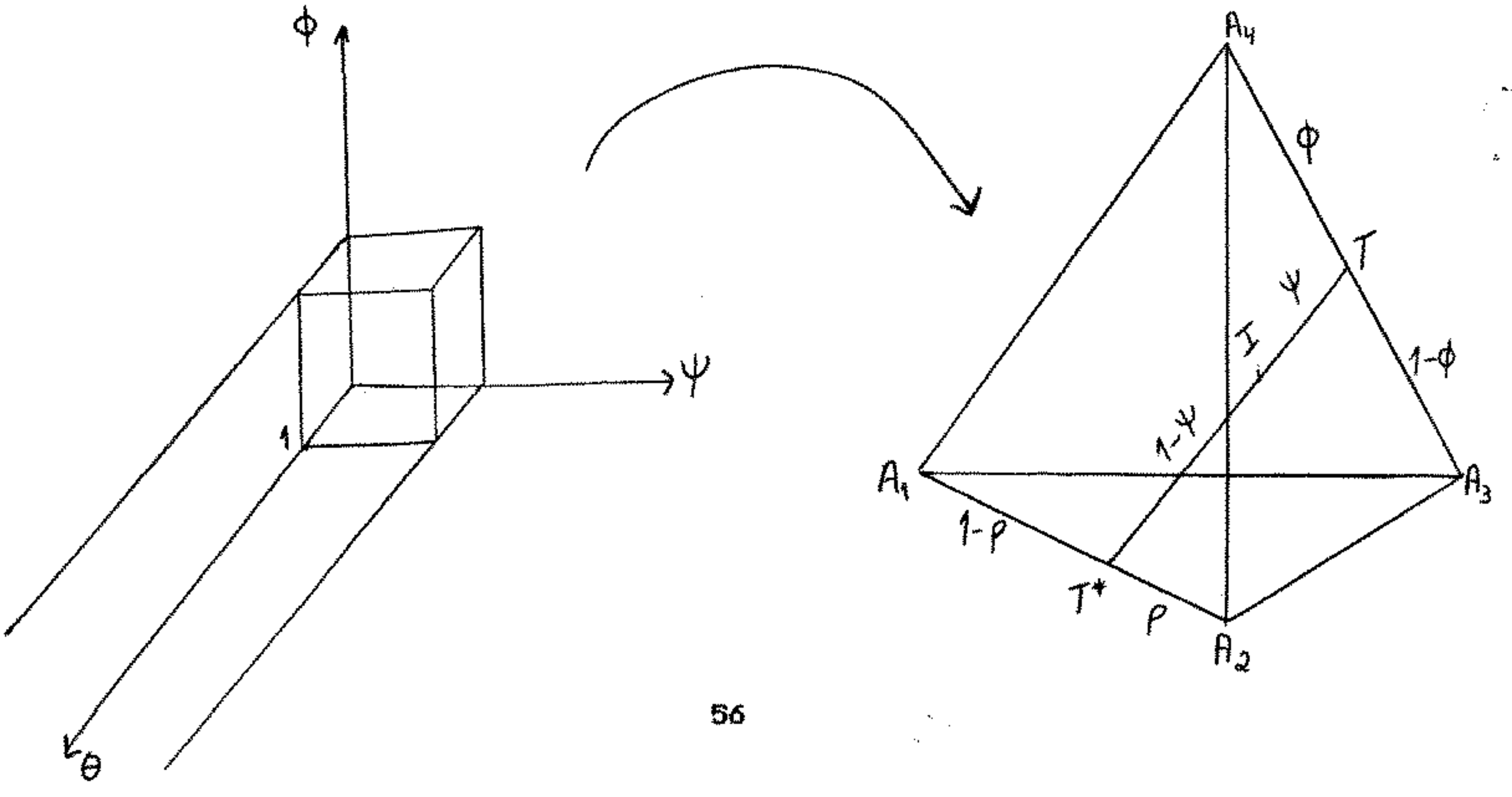
De fato, as superfícies de constante θ são hiperbolóides de uma folha, os quais são também superfícies duplamente regradas. Note que o plano, o parabolóide hiperbólico, e o hiperbolóide de uma folha são as únicas superfícies duplamente regradas, Fienberg e Gilbert (1970).

Uma tabela 2x2 exibe perfeita associação "positiva" quando todas as probabilidades estão nas células (1,1) e (2,2), e $\theta = \infty$. O local de tais tabelas é o segmento de linha A_1A_4, o qual é o limite das superfícies de associação constante θ quando θ tende ao ∞.

Semelhantemente, uma tabela com perfeita associação "negativa" tem todas as suas probabilidades nas células (1,2) e (2,1), e $\theta = 0$. O local de tais tabelas é o segmento de linha A_2A_3, o qual é o limite das superfícies de constante θ, quando θ tende a zero.

Finalmente, as linhas TT^* não são o local de pontos correspondendo a tabelas com marginais colunas constantes, como eram as linhas TT'. Semelhantemente, as linhas SS^*, não correspondem a tabelas com totais marginais linhas constantes.

Estamos saindo do cubo "alongado" para o tetraedro como mostram as figuras abaixo:



TABELAS COM MARGINAIS FIXADAS

Tabelas com ambos totais marginais linha e coluna fixados são de interesse estatístico por causa do seu uso no teste exato de Fisher. Portanto, encontraremos o local correspondente para tais tabelas.

Considere dois pontos quaisquer no tetraedro,

$$P_1 = (a_1 \, , \psi - a_1 \, , \phi - a_1 \, , 1- \psi - \phi + a_1)$$

e

$$P_2 = (a_2 \, , \psi - a_2 \, , \phi - a_2 \, , 1- \psi - \phi + a_2)$$

os quais correspondem, respectivamente, as tabelas com totais marginais linhas $(\psi,1-\psi)$ e totais marginais colunas $(\phi,1-\phi)$ conforme abaixo:

a_1	$\psi - a_1$	ψ
$\phi - a_1$	$1-\psi-\phi+a_1$	$1-\psi$
ϕ	$1-\phi$	

e

a_2	$\psi - a_2$	ψ
$\phi - a_2$	$1-\psi-\phi+a_2$	$1-\psi$
ϕ	$1-\phi$	

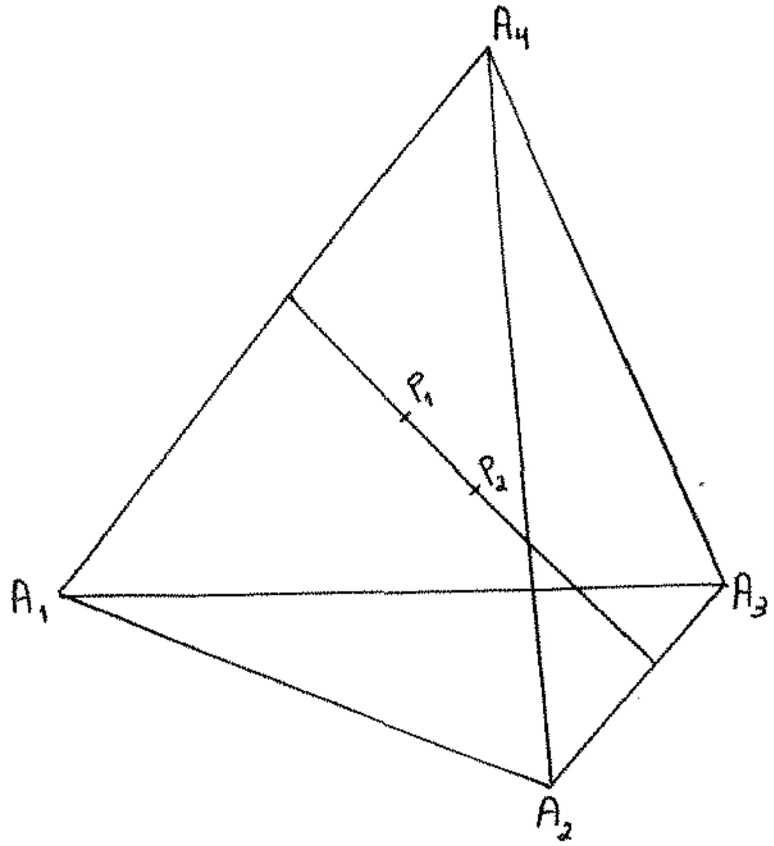
Podemos escrever P_1 e P_2 , da seguinte maneira:

$$P_1 = (0,\psi,\phi,1-\psi-\phi) + a_1(1,-1,-1,1)$$

e

$$P_2 = (0,\psi,\phi,1-\psi-\phi) + a_2(1,-1,-1,1)$$

A direção da linha P_1P_2 é dada pelo vetor (1,-1,-1,1), o qual é independente de a_1 e a_2. Assim todos os pontos (P) com totais marginais $(\psi,1-\psi)$ e $(\phi,1-\phi)$ estão sobre a linha P_1P_2. Mais ainda, é claro que a linha P_1P_2 é ortogonal as linhas A_1A_4 e A_2A_3, as quais têm direções dadas vetores (1,0,0,-1) e (0,1,-1,0) respectivamente.



Assim todas as tabelas com o mesmo conjunto de marginais (fixadas) correspondem a pontos sobre uma linha reta ortogonal a A_1A_4 e A_2A_3. Note que esta linha de marginais fixadas que tem direção (1.-1,-1,1) é paralela a linha que une os pontos médios de A_1A_4 e A_2A_3 e portanto, a linha através de P_1P_2 não é perpendicular a superfície de independência, a menos que as marginais sejam todas iguais, Fienberg e Gilbert (1970).

Todas as tabelas com o mesmo conjunto de marginais fixadas correspondem a pontos sobre uma linha paralela a essa linha que une os pontos médios de A_1A_4 e A_2A_3.

Assim temos que um valor especifico de θ e um par de marginais determinam completamente uma tabela 2x2. A geometria equivalente a esta afirmação é a seguinte: primeiro vamos para o ponto sobre a superfície de independência cuja tabela correspondente tem as marginais fixadas. Então nos deslocamos ao longo da linha reta, que passa por esse ponto, ortogonal a A_1A_4 e A_2A_3 até cruzarmos a superfície correspondente a θ. Este ponto de interseção corresponde a tabela 2x2 determinada por θ e pelo par de marginais fixadas.

CAPITULO 4. EXEMPLOS

No capítulo 2, a 1ª transformação que apresentamos foi a proposta por Fisher e é para esta transformação que daremos dois exemplos.

O objetivo de apresentar esses exemplos é para chamar a atenção de que é realmente importante encontrar uma transformação que fatore a multinomial em partes de maneira que uma estatística dependa apenas do parâmetro de interesse θ e as outras não dependam de θ.

O problema de encontrar uma tal transformação é um problema em aberto e esperamos que com este trabalho mais pessoas se interessem em pesquisar nesta direção.

De acordo com a parametrização proposta por Fisher, a fatoração da multinomial é dada por:

$$P(x_1,x_2,x_3/\theta,\phi,\psi) = P(x_1+x_2/\psi).P(x_1+x_3/x_1+x_2,\theta,\phi).P(x_1/x_1+x_3,x_1+x_2,\theta)$$

A distribuição marginal de x_1+x_3 condicionada a x_1+x_2, envolve o parâmetro θ e portanto deve conter informação sobre ele.

Mostraremos através de dois exemplos que a distribuição de θ, depende da marginal x_1+x_3, assim temos que quando mudamos o valor dessa marginal a distribuição de θ também se altera.

Considere uma população de indivíduos, classificada segundo duas características A e B, gerando a tabela:

	A	A^c
B	p_1	p_2
B^c	p_3	p_4

onde $p_i \geq 0$ e $\sum_{i=1}^{4} p_i = 1$.

Toma-se n_1 indivíduos com característica B e $n-n_1$ indivíduos com característica B^c. Dentre estes, observa-se o número de indivíduos com característica A, gerando a tabela:

	A	A^c	
B	x_1		n_1
B^c	x_3		$n-n_1$

Os dados (x_1,x_3) seguem a distribuição:

$$x_1 \sim B\left(n_1 , \frac{p_1}{p_1+p_2}\right)$$

$$x_3 \sim B\left(n-n_1 , \frac{p_3}{p_3+p_4}\right)$$

independentes.

O modelo binomial pode ser considerado como uma redução do multinomial, no sentido de que a distribuição multinomial pode ser escrita como:

$$P(x_1,x_2,x_2/p_1,p_2,p_3) \quad = \quad \frac{n!\,(x_1+x_2)!\,(n-x_1-x_2)!}{(x_1+x_2)!\,(n-x_1-x_2)!\,x_1!\,x_2!\,x_3!\,(n-x_1-x_2-x_3)!}$$

$$.\,p_1^{x_1}.\,p^{x_2}.\,p_3^{x_3}.\,(1-p_1-p_2-p_3)^{n-x_1-x_2-x_3} \quad =$$

$$= \binom{n}{x_1+x_2}\binom{x_1+x_2}{x_1}\binom{n-x_1-x_2}{x_3}\left(\frac{p_1}{p_1+p_2}\right)^{x_1}\left(\frac{p_2}{p_1+p_2}\right)^{x_2}.\left(\frac{p_3}{1-p_1-p_2}\right)^{x_3}.$$

$$.\left(\frac{1-p_1-p_2-p_3}{1-p_1-p_2}\right)^{n-x_1-x_2-x_3}(p_1+p_2)^{x_1+x_2}(1-p_1-p_2)^{n-x_1-x_2} \quad =$$

$$= \binom{n}{x_1+x_2}\,(p_1+p_2)^{x_1+x_2}(1-p_1-p_2)^{n-x_1-x_2}\,\binom{x_1+x_2}{x_1}\left(\frac{p_1}{p_1+p_2}\right)^{x_1}\left(\frac{p_2}{p_1+p_2}\right)^{x_2}.$$

$$.\,\binom{n-x_1-x_2}{x_3}\left(\frac{p_3}{1-p_1-p_2}\right)^{x_3}\left(\frac{1-p_1-p_2-p_3}{1-p_1-p_2}\right)^{n-x_1-x_2-x_3} \quad =$$

$$= B\left(x_1+x_2 \;/\; n,p_1+p_2\right)\;.\;B\left(x_1 \;/\; x_1+x_2,\frac{p_1}{p_1+p_2}\right)\;.\;B\left(x_3 \;/\; n-x_1-x_2,\frac{p_3}{1-p_1-p_2}\right)$$

EXEMPLO 1. Consideremos duas distribuições binomiais com tamanhos iguais a 5 e probabilidades iguais a $\frac{1}{2}$.

Temos a seguinte tabela:

x_1	x_2	5
x_3	x_4	5
m_1	m_2	10

Os possíveis valores para $\begin{cases} x_1 = 0, 1, 2, 3, 4, 5 \\ x_3 = 0, 1, 2, 3, 4, 5 \end{cases}$.

Considerando as marginais m_1 e m_2 fixadas, todos os possíveis valores são, respectivamente:

$$\begin{cases} m_1 = 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, 0 \\ m_2 = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 \end{cases}$$

Podemos gerar 36 tabelas 2x2, classificando-as pelos valores marginais m_1 e m_2 .

A probabilidade de cada tabela será:

$$\binom{5}{x_1}\left(\frac{1}{2}\right)^5 \cdot \binom{5}{x_3}\left(\frac{1}{2}\right)^5 = \frac{\binom{5}{x_1}\binom{5}{x_3}}{1024}$$

A probabilidade de cada tabela para valores dados das marginais m_1 e m_2 $(x_1+ x_2= m_1)$ segue a distribuição hipergeométrica dada por:

$$\frac{\binom{5}{x_1}\binom{5}{x_2}}{\binom{10}{m_1}}$$

As frequências das 36 tabelas, classificadas pelos valores das marginais m_1 e m_2, estão relacionadas no Apêndice B.1 e abaixo reproduzimos as tabelas para as marginais $m_1= 7$ e $m_2= 3$, Yates (1984):

(5,0,2,3)

5	0	5
2	3	5
7	3	

$$\binom{5}{5}\cdot\binom{5}{2} = 10 \qquad p_1-p_2= 0{,}6 \qquad \theta = \infty$$

(4,1,3,2)

4	1	5
3	2	5
7	3	

$$\binom{5}{4}\cdot\binom{5}{3} = 50 \qquad p_1-p_2= 0{,}2 \qquad \theta = \frac{8}{3}$$

(3,2,4,1)

3	2	5
4	1	5
7	3	

$$\binom{5}{3}\cdot\binom{5}{4} = 50 \qquad p_1-p_2=-0{,}2 \qquad \theta = \frac{3}{8}$$

(2,3,5,0)

2	3	5
5	0	5
7	3	

$\binom{5}{2} \cdot \binom{5}{5} = 10$

$p_1 - p_2 = -0{,}6$

$\theta = 0$

TABELA 1. Frequencias relativas das 36 tabelas 2x2 geradas por amostras de duas distribuições binomiais com: $n_1 = n_2 = 5$ e $p = 0,5$, classificadas pelos valores das marginais m_1 e m_2 .

	marginais m_1 , m_2											
$p_1 - p_2$	10,0	9,1	8,2	7,3	6,4	5,5	4,6	3,7	2,8	1,9	0,10	Total
-1.0						$\theta = 0$ 1 (0.004)						1
-0.8					$\theta = 0$ 5 (0.024)		$\theta = 0$ 5 (0.024)					10
-0.6				$\theta = 0$ 10 (0.083)		$\theta = 1/16$ 25 (0.099)		$\theta = 0$ 10 (0.083)				45
-0.4			$\theta = 0$ 10 (0.222)		$\theta = 1/6$ 50 (0.238)		$\theta = 1/6$ 50 (0.238)		$\theta = 0$ 10 (0.222)			120
-0.2		$\theta = 0$ 5 (0.5)		$\theta = 3/8$ 50 (0.417)		$\theta = 4/9$ 100 (0.397)		$\theta = 3/8$ 50 (0.417)		$\theta = 0$ 5 (0.5)		210
0.0	$\theta = 0/0$ 1 (1.0)		$\theta = 1$ 25 (0.556)		$\theta = 1$ 100 (0.476)		$\theta = 1$ 100 (0.476)		$\theta = 1$ 25 (0.556)		$\theta = 0/0$ 1 (1.0)	252
+0.2		$\theta = \infty$ 5 (0.5)		$\theta = 8/3$ 50 (0.417)		$\theta = 9/4$ 100 (0.397)		$\theta = 8/3$ 50 (0.417)		$\theta = \infty$ 5 (0.5)		210
+0.4			$\theta = \infty$ 10 (0.222)		$\theta = 6$ 50 (0.238)		$\theta = 6$ 50 (0.238)		$\theta = \infty$ 10 (0.222)			120
+0.6				$\theta = \infty$ 10 (0.083)		$\theta = 16$ 25 (0.099)		$\theta = \infty$ 10 (0.083)				45
+0.8					$\theta = \infty$ 5 (0.024)		$\theta = \infty$ 5 (0.024)					10
+1.0						$\theta = \infty$ 1 (0.004)						1
Total	1	10	45	120	210	252	210	120	45	10	1	1024

A primeira coluna contém a tabela (5,0,5,0), a segunda as tabelas (4,1,5,0), (5,0,4,1), etc. Os números em parênteses são os elementos da distribuição hipergeométrica para valores dados das marginais m_1 e m_2 .

Observamos pela tabela 1, anterior, que a distribuição de θ depende das marginais m_1 e m_2 , ou seja quando mudamos as valores das marginais m_1 e m_2 a distribuição de θ se altera.

Assim a transformação proposta por Fisher não é a que procuramos pois a distribuição da marginal x_1+x_3 condicionada a x_1+x_2 envolve o papâmetro θ e assim a distribuição de θ se altera quando mudamos as marginais.

Distribuição exata de probabilidade do odds ratio (θ) quando se consideram duas binomiais com tamanhos iguais a 5 e probabilidades iguais a $\frac{1}{2}$.

θ	f	P(θ)
0/0	2	0.002
0	61	0.060
1/16	25	0.024
1/6	100	0.098
3/8	100	0.098
4/9	100	0.098
1	250	0.242
9/4	100	0.098
8/3	100	0.098
6	100	0.098
16	25	0.024
∞	61	0.060
Σ	1024	1.000

Para as tabelas com $\theta = 0$, $0/0$ e ∞ respectivamente, faremos a seguinte modificação proposta por Haldane e Anscombe (Gart e Zweifel , 1967).

Chamaremos θ^* o valor de θ modificado segundo Haldame e Anscombe, o qual é dado por:

$$\begin{array}{|c|c|} \hline a & b \\ \hline c & d \\ \hline \end{array} \quad \rightarrow \quad \begin{array}{|c|c|} \hline a + \frac{1}{2} & b + \frac{1}{2} \\ \hline c + \frac{1}{2} & d + \frac{1}{2} \\ \hline \end{array}$$

$$\theta = \frac{a\ d}{b\ c} \quad \rightarrow \quad \theta^* = \frac{\left(a + \frac{1}{2}\right)\left(d + \frac{1}{2}\right)}{\left(b + \frac{1}{2}\right)\left(c + \frac{1}{2}\right)}$$

Ou seja:

$$\theta^* = \begin{cases} \dfrac{a\ d}{b\ c} & ;\ a,b,c,d > 0 \\[2ex] \dfrac{\left(a + \frac{1}{2}\right)\left(d + \frac{1}{2}\right)}{\left(b + \frac{1}{2}\right)\left(c + \frac{1}{2}\right)} & ;\ a \text{ ou } b \text{ ou } c \text{ ou } d = 0 \end{cases}$$

Assim a distribuição exata de probabilidade de θ^* quando se consideram duas distribuições binomiais com tamanhos iguais a 5 e probabilidades iguais a $\frac{1}{2}$ é dada por:

θ^*	f	$P(\theta^*)$
0.27	10	0.0098
0.13	20	0.0195
0.06	20	0.0195
0.03	10	0.0098
0.008	1	0.0010
1/16	25	0.0244
1/6	100	0.0977
3/8	100	0.0977
4/9	100	0.0977
1	252	0.2458
9/4	100	0.0977
8/3	100	0.0977
6	100	0.0977
16	25	0.0244
3.67	10	0.0098
7.86	20	0.0195
15.40	20	0.0195
33.00	10	0.0098
121.00	1	0.0010
Σ	1024	1.0000

Calculando a média e a variância de θ^* temos:

$$E(\theta^*) = 2.738 \quad \text{e} \quad \sigma = 5.886$$

EXEMPLO 2. Consideremos duas distribuições binomiais com tamanhos iguais a 10 e probabilidades iguais a $\frac{1}{2}$.

Temos a seguinte tabela:

x_1	x_2	10
x_3	x_4	10
m_1	m_2	20

Os possiveis valores para $\begin{cases} x_1 = 0,\ 1,\ 2,\ \dots,\ 10 \\ x_3 = 0,\ 1,\ 2,\ \dots,\ 10 \end{cases}$

Considerando as marginais m_1 e m_2 fixadas, todos os possiveis valores são, respectivamente:

$$\begin{cases} m_1 = 20,\ 19,\ 18,\ 17,\ 16,\ \dots,\ 3,\ 2,\ 1,\ 0 \\ m_2 = 0,\ 1,\ 2,\ 3,\ \dots,\ 16,\ 17,\ 18,\ 19,\ 20 \end{cases}$$

Podemos gerar 121 tabelas 2x2, classificando-as pelos valores marginais m_1 e m_2 .

A probabilidade de cada tabela será:

$$\binom{10}{x_1}\left(\frac{1}{2}\right)^{10}\cdot\binom{10}{x_3}\left(\frac{1}{2}\right)^{10} = \frac{\binom{10}{x_1}\binom{10}{x_3}}{1048576}$$

A probabilidade de cada tabela para valores dados das marginais m_1 e m_2 ($x_1 + x_3 = m_1$) segue a distribuição hipergeométrica dada por:

$$\frac{\binom{10}{x_1}\binom{10}{x_3}}{\binom{20}{m_1}}$$

As frequências das 121 tabelas 2x2 geradas por amostras de duas distribuições binomiais $n_1 = n_2 = 10$ e $p = \frac{1}{2}$, classificadas pelos valores das marginais m_1 e m_2, estão relacionadas no Apêndice B.2 e a seguir reproduzimos as tabelas para as marginais $m_1 = 17$ e $m_2 = 3$:

(10,0, 7,3)

10	0	10
7	3	10
17	3	

$$\binom{10}{10}\cdot\binom{10}{7} = 120 \qquad p_1 - p_2 = 0.3 \qquad \theta = \infty$$

(9,1, 8,2)

9	1	10
8	2	10
17	3	

$$\binom{10}{9}\cdot\binom{10}{8} = 450 \qquad p_1 - p_2 = 0,1 \qquad \theta = \frac{9}{4}$$

$(8,2,\ 9,1)$

8	2	10
9	1	10
17	3	

$$\binom{10}{8}\cdot\binom{10}{9} = 450 \qquad p_1-p_2=-0,1 \qquad \theta = \frac{4}{9}$$

$(7,3,10,0)$

7	3	10
10	0	10
17	3	

$$\binom{10}{7}\cdot\binom{10}{10} = 120 \qquad p_1-p_2=-0,3 \qquad \theta = 0$$

Observamos que a tabela 1 do exemplo 1 é simétrica em relação aos valores centrais 5,5 das marginais m_1 e m_2, e por causa desta simetria é que apresentaremos a tabela 2, a seguir, com as frequências relativas das 66 tabelas 2x2 ,de um total de 121, que podem ser geradas por amostras de duas distribuições binomiais com $n_1 = n_2 = 10$ e $p = \frac{1}{2}$, classificadas pelos valores das marginais m_1 e m_2.

A primeira coluna da tabela 2, a seguir, contém a tabela (10,0,10,0), a segunda as tabelas (9,1,10,0), (10,0,9,1), etc. Os números em parênteses são os elementos da distribuição hipergeométrica para os valores dados das marginais m_1 e m_2.

Este exemplo 2 vem ilustrar que ao mudarmos os valores das marginais a distribuição de θ se altera.

Assim é realmente importante encontrarmos uma transformação que fatore a multinomial em partes de modo que uma estatística dependa apenas do parâmetro de interesse θ e as outras independam de θ.

TABELA 2. Frequências relativas das 66 tabelas 2x2, de um total de 121, geradas por amostras de duas distribuições binomiais com $n_1 = n_2 = 10$ e $p = \frac{1}{2}$, classificadas pelos valores das marginais m_1 e m_2, onde o total das linhas corresponde ao total para todas as 121 tabelas, foi considerada a simetria.

$p_1 - p_2$	20,0	19,1	18,2	17,3	16,4	15,5	14,6	13,7	12,8	11,9	10,10	Total
-1.0											$\theta = 0$ 1 (0.000005)	1
-0.9										$\theta = 0$ 10 (0.00006)		20
-0.8									$\theta = 0$ 45 (0.0004)		$\theta = 1/81$ 100 (0.0005)	190
-0,7								$\theta = 0$ 120 (0.002)		$\theta = 1/36$ 450 (0.003)		1140
-0.6							$\theta = 0$ 210 (0.005)		$\theta = 1/21$ 1200 (0.010)		$\theta = 1/16$ 2025 (0.011)	4845
-0.5						$\theta = 0$ 252 (0.016)		$\theta = 2/27$ 2100 (0.027)		$\theta = 3/28$ 5400 (0.032)		15504
-0.4					$\theta = 0$ 210 (0.043)		$\theta = 1/9$ 2520 (0.065)		$\theta = 1/6$ 9450 (0.075)		$\theta = 9/49$ 14400 (0.078)	38760
-0.3				$\theta = 0$ 120 (0.105)		$\theta = 1/6$ 2100 (0.135)		$\theta = 1/4$ 11340 (0.146)		$\theta = 2/7$ 25200 (0.150)		77520
-0.2			$\theta = 0$ 45 (0.237)		$\theta = 7/27$ 1200 (0.248)		$\theta = 3/8$ 9450 (0.244)		$\theta = 3/7$ 30240 (0.240)		$\theta = 4/9$ 44100 (0.239)	125970
-0.1		$\theta = 0$ 10 (0.5)		$\theta = 4/9$ 450 (0.395)		$\theta = 7/12$ 5400 (0.348)		$\theta = 9/14$ 25200 (0.325)		$\theta = 2/3$ 52920 (0.315)		167960
0.0	$\theta = 0/0$ 1 (1.0)		$\theta = 1$ 100 (0.526)		$\theta = 1$ 2025 (0.418)		$\theta = 1$ 14400 (0.372)		$\theta = 1$ 44100 (0.350)		$\theta = 1$ 63540 (0.344)	184756
+0.1		$\theta = \infty$ 10 (0.5)		$\theta = 9/4$ 450 (0.395)		$\theta = 12/7$ 5400 (0.348)		$\theta = 14/9$ 25200 (0.325)		$\theta = 3/2$ 52920 (0.315)		167960
+0.2			$\theta = \infty$ 45 (0.237)		$\theta = 27/7$ 1200 (0.248)		$\theta = 8/3$ 9450 (0.244)		$\theta = 7/3$ 30240 (0.240)		$\theta = 9/4$ 44100 (0.239)	125970
+0.3				$\theta = \infty$ 120 (0.105)		$\theta = 6$ 2100 (0.135)		$\theta = 4$ 11340 (0.146)		$\theta = 7/2$ 25200 (0.150)		77520
+0.4					$\theta = \infty$ 210 (0.043)		$\theta = 9$ 2520 (0.065)		$\theta = 6$ 9450 (0.075)		$\theta = 49/9$ 14400 (0.078)	38760
+0.5						$\theta = \infty$ 252 (0.016)		$\theta = 27/2$ 2100 (0.027)		$\theta = 28/3$ 5400 (0.032)		15504
+0.6							$\theta = \infty$ 210 (0.005)		$\theta = 21$ 1200 (0.010)		$\theta = 16$ 2025 (0.011)	4845
+0.7								$\theta = \infty$ 120 (0.002)		$\theta = 36$ 450 (0.003)		1140
+0.8									$\theta = \infty$ 45 (0.0004)		$\theta = 81$ 100 (0.0005)	190
+0.9										$\theta = \infty$ 10 (0.00006)		20
+1.0											$\theta = \infty$ 1 (0.000005)	1
Σ	1	20	190	1140	4845	15504	38760	77520	125970	167960	184756	1048576

Distribuição exata de probabilidade do odds ratio (θ) quando se consideram duas binomiais com tamanhos iguais a 10 e probabilidades iguais a $\frac{1}{2}$.

θ	f	P(θ)	θ	f	P(θ)
0/0	2	0.0000	1	184754	0.1762
0	2045	0.0020	3/2	105840	0.1009
1/81	100	0.0001	14/9	50400	0.0481
1/36	900	0.0009	12/71	10800	0.0103
1/21	2400	0.0023	9/4	45000	0.0429
1/16	2025	0.0019	7/3	60480	0.0577
2/27	4200	0.0040	8/3	18900	0.0180
3/28	10800	0.0103	7/2	50400	0.0481
1/9	5040	0.0048	27/7	2400	0.0023
1/6	23100	0.0220	4	22680	0.0216
9/49	14400	0.0137	49/9	14400	0.0137
1/4	22680	0.0216	6	23100	0.0220
7/27	2400	0.0023	9	5040	0.0048
2/7	50400	0.0481	28/3	10800	0.0103
3/8	18900	0.0180	27/2	4200	0.0040
3/7	60480	0.0577	16	2025	0.0019
4/9	45000	0.0429	21	2400	0.0023
7/12	10800	0.0103	36	900	0.0009
9/14	50400	0.0481	81	100	0.0001
2/3	105840	0.1009	∞	2045	0.0020
			Σ	1048576	1.0000

Para as tabelas com θ = 0 , 0/0 e ∞ respectivamente, faremos a seguinte modificação proposta por Haldane e Anscombe (Gart e Zweifel , 1967).

Chamaremos θ^* o valor de θ modificado segundo Haldame e Anscombe, o qual é dado por:

a	b
c	d

$\rightarrow$

$a + \frac{1}{2}$	$b + \frac{1}{2}$
$c + \frac{1}{2}$	$d + \frac{1}{2}$

$$\theta = \frac{a\ d}{b\ c} \quad \rightarrow \quad \theta^* = \frac{\left(a + \frac{1}{2}\right)\left(d + \frac{1}{2}\right)}{\left(b + \frac{1}{2}\right)\left(c + \frac{1}{2}\right)}$$

Ou seja:

$$\theta^* = \begin{cases} \dfrac{a\ d}{b\ c} & ;\ a,b,c,d > 0 \\[2ex] \dfrac{\left(a + \frac{1}{2}\right)\left(d + \frac{1}{2}\right)}{\left(b + \frac{1}{2}\right)\left(c + \frac{1}{2}\right)} & ;\ a \text{ ou } b \text{ ou } c \text{ ou } d = 0 \end{cases}$$

Assim a distribuição exata de probabilidade de θ^* quando se consideram duas distribuições binomiais com tamanhos iguais a 10 e probabilidades iguais a $\frac{1}{2}$ é dada por:

θ^*	f	$P(\theta^*)$
0.302	20	0.0000
0.162	90	0.0001
0.102	240	0.0002
0.069	420	0.0004
0.048	504	0.0005
0.033	420	0.0004
0.022	240	0.0002
0.014	90	0.0001
0.008	20	0.0000
0.002	1	0.0000
1/81	100	0.0001
1/36	900	0.0009
1/21	2400	0.0023
1/16	2025	0.0019
2/27	4200	0.0040
3/28	10800	0.0103
1/9	5040	0.0048
1/6	23100	0.0220
9/49	14400	0.0137
1/4	22680	0.0216
7/27	2400	0.0023
2/7	50400	0.0481
3/8	18900	0.0180
3/7	60480	0.0577
4/9	45000	0.0429
7/12	10800	0.0103
9/14	50400	0.0481
2/3	105840	0.1009

θ^*	f	$P(\theta^*)$
1	184756	0.1762
3/2	105840	0.1009
14/9	50400	0.0481
12/71	10800	0.0103
9/4	45000	0.0429
7/3	60480	0.0577
8/3	18900	0.0180
7/2	50400	0.0481
27/7	2400	0.0023
4	22680	0.0216
49/9	14400	0.0137
6	23100	0.0220
9	5040	0.0048
28/3	10800	0.0103
27/2	4200	0.0040
16	2025	0.0019
21	2400	0.0023
36	900	0.0009
81	100	0.0001
3.316	20	0.0000
6.176	90	0.0001
9.800	240	0.0002
14.538	420	0.0004
21.000	504	0.0005
30.333	420	0.0004
45.000	240	0.0002
71.400	90	0.0001
133.000	20	0.0000
441.000	1	0.0000
Σ	1048576	1.0000

Calculando a média e a variância de θ^* temos:

$$E(\theta^*) = 1.714 \quad \text{e} \quad \sigma = 2.812$$

APÊNDICE A

A.1. Cálculo do jacobiano da 1ª transformação do capítulo 2.

Considere a transformação abaixo :

$$\theta = \frac{p_1 \; p_4}{p_2 \; p_3} \quad , \quad \phi = \frac{p_3}{1-p_1-p_2} \quad , \; \psi = p_1 + p_2$$

Temos que sua inversa é dada por:

$$p_1 = \frac{\theta \; \phi \; \psi}{1- \; \phi + \phi \; \theta}$$

$$p_2 = \frac{\psi \; (1-\phi)}{1-\phi \; + \; \phi \; \theta}$$

$$p_3 = \phi(1-\psi)$$

Derivando p_1 em relação a θ , ϕ e ψ temos:

$$\frac{\partial \; p_1}{\partial \; \theta} = \frac{\phi \; \psi \; (1- \; \phi \; + \; \theta \; \phi \;) - \theta \; \phi \; \psi \; \phi}{(1- \; \phi \; + \; \theta \; \phi \;)^2} = \frac{\phi \; \psi \; (1-\phi)}{(1- \; \phi \; + \; \theta \; \phi \;)^2}$$

$$\frac{\partial \; p_1}{\partial \; \phi} = \frac{\theta \; \psi \; (1- \; \phi \; + \; \theta \; \phi \;) - \theta \; \phi \; \psi \; (\theta \; -1)}{(1- \; \phi \; + \; \theta \; \phi \;)^2} = \frac{\theta \; \psi}{(1- \; \phi \; + \; \theta \; \phi \;)^2}$$

$$\frac{\partial \; p_1}{\partial \; \psi} = \frac{\theta \; \phi}{1- \; \phi \; + \; \theta \; \phi}$$

Derivando p_2 em relação a θ, ϕ e ψ temos:

$$\frac{\partial p_2}{\partial \theta} = \frac{-\psi\ (1-\ \phi\ )\ \phi}{(1-\ \phi + \theta\ \phi\ )^2}$$

$$\frac{\partial p_2}{\partial \phi} = \frac{-\ \psi\ (1-\ \phi + \theta\ \phi\ ) - \psi\ (1-\ \phi)(\theta\ -1)}{(1-\ \phi\ + \theta\ \phi\ )^2} = \frac{-\ \theta\ \psi}{(1-\ \phi\ + \theta\ \phi\ )^2}$$

$$\frac{\partial p_2}{\partial \psi} = \frac{1-\ \phi}{1-\ \phi\ + \theta\ \phi}$$

Derivando p_3 em relação a θ, ϕ e ψ temos:

$$\frac{\partial p_3}{\partial \theta}\ 0 \qquad\qquad \frac{\partial p_2}{\partial \phi} = 1 - \psi \qquad\qquad \frac{\partial p_3}{\partial \psi} = -\ \phi$$

Então o jacobiano é dado por:

$$J(\theta,\phi,\psi) = \begin{vmatrix} \dfrac{\phi\ \psi\ (1-\phi)}{(1-\ \phi\ + \theta\ \phi\ )^2} & \dfrac{\theta\ \psi}{(1-\ \phi\ + \theta\ \phi\ )^2} & \dfrac{\theta\ \phi}{1-\ \phi\ + \theta\ \phi} \\[2ex] \dfrac{-\ \psi\ (1-\ \phi\ )\ \phi}{(1-\ \phi\ + \theta\ \phi\ )^2} & \dfrac{-\ \theta\ \psi}{(1-\ \phi\ + \theta\ \phi\ )^2} & \dfrac{1-\ \phi}{1-\ \phi\ + \theta\ \phi} \\[2ex] 0 & 1\ -\ \psi & -\ \phi \end{vmatrix}$$

$$= \frac{\theta \psi^2 \phi^2 (1-\phi)}{(1-\phi+\theta\phi)^4} - \frac{\theta \phi^2 \psi (1-\phi)(1-\psi)}{(1-\phi+\theta\phi)^3} - \frac{\phi \psi (1-\phi)^2 (1-\psi)}{(1-\phi+\theta\phi)^3} +$$

$$- \frac{\theta \phi^2 \psi^2 (1-\phi)}{(1-\phi+\theta\phi)^4} = \frac{-\phi \psi (1-\phi)(1-\psi)(1-\phi+\theta\phi)}{(1-\phi+\theta\phi)^3} =$$

$$= \frac{-\phi \psi (1-\phi)(1-\psi)}{(1-\phi+\theta\phi)^2} \neq 0 \qquad \blacksquare$$

A.2. Cálculo do jacobiano da 2ª transformação do capitulo 2.

Seja a transformação abaixo :

$$\gamma_1 = p_1 \quad , \quad \gamma_2 = \frac{p_2}{1 - p_1} \quad , \quad \gamma_3 = \frac{p_3}{1 - p_1 - p_2}$$

Temos que sua inversa é dada por :

$$p_1 = \gamma_1$$

$$p_2 = \gamma_2(1-\gamma_1)$$

$$p_3 = \gamma_3(1-\gamma_2)(1-\gamma_1)$$

Assim o jacobiano é dado por:

$$J\,(\,\gamma_1,\gamma_2,\gamma_3\,) = \begin{vmatrix} 1 & 0 & 0 \\ -\gamma_2 & 1-\gamma_1 & 0 \\ -\gamma_3(1-\gamma_2) & -\gamma_3(1-\gamma_1) & (1-\gamma_1)(1-\gamma_2) \end{vmatrix} =$$

$$= (1-\gamma_1)^2(1-\gamma_2) \neq 0 \qquad \blacksquare$$

A.3. Cálculo do jacobiano da 3ª transformação do capitulo 2.

Considere a transformação abaixo :

$$\psi = p_1+p_2 \quad , \delta = \frac{p_1}{p_1+p_2} \quad , \quad \phi = \frac{p_3}{1-p_1-p_2}$$

Temos que:

$$\psi = p_1+p_2$$

$$\delta = \frac{p_1}{p_1+p_2}$$

$$\phi = \frac{p_3}{1-p_1-p_2}$$

Assim a sua inversa é dada por :

$$p_1 = \psi\,\delta$$

$$p_2 = \psi\,(1-\delta)$$

$$p_3 = \phi\,(1-\psi)$$

Então calculando o jacobiano temos:

$$J(p_1, p_2, p_3) = \begin{vmatrix} 1 & 1 & 0 \\ \dfrac{p_2}{(p_1+p_2)^2} & \dfrac{-p_1}{(p_1+p_2)^2} & 0 \\ \dfrac{p_3}{(1-p_1-p_2)^2} & \dfrac{p_3}{(1-p_1-p_2)^2} & \dfrac{1}{1-p_1-p_2} \end{vmatrix} =$$

$$= \frac{-p_1}{(p_1+p_2)^2(1-p_1-p_2)} + \frac{-p_2}{(p_1+p_2)^2(1-p_1-p_2)} = \frac{-p_1-p_2}{(p_1+p_2)^2(1-p_1-p_2)}$$

$$= \frac{-1}{(p_1+p_2)(1-p_1-p_2)} \neq 0 \qquad \blacksquare$$

A.4. Cálculo do jacobiano da 4ª transformação do capítulo 2.

Considerando a transformação abaixo :

$$\psi = p_1 + p_2 \quad , \quad \alpha = \ln \frac{p_3}{p_4} \quad , \quad \lambda = \ln \frac{p_1\ p_4}{p_2\ p_3}$$

Então temos que sua inversa é dada por :

$$p_1 = \frac{\psi\ e^{\lambda+\alpha}}{1 + e^{\lambda+\alpha}} \quad , \quad p_2 = \frac{\psi}{1 + e^{\lambda+\alpha}} \quad , \quad p_3 = \frac{e^{\alpha} (1 - \psi)}{1 + e^{\alpha}}$$

Derivando p_1, p_2 e p_3 em relação a λ temos:

$$\frac{\partial p_1}{\partial \lambda} = \frac{\psi\ e^{\lambda+\alpha} \left(1 + e^{\lambda+\alpha}\right) - \psi\ e^{\lambda+\alpha}\ e^{\lambda+\alpha}}{\left(1 + e^{\lambda+\alpha} \right)^2} = \frac{\psi\ e^{\lambda+\alpha}}{\left(1 + e^{\lambda+\alpha} \right)^2}$$

$$\frac{\partial p_2}{\partial \lambda} = \frac{-\ \psi\ e^{\lambda+\alpha}}{\left(1 + e^{\lambda+\alpha} \right)^2} \qquad \qquad \frac{\partial p_2}{\partial \lambda} = 0$$

Derivando p_1, p_2 e p_3 em relação a α temos:

$$\frac{\partial p_1}{\partial \alpha} = \frac{\psi\ e^{\lambda+\alpha} \left(1 + e^{\lambda+\alpha}\right) - \psi\ e^{\lambda+\alpha}\ e^{\lambda+\alpha}}{\left(1 + e^{\lambda+\alpha} \right)^2} = \frac{\psi\ e^{\lambda+\alpha}}{\left(1 + e^{\lambda+\alpha} \right)^2}$$

$$\frac{\partial\, p_2}{\partial\, \alpha} = \frac{-\,\psi\, e^{\lambda+\alpha}}{\left(1 + e^{\lambda+\alpha}\right)^2}$$

$$\frac{\partial\, p_3}{\partial\, \alpha} = \frac{e^{\alpha}\,(1-\,\psi)\left(1+\,e^{\alpha}\right) - e^{\alpha}\,(1-\,\psi)\,e^{\alpha}}{\left(1\,+\,e^{\alpha}\right)^2} = \frac{e^{\alpha}\,(1-\,\psi)}{\left(1\,+\,e^{\alpha}\right)^2}$$

Assim calculando o jacobiano temos:

$$J(\,\psi,\lambda,\alpha\,) = \begin{vmatrix} \dfrac{e^{\lambda+\alpha}}{1\,+\,e^{\lambda+\alpha}} & \dfrac{\psi\,e^{\lambda+\alpha}}{\left(1\,+\,e^{\lambda+\alpha}\right)^2} & \dfrac{\psi\,e^{\lambda+\alpha}}{\left(1\,+\,e^{\lambda+\alpha}\right)^2} \\ \dfrac{1}{1\,+\,e^{\lambda+\alpha}} & \dfrac{-\,\psi\,e^{\lambda+\alpha}}{\left(1\,+\,e^{\lambda+\alpha}\right)^2} & \dfrac{-\,\psi\,e^{\lambda+\alpha}}{\left(1\,+\,e^{\lambda+\alpha}\right)^2} \\ \dfrac{-\,e^{\alpha}}{1\,+\,e^{\alpha}} & 0 & \dfrac{e^{\alpha}\,(\,1\,-\,\psi\,)}{\left(\,1\,+\,e^{\alpha}\right)^2} \end{vmatrix} =$$

$$= \frac{-\psi\,(1-\,\psi)\,e^{\alpha}\,e^{\lambda+\alpha}\,e^{\lambda+\alpha}}{\left(1+\,e^{\alpha}\right)^2\left(1+\,e^{\lambda+\alpha}\right)^3} + \frac{\psi^2\,e^{\lambda+\alpha}\,e^{\lambda+\alpha}\,e^{\alpha}}{\left(1+\,e^{\lambda+\alpha}\right)^4\left(1+\,e^{\alpha}\right)} +$$

$$+ \frac{-\,\psi^2\,e^{\lambda+\alpha}\,e^{\lambda+\alpha}\,e^{\alpha}}{\left(1+\,e^{\lambda+\alpha}\right)^4\left(1+\,e^{\alpha}\right)} + \frac{-\psi\,(1-\,\psi)\,e^{\alpha}\,e^{\lambda+\alpha}}{\left(1+\,e^{\alpha}\right)^2\left(1+\,e^{\lambda+\alpha}\right)^3} =$$

$$= \frac{-\psi\,(1-\,\psi)\,e^{\alpha}\,e^{\lambda+\alpha}\left(e^{\lambda+\alpha}+\,1\right)}{\left(1+\,e^{\alpha}\right)^2\left(1+\,e^{\lambda+\alpha}\right)^3} = \frac{-\psi\,(1-\,\psi)\,e^{\alpha}\,e^{\lambda+\alpha}}{\left(1+\,e^{\alpha}\right)^2\left(1+\,e^{\lambda+\alpha}\right)^2}$$

$\neq 0$ ■

A.5. Cálculo do jacobiano da 5ª transformação do capítulo 2.

Considere a transformação abaixo :

$$\theta = \frac{p_1\ p_4}{p_2\ p_3} \quad , \quad \beta = \frac{p_2}{p_4} \quad , \quad \gamma = \frac{p_3}{p_4}$$

Assim temos que sua inversa é dada por :

$$p_1 = \frac{\theta\ \beta\ \gamma}{1 + \beta + \gamma + \theta\ \beta\ \gamma}$$

$$p_2 = \frac{\beta}{1 + \beta + \gamma + \theta\ \beta\ \gamma}$$

$$p_3 = \frac{\gamma}{1 + \beta + \gamma + \theta\ \beta\ \gamma}$$

Temos:

$$\theta = \frac{p_1\ p_4}{p_2\ p_3} = \frac{p_1(1-p_1-p_2-p_3)}{p_2\ p_3} = \frac{p_1-p_1^2-p_1p_2-p_1p_3}{p_2\ p_3}$$

$$\beta = \frac{p_2}{p_4} = \frac{p_2}{1-p_1-p_2-p_3}$$

$$\gamma = \frac{p_3}{p_4} = \frac{p_3}{1-p_1-p_2-p_3}$$

Derivando θ em relação a p_1, p_2 e p_3 temos:

$$\frac{\partial\, \theta}{\partial\, p_1} = \frac{1-2p_1-p_2-p_3}{p_2\; p_3}$$

$$\frac{\partial\, \theta}{\partial\, p_2} = \frac{-\; p_1p_2p_3 \;-\; p_1(1-p_1-p_2-p_3)p_3}{(p_2\;\; p_3)^2} = \frac{-\; p_1(1-p_1-p_3)}{p_2^2\; p_3}$$

$$\frac{\partial\, \theta}{\partial\, p_3} = \frac{-\; p_1p_2p_3 \;-\; p_1(1-p_1-p_2-p_3)p_2}{(p_2\;\; p_3)^2} = \frac{-\; p_1(1-p_1-p_2)}{p_2\; p_3^2}$$

Derivando β em relação a p_1, p_2 e p_3 temos:

$$\frac{\partial\, \beta}{\partial\, p_1} = \frac{p_2}{(1-p_1-p_2-p_3)^2}$$

$$\frac{\partial\, \beta}{\partial\, p_2} = \frac{(1-p_1-p_2-p_3)+p_2}{(1-p_1-p_2-p_3)^2} = \frac{1-p_1-p_3}{(1-p_1-p_2-p_3)^2}$$

$$\frac{\partial\, \beta}{\partial\, p_3} = \frac{p_2}{(1-p_1-p_2-p_3)^2}$$

Derivando γ em relação a p_1, p_2 e p_3 temos:

$$\frac{\partial\, \gamma}{\partial\, p_1} = \frac{p_3}{(1-p_1-p_2-p_3)^2}$$

$$\frac{\partial\, \gamma}{\partial\, p_2} = \frac{p_3}{(1-p_1-p_2-p_3)^2}$$

$$\frac{\partial \gamma}{\partial p_3} = \frac{(1-p_1-p_2-p_3)+p_3}{(1-p_1-p_2-p_3)^2} = \frac{1-p_1-p_2}{(1-p_1-p_2-p_3)^2}$$

Então o jacobiano é dado por:

$$J(p_1,p_2,p_3) = \begin{vmatrix} \dfrac{1-2p_1-p_2-p_3}{p_2\ p_3} & \dfrac{-\ p_1(1-p_1-p_3)}{p_2^2\ p_3} & \dfrac{-\ p_1(1-p_1-p_2)}{p_2\ p_3^2} \\[2ex] \dfrac{p_2}{(1-p_1-p_2-p_3)^2} & \dfrac{1-p_1-p_3}{(1-p_1-p_2-p_3)^2} & \dfrac{p_2}{(1-p_1-p_2-p_3)^2} \\[2ex] \dfrac{p_3}{(1-p_1-p_2-p_3)^2} & \dfrac{p_3}{(1-p_1-p_2-p_3)^2} & \dfrac{1-p_1-p_2}{(1-p_1-p_2-p_3)^2} \end{vmatrix}$$

$$= \frac{(1-2p_1-p_2-p_3)(1-p_1-p_3)(1-p_1-p_2)}{p_2p_3(1-p_1-p_2-p_3)^4} - \frac{p_1(1-p_1-p_2)}{p_3(1-p_1-p_2-p_3)^4} -$$

$$- \frac{p_1(1-p_1-p_3)}{p_2(1-p_1-p_2-p_3)^4} + \frac{p_1(1-p_1-p_3)(1-p_1-p_2)}{p_2p_3(1-p_1-p_2-p_3)^4} -$$

$$- \frac{1-2p_1-p_2-p_3}{(1-p_1-p_2-p_3)^4} + \frac{p_1(1-p_1-p_3)(1-p_1-p_2)}{p_2p_3(1-p_1-p_2-p_3)^4} =$$

$$
= \frac{(1-p_2-p_3)(1-p_1-p_3)(1-p_1-p_2)}{p_2 p_3 (1-p_1-p_2-p_3)^4} + \frac{-\, p_1 p_2 (1-p_1-p_2) - p_1 p_3 (1-p_1-p_3)}{p_2 p_3 (1-p_1-p_2-p_3)^4} -
$$

$$
- \frac{p_2 p_3 (1-2p_1-p_2-p_3)}{p_2 p_3 (1-p_1-p_2-p_3)^4} =
$$

$$
= \frac{(1-p_1-p_2)(1-p_1-p_3)(1-p_2-p_3) - p_1 p_2 (1-p_1-p_2) - p_1 p_3 (1-p_1-p_3)}{p_2 p_3 (1-p_1-p_2-p_3)^4} -
$$

$$
- \frac{p_2 p_3 (1-2p_1-p_2-p_3)}{p_2 p_3 (1-p_1-p_2-p_3)^4} =
$$

$$
= \frac{(1-p_1-p_2)\left[1-p_2-p_3-p_1+p_1 p_2+p_1 p_3-p_3+p_2 p_3+p_3^2\right] - p_1 p_2 + p_1^2 p_2}{p_2 p_3 (1-p_1-p_2-p_3)^4} +
$$

$$
+ \frac{p_1 p_2^2 - p_1 p_3 + p_1^2 p_3 + p_1 p_3^2 - p_2 p_3 + 2p_1 p_2 p_3 + p_2^2 p_3 + p_2 p_3^2}{p_2 p_3 (1-p_1-p_2-p_3)^4} =
$$

$$= \frac{1-p_2-p_3-p_1+p_1p_2+p_1p_3-p_3+p_2p_3+p_3^2 - p_1 + p_1p_2 + p_1p_3 + p_1^2}{p_2p_3(1-p_1-p_2-p_3)^4} +$$

$$+ \frac{- p_1^2p_2 - p_1^2p_3 + p_1p_3 - p_1p_2p_3 - p_1p_3^2 - p_2 + p_2^2 + p_2p_3 + p_1p_2}{p_2p_3(1-p_1-p_2-p_3)^4} +$$

$$+ \frac{- p_1p_2^2 - p_1p_2p_3 + p_2p_3 - p_2^2p_3 - p_2p_3^2 - p_1p_2 + p_1^2p_2}{p_2p_3(1-p_1-p_2-p_3)^4} +$$

$$+ \frac{+ p_1p_2^2 - p_1p_3 + p_1^2p_3 + p_1p_3^2 - p_2p_3 + 2p_1p_2p_3 + p_2^2p_3 + p_2p_3^2}{p_2p_3(1-p_1-p_2-p_3)^4} =$$

$$= \frac{1 - 2p_1 - 2p_2 - 2p_3 + 2p_1p_2 + 2p_1p_3 + 2p_2p_3 + p_1^2 + p_2^2 + p_3^2}{p_2p_3(1-p_1-p_2-p_3)^4} =$$

$$= \frac{(1-p_1-p_2-p_3)^2}{p_2p_3(1-p_1-p_2-p_3)^4} = \frac{1}{p_2p_3(1-p_1-p_2-p_3)^2} \neq 0 \qquad \blacksquare$$

A.6. Cálculo do jacobiano da 6ª transformação do capítulo 2.

Considere a transformação abaixo :

$$\omega_1 = \frac{p_1}{1 - p_1 - p_2 - p_3}$$

$$\omega_2 = \frac{p_2}{1 - p_2 - p_3}$$

$$\omega_3 = \frac{p_3}{1 - p_3}$$

Temos que sua inversa é dada por :

$$p_1 = \frac{\omega_1}{(\omega_1+1)(\omega_2+1)(\omega_3+1)}$$

$$p_2 = \frac{\omega_2}{(\omega_2+1)(\omega_3+1)}$$

$$p_3 = \frac{\omega_3}{(\omega_3+1)}$$

Derivando p_1 em relação a ω_1,ω_2 e ω_3, temos:

$$\frac{\partial p_1}{\partial \omega_1} = \frac{(\omega_1+1)(\omega_2+1)(\omega_3+1) - \omega_1(\omega_2+1)(\omega_3+1)}{(\omega_1+1)^2(\omega_2+1)^2(\omega_3+1)^2} = \frac{1}{(\omega_1+1)^2(\omega_2+1)(\omega_3+1)}$$

$$\frac{\partial p_1}{\partial \omega_2} = \frac{-\omega_1(\omega_1+1)(\omega_3+1)}{(\omega_1+1)^2(\omega_2+1)^2(\omega_3+1)^2} = \frac{-\omega_1}{(\omega_1+1)(\omega_2+1)^2(\omega_3+1)}$$

$$\frac{\partial\, p_1}{\partial\, \omega_3} = \frac{-\,\omega_1(\omega_1+1)(\omega_2+1)}{(\omega_1+1)^2(\omega_2+1)^2(\omega_3+1)^2} = \frac{-\,\omega_1}{(\omega_1+1)(\omega_2+1)(\omega_3+1)^2}$$

Derivando p_2 em relação a ω_1, ω_2 e ω_3 temos:

$$\frac{\partial\, p_2}{\partial\, \omega_1} = 0$$

$$\frac{\partial\, p_2}{\partial\, \omega_2} = \frac{(\omega_2+1)(\omega_3+1) - \omega_2(\omega_3+1)}{(\omega_2+1)^2(\omega_3+1)^2} = \frac{1}{(\omega_2+1)^2(\omega_3+1)}$$

$$\frac{\partial\, p_2}{\partial\, \omega_3} = \frac{-\,\omega_2(\omega_2+1)}{(\omega_2+1)^2(\omega_3+1)^2} = \frac{-\,\omega_2}{(\omega_2+1)(\omega_3+1)^2}$$

Derivando p_3 em relação a ω_1, ω_2 e ω_3 temos:

$$\frac{\partial\, p_3}{\partial\, \omega_1} = 0$$

$$\frac{\partial\, p_3}{\partial\, \omega_2} = 0$$

$$\frac{\partial\, p_3}{\partial\, \omega_3} = \frac{(\omega_3+1) - \omega_3}{(\omega_3+1)^2} = \frac{1}{(\omega_3+1)^2}$$

Assim o Jacobiano em relação a ω_1, ω_2 e ω_3 é dado por:

$$\begin{vmatrix} \dfrac{1}{(\omega_1+1)^2(\omega_2+1)(\omega_3+1)} & \dfrac{-\omega_1}{(\omega_1+1)(\omega_2+1)^2(\omega_3+1)} & \dfrac{-\omega_1}{(\omega_1+1)(\omega_2+1)(\omega_3+1)^2} \\ 0 & \dfrac{1}{(\omega_2+1)^2(\omega_3+1)} & \dfrac{-\omega_2}{(\omega_2+1)(\omega_3+1)^2} \\ 0 & 0 & \dfrac{1}{(\omega_3+1)^2} \end{vmatrix} =$$

$$= \frac{1}{(\omega_1+1)^2(\omega_2+1)^3(\omega_3+1)^4} \neq 0 \qquad \blacksquare$$

A.7. Cálculo da matriz de informação de Fisher para a 6ª transformação do capitulo 2.

Seja a transformação:

$$p_1 = \frac{\omega_1}{(\omega_1+1)(\omega_2+1)(\omega_3+1)}$$

$$p_2 = \frac{\omega_2(\omega_1+1)}{(\omega_1+1)(\omega_2+1)(\omega_3+1)}$$

$$p_3 = \frac{\omega_3(\omega_2+1)(\omega_1+1)}{(\omega_1+1)(\omega_2+1)(\omega_3+1)}$$

Assim temos:

$$p_4 = \frac{1}{(\omega_1+1)(\omega_2+1)(\omega_3+1)}$$

A multinomial pode ser escrita como:

$$L = P(x_1,x_2,x_3/\omega_1,\omega_2,\omega_3) = \frac{n!}{x_1!x_2!x_3!(n-x_1-x_2-x_3)!}$$

$$\cdot \frac{\omega_1^{x_1}\ \omega_2^{x_2}\ \omega_3^{x_3}\ (w_1+1)^{x_2}\ (\omega_2+1)^{x_3}\ (\omega_1+1)^{x_3}}{(\omega_1+1)^n\ (\omega_2+1)^n\ (\omega_3+1)^n}$$

Tomando o logaritmo de L, temos:

$$\log L = \log\left[\frac{n!}{x_1!x_2!x_3!(n-x_1-x_2-x_3)!}\right] + x_1\log\omega_1 + x_2\log\omega_2 + x_3\log\omega_3 +$$

$$+ x_2\log(\omega_1+1) + x_3\log(\omega_2+1) + x_3\log(\omega_1+1) - n\log(\omega_1+1) -$$

$$- n\log(\omega_2+1) - n\log(\omega_3+1)$$

Serão calculadas as seguintes derivadas para a obtenção dos elementos da correspondente matriz de informação de Fisher:

a) $$\frac{\partial \log L}{\partial \omega_1} = \frac{x_1}{\omega_1} + \frac{x_2}{\omega_1+1} + \frac{x_3}{\omega_1+1} - \frac{n}{\omega_1+1}$$

$$\frac{\partial^2 \log L}{\partial \omega_1^2} = -\frac{x_1}{\omega_1^2} - \frac{x_2}{(\omega_1+1)^2} - \frac{x_3}{(\omega_1+1)^2} + \frac{n}{(\omega_1+1)^2}$$

$$\frac{\partial^2 \log L}{\partial \omega_1 \partial \omega_2} = 0 \qquad\qquad \frac{\partial^2 \log L}{\partial \omega_1 \partial \omega_3} = 0$$

b) $$\frac{\partial \log L}{\partial \omega_2} = \frac{x_2}{\omega_2} + \frac{x_3}{\omega_2+1} - \frac{n}{\omega_2+1}$$

$$\frac{\partial^2 \log L}{\partial \omega_2^2} = -\frac{x_2}{\omega_2^2} - \frac{x_3}{(\omega_2+1)^2} + \frac{n}{(\omega_2+1)^2}$$

$$\frac{\partial^2 \log L}{\partial \omega_2 \partial \omega_1} = 0 \qquad\qquad \frac{\partial^2 \log L}{\partial \omega_2 \partial \omega_3} = 0$$

c) $$\frac{\partial \log L}{\partial \omega_3} = \frac{x_3}{\omega_3} - \frac{n}{\omega_3+1}$$

$$\frac{\partial^2 \log L}{\partial \omega_3^2} = - \frac{x_3}{\omega_3^2} + \frac{n}{(\omega_3+1)^2}$$

$$\frac{\partial^2 \log L}{\partial \omega_3 \partial \omega_1} = 0 \qquad\qquad \frac{\partial^2 \log L}{\partial \omega_3 \partial \omega_2} = 0$$

Então a matriz de informação de Fisher é dada por:

$$I(\omega) = E \begin{bmatrix} -\dfrac{\partial^2 \log L}{\partial \omega_1^2} & -\dfrac{\partial^2 \log L}{\partial \omega_1 \partial \omega_2} & -\dfrac{\partial^2 \log L}{\partial \omega_1 \partial \omega_3} \\ -\dfrac{\partial^2 \log L}{\partial \omega_2 \partial \omega_1} & -\dfrac{\partial^2 \log L}{\partial \omega_2^2} & -\dfrac{\partial^2 \log L}{\partial \omega_2 \partial \omega_3} \\ -\dfrac{\partial^2 \log L}{\partial \omega_3 \partial \omega_1} & -\dfrac{\partial^2 \log L}{\partial \omega_3 \partial \omega_2} & -\dfrac{\partial^2 \log L}{\partial \omega_3^2} \end{bmatrix}$$

Mas sabemos que:

$$E(x_1) = n\ p_1 = \frac{n\ \omega_1}{(\omega_1+1)(\omega_2+1)(\omega_3+1)}$$

$$E(x_2) = n\ p_2 = \frac{n\ \omega_2}{(\omega_2+1)(\omega_3+1)}$$

$$E(x_3) = n\ p_3 = \frac{n\ \omega_3}{\omega_3+1}$$

Assim temos:

$$E\left(-\frac{\partial^2 \log\ L}{\partial\ \omega_1^2}\right) = E\left(\frac{x_1}{\omega_1^2} + \frac{x_2}{(\omega_1+1)^2} + \frac{x_3}{(\omega_1+1)^2} - \frac{n}{(\omega_1+1)^2}\right) =$$

$$= \frac{n\ \omega_1}{\omega_1^2(\omega_1+1)(\omega_2+1)(\omega_3+1)} + \frac{n\ \omega_2}{(\omega_1+1)^2(\omega_2+1)(\omega_3+1)} + \frac{n\ \omega_3}{(\omega_1+1)^2(\omega_3+1)} - \frac{n}{(\omega_1+1)^2}$$

$$= \frac{n(\omega_1+1) + n\ \omega_1\omega_2 + n\ \omega_1\omega_3(\omega_2+1) - n\ \omega_1(\omega_2+1)(\omega_3+1)}{\omega_1(\omega_1+1)^2(\omega_2+1)(\omega_3+1)} =$$

$$= \frac{n\ \omega_1 + n + n\ \omega_1\omega_2 + n\ \omega_1\omega_2\omega_3 + n\ \omega_1 w_3 - n\ \omega_1\omega_2\omega_3}{\omega_1(\omega_1+1)^2(\omega_2+1)(\omega_3+1)} +$$

$$+ \frac{-\ n\ w_1\omega_2 - n\ \omega_1\omega_3 - n\ \omega_1}{\omega_1(\omega_1+1)^2(\omega_2+1)(\omega_3+1)} =$$

$$= \frac{n}{\omega_1(\omega_1+1)^2(\omega_2+1)(\omega_3+1)}$$

$$E\left(-\frac{\partial^2 \log L}{\partial \omega_2^2}\right) = E\left(\frac{x_2}{\omega_2^2} + \frac{x_3}{(\omega_2+1)^2} - \frac{n}{(\omega_2+1)^2}\right) =$$

$$= \frac{n\ \omega_2}{\omega_2^2(\omega_2+1)(\omega_3+1)} + \frac{n\ \omega_3}{(\omega_2+1)^2(\omega_3+1)} - \frac{n}{(\omega_2+1)^2} =$$

$$= \frac{n(\omega_2+1) + n\ \omega_2\omega_3 - n\ \omega_2(\omega_3+1)}{\omega_2(\omega_2+1)^2(\omega_3+1)} =$$

$$= \frac{n\ w_2 + n + n\ \omega_2\omega_3 - n\ \omega_2 w_3 - n\ \omega_2}{\omega_2(\omega_2+1)^2(\omega_3+1)} =$$

$$= \frac{n}{\omega_2(\omega_2+1)^2(\omega_3+1)}$$

$$E\left(-\frac{\partial^2 \log L}{\partial \omega_3^2}\right) = E\left(\frac{x_3}{\omega_3^2} - \frac{n}{(\omega_3+1)^2}\right) =$$

$$= \frac{n\ \omega_3}{\omega_3^2(\omega_3+1)} - \frac{n}{(\omega_3+1)^2} = \frac{n\ (\omega_3+1) - n\ \omega_3}{\omega_3(\omega_3+1)^2} = \frac{n}{\omega_3(\omega_3+1)^2}$$

Assim a matriz de informação de Fisher é dada por:

$$I(\omega) = \begin{bmatrix} \dfrac{n}{\omega_1(\omega_1+1)^2(\omega_2+1)(\omega_3+1)} & 0 & 0 \\ 0 & \dfrac{n}{\omega_2(\omega_2+1)^2(\omega_3+1)} & 0 \\ 0 & 0 & \dfrac{n}{\omega_3(\omega_3+1)^2} \end{bmatrix}$$

Observamos que $I(\omega)$ é uma matriz diagonal. ∎

APÊNDICE B

B.1. Cálculo das frequências relativas das tabelas do exemplo 1 do capítulo 4.

As frequencias relativas das 36 tabelas, classificadas pelos valores das marginais m_1 e m_2 ,estão relacionadas abaixo:

(5,0,5,0)

5	0	5
5	0	5
10	0	

$$\binom{5}{5}\cdot\binom{5}{5} = 1 \qquad p_1 - p_2 = 0{,}0 \qquad \theta = \frac{0}{0}$$

(5,0,4,1)

5	0	5
4	1	5
9	1	

$$\binom{5}{5}\cdot\binom{5}{4} = 5 \qquad p_1 - p_2 = 0{,}2 \qquad \theta = \infty$$

(4,1,5,0)

4	1	5
5	0	5
9	1	

$$\binom{5}{4}\cdot\binom{5}{5} = 5 \qquad p_1 - p_2 = -0{,}2 \qquad \theta = 0$$

(5,0,3,2)

5	0	5
3	2	5
8	2	

$$\binom{5}{5}\cdot\binom{5}{3} = 10 \qquad p_1 - p_2 = 0{,}4 \qquad \theta = \infty$$

(4,1,4,1)

4	1	5
4	1	5
8	2	

$$\binom{5}{4}\cdot\binom{5}{4} = 25 \qquad p_1 - p_2 = 0{,}0 \qquad \theta = 1$$

(3,2,5,0)

3	2	5
5	0	5
8	2	

$$\binom{5}{3}\cdot\binom{5}{5} = 10 \qquad p_1 - p_2 = -0{,}4 \qquad \theta = 0$$

(5,0,2,3)

5	0	5
2	3	5
7	3	

$$\binom{5}{5}\cdot\binom{5}{2} = 10 \qquad p_1 - p_2 = 0{,}6 \qquad \theta = \infty$$

(4,1,3,2)

4	1	5
3	2	5
7	3	

$$\binom{5}{4}\cdot\binom{5}{3} = 50 \qquad p_1 - p_2 = 0{,}2 \qquad \theta = \frac{8}{3}$$

(3,2,4,1)

3	2	5
4	1	5
7	3	

$$\binom{5}{3}\cdot\binom{5}{4} = 50 \qquad p_1 - p_2 = -0{,}2 \qquad \theta = \frac{3}{8}$$

(2,3,5,0)

2	3	5
5	0	5
7	3	

$$\binom{5}{2}\cdot\binom{5}{5} = 10 \qquad p_1 - p_2 = -0{,}6 \qquad \theta = 0$$

(5,0,1,4)

5	0	5
1	4	5
6	4	

$$\binom{5}{5}\cdot\binom{5}{1} = 5 \qquad p_1 - p_2 = 0{,}8 \qquad \theta = \infty$$

(4,1,2,3)

4	1	5
2	3	5
6	4	

$$\binom{5}{4}\cdot\binom{5}{2} = 50 \qquad p_1 - p_2 = 0{,}4 \qquad \theta = 6$$

(3,2,3,2)

3	2	5
3	2	5
6	4	

$$\binom{5}{3}\cdot\binom{5}{3} = 100 \qquad p_1 - p_2 = 0{,}0 \qquad \theta = 1$$

(2,3,4,1)

2	3	5
4	1	5
6	4	

$$\binom{5}{2}\cdot\binom{5}{4} = 50 \qquad p_1 - p_2 = -0{,}4 \qquad \theta = \frac{1}{6}$$

(1,4,5,0)

1	4	5
5	0	5
6	4	

$$\binom{5}{1}\cdot\binom{5}{5} = 5 \qquad p_1 - p_2 = -0{,}8 \qquad \theta = 0$$

(5,0,0,5)

5	0	5
0	5	5
5	5	

$$\binom{5}{5}\cdot\binom{5}{0} = 1 \qquad p_1 - p_2 = 1{,}0 \qquad \theta = \infty$$

(4,1,1,4)

4	1	5
1	4	5
5	5	

$$\binom{5}{4}\cdot\binom{5}{1} = 25 \qquad p_1 - p_2 = 0{,}6 \qquad \theta = 16$$

(3,2,2,3)

3	2	5
2	3	5
5	5	

$$\binom{5}{3}\cdot\binom{5}{2} = 100 \qquad p_1 - p_2 = 0{,}2 \qquad \theta = \frac{9}{4}$$

(2,3,3,2)

2	3	5
3	2	5
5	5	

$$\binom{5}{2} \cdot \binom{5}{3} = 100 \qquad p_1 - p_2 = -0{,}2 \qquad \theta = \frac{4}{9}$$

(1,4,4,1)

1	4	5
4	1	5
5	5	

$$\binom{5}{1} \cdot \binom{5}{4} = 25 \qquad p_1 - p_2 = -0{,}6 \qquad \theta = \frac{1}{16}$$

(0,5,5,0)

0	5	5
5	0	5
5	5	

$$\binom{5}{0} \cdot \binom{5}{5} = 1 \qquad p_1 - p_2 = -1{,}0 \qquad \theta = 0$$

B.2. Cálculo das frequências relativas das tabelas do exemplo 2 do capítulo 4.

As frequencias relativas das 121 tabelas 2x2 geradas por amostras de duas distribuições binomiais $n_1 = n_2 = 10$ e $p = \frac{1}{2}$, classificadas pelos valores das marginais m_1 e m_2 , estão relacionadas abaixo:

(10,0,10,0)

10	0	10
10	0	10
20	0	

$$\binom{10}{10} \cdot \binom{10}{10} = 1 \qquad p_1 - p_2 = 0{,}0 \qquad \theta = \frac{0}{0}$$

(10,0, 9,1)

10	0	10
9	1	10
19	1	

$$\binom{10}{10} \cdot \binom{10}{9} = 10 \qquad p_1 - p_2 = 0{,}1 \qquad \theta = \infty$$

(9,1,10,0)

9	1	10
10	0	10
19	1	

$$\binom{10}{9} \cdot \binom{10}{10} = 10 \qquad p_1 - p_2 = -0{,}1 \qquad \theta = 0$$

(10,0, 8,2)

10	0	10
8	2	10
18	2	

$$\binom{10}{10}\cdot\binom{10}{8} = 45 \qquad p_1-p_2 = 0{,}2 \qquad \theta = \infty$$

(9,1, 9,1)

9	1	10
9	1	10
18	2	

$$\binom{10}{9}\cdot\binom{10}{9} = 100 \qquad p_1-p_2 = 0{,}0 \qquad \theta = 1$$

(8,2,10,0)

8	2	10
10	0	10
18	2	

$$\binom{10}{8}\cdot\binom{10}{10} = 45 \qquad p_1-p_2 = -0{,}2 \qquad \theta = 0$$

(10,0, 7,3)

10	0	10
7	3	10
17	3	

$$\binom{10}{10}\cdot\binom{10}{7} = 120 \qquad p_1-p_2 = 0{,}3 \qquad \theta = \infty$$

(9,1, 8,2)

9	1	10
8	2	10
17	3	

$$\binom{10}{9}\cdot\binom{10}{8} = 450 \qquad p_1-p_2 = 0{,}1 \qquad \theta = \frac{9}{4}$$

(8,2, 9,1)

8	2	10
9	1	10
17	3	

$$\binom{10}{8}\cdot\binom{10}{9} = 450 \qquad p_1-p_2=-0,1 \qquad \theta = \frac{4}{9}$$

(7,3,10,0)

7	3	10
10	0	10
17	3	

$$\binom{10}{7}\cdot\binom{10}{10} = 120 \qquad p_1-p_2=-0,3 \qquad \theta = 0$$

(10,0, 6,4)

10	0	10
6	4	10
16	4	

$$\binom{10}{10}\cdot\binom{10}{6} = 210 \qquad p_1-p_2= 0,4 \qquad \theta = \infty$$

(9,1, 7,3)

9	1	10
7	3	10
16	4	

$$\binom{10}{9}\cdot\binom{10}{7} = 1200 \qquad p_1-p_2= 0,2 \qquad \theta = \frac{27}{7}$$

(8,2, 8,2)

8	2	10
8	2	10
16	4	

$$\binom{10}{8}\cdot\binom{10}{8} = 2025 \qquad p_1-p_2= 0,0 \qquad \theta = 1$$

(7,3, 9,1)

7	3	10
9	1	10
16	4	

$$\binom{10}{7}\cdot\binom{10}{9} = 1200 \qquad p_1 - p_2 = -0,2 \qquad \theta = \frac{7}{27}$$

(6,4,10,0)

6	4	10
10	0	10
16	4	

$$\binom{10}{6}\cdot\binom{10}{10} = 210 \qquad p_1 - p_2 = -0,4 \qquad \theta = 0$$

(10,0, 5,5)

10	0	10
5	5	10
15	5	

$$\binom{10}{10}\cdot\binom{10}{5} = 252 \qquad p_1 - p_2 = 0,5 \qquad \theta = \infty$$

(9,1, 6,4)

9	1	10
6	4	10
15	5	

$$\binom{10}{9}\cdot\binom{10}{6} = 2100 \qquad p_1 - p_2 = 0,3 \qquad \theta = 6$$

(8,2, 7,3)

8	2	10
7	3	10
15	5	

$$\binom{10}{8}\cdot\binom{10}{7} = 5400 \qquad p_1 - p_2 = 0,1 \qquad \theta = \frac{12}{7}$$

(7,3, 8,2)

7	3	10
8	2	10
15	5	

$p_1-p_2=-0{,}1$

$$\binom{10}{7}\cdot\binom{10}{8} = 5400$$

$\theta = \frac{7}{12}$

(6,4, 9,1)

6	4	10
9	1	10
15	5	

$p_1-p_2=-0{,}3$

$$\binom{10}{6}\cdot\binom{10}{9} = 2100$$

$\theta = \frac{1}{6}$

(5,5,10,0)

5	5	10
10	0	10
15	5	

$p_1-p_2=-0{,}5$

$$\binom{10}{5}\cdot\binom{10}{10} = 252$$

$\theta = 0$

(10,0, 4,6)

10	0	10
4	6	10
14	6	

$p_1-p_2= 0{,}6$

$$\binom{10}{4}\cdot\binom{10}{6} = 210$$

$\theta = \infty$

(9,1, 5,5)

9	1	10
5	5	10
14	6	

$p_1-p_2= 0{,}4$

$$\binom{10}{9}\cdot\binom{10}{5} = 2520$$

$\theta = 9$

(8,2, 6,4)

8	2	10
6	4	10
14	6	

$$p_1 - p_2 = 0{,}2$$
$$\binom{10}{8}\cdot\binom{10}{6} = 9450$$
$$\theta = \frac{8}{3}$$

(7,3, 7,3)

7	3	10
7	3	10
14	6	

$$p_1 - p_2 = 0{,}0$$
$$\binom{10}{7}\cdot\binom{10}{7} = 14400$$
$$\theta = 1$$

(6,4, 8,2)

6	4	10
8	2	10
14	6	

$$p_1 - p_2 = -0{,}2$$
$$\binom{10}{6}\cdot\binom{10}{8} = 9450$$
$$\theta = \frac{3}{8}$$

(5,5, 9,1)

5	5	10
9	1	10
14	6	

$$p_1 - p_2 = -0{,}4$$
$$\binom{10}{5}\cdot\binom{10}{9} = 2520$$
$$\theta = \frac{1}{9}$$

(4,6,10,0)

4	6	10
10	0	10
14	6	

$$p_1 - p_2 = -0{,}6$$
$$\binom{10}{4}\cdot\binom{10}{10} = 210$$
$$\theta = 0$$

(10,0, 3,7)

10	0	10
3	7	10
13	7	

$$p_1-p_2 = 0{,}7$$
$$\left[\begin{matrix}10\\10\end{matrix}\right]\cdot\left[\begin{matrix}10\\3\end{matrix}\right] = 120$$
$$\theta = \infty$$

(9,1, 4,6)

9	1	10
4	6	10
13	7	

$$p_1-p_2 = 0{,}5$$
$$\left[\begin{matrix}10\\9\end{matrix}\right]\cdot\left[\begin{matrix}10\\4\end{matrix}\right] = 2100$$
$$\theta = \frac{27}{2}$$

(8,2, 5,5)

8	2	10
5	5	10
13	7	

$$p_1-p_2 = 0{,}3$$
$$\left[\begin{matrix}10\\8\end{matrix}\right]\cdot\left[\begin{matrix}10\\5\end{matrix}\right] = 11340$$
$$\theta = 4$$

(7,3, 6,4)

7	3	10
6	4	10
13	7	

$$p_1-p_2 = 0{,}1$$
$$\left[\begin{matrix}10\\7\end{matrix}\right]\cdot\left[\begin{matrix}10\\6\end{matrix}\right] = 25200$$
$$\theta = \frac{14}{9}$$

(6,4, 7,3)

6	4	10
7	3	10
13	7	

$$p_1-p_2 = -0{,}1$$
$$\left[\begin{matrix}10\\6\end{matrix}\right]\cdot\left[\begin{matrix}10\\7\end{matrix}\right] = 25200$$
$$\theta = \frac{9}{14}$$

(5,5, 8,2)

5	5	10
8	2	10
13	7	

$p_1 - p_2 = -0{,}3$

$$\binom{10}{5}\cdot\binom{10}{8} = 11340$$

$\theta = \frac{1}{4}$

(4,6, 9,1)

4	6	10
9	1	10
13	7	

$p_1 - p_2 = -0{,}5$

$$\binom{10}{4}\cdot\binom{10}{9} = 2100$$

$\theta = \frac{2}{27}$

(3,7,10,0)

3	7	10
10	0	10
13	7	

$p_1 - p_2 = -0{,}7$

$$\binom{10}{3}\cdot\binom{10}{10} = 120$$

$\theta = 0$

(10,0, 2,8)

10	0	10
2	8	10
12	8	

$p_1 - p_2 = 0{,}8$

$$\binom{10}{10}\cdot\binom{10}{2} = 45$$

$\theta = \infty$

(9,1, 3,7)

9	1	10
3	7	10
12	8	

$p_1 - p_2 = 0{,}6$

$$\binom{10}{9}\cdot\binom{10}{3} = 1200$$

$\theta = 21$

(8,2, 4,6)

8	2	10
4	6	10
12	8	

$$\binom{10}{8}\cdot\binom{10}{4} = 9450 \qquad p_1-p_2= 0{,}4 \qquad \theta = 6$$

(7,3, 5,5)

7	3	10
5	5	10
12	8	

$$\binom{10}{7}\cdot\binom{10}{5} = 30240 \qquad p_1-p_2= 0{,}2 \qquad \theta = \frac{7}{3}$$

(6,4, 6,4)

6	4	10
6	4	10
12	8	

$$\binom{10}{6}\cdot\binom{10}{6} = 44100 \qquad p_1-p_2= 0{,}0 \qquad \theta = 1$$

(5,5, 7,3)

5	5	10
7	3	10
12	8	

$$\binom{10}{5}\cdot\binom{10}{7} = 30240 \qquad p_1-p_2=-0{,}2 \qquad \theta = \frac{3}{7}$$

(4,6, 8,2)

4	6	10
8	2	10
12	8	

$$\binom{10}{4}\cdot\binom{10}{8} = 9450 \qquad p_1-p_2=-0{,}4 \qquad \theta = \frac{1}{6}$$

(3,7, 9,1)

3	7	10
9	1	10
12	8	

$$\binom{10}{3}\cdot\binom{10}{9} = 1200 \qquad p_1-p_2=-0,6 \qquad \theta = \frac{1}{21}$$

(2,8,10,0)

2	8	10
10	0	10
12	8	

$$\binom{10}{2}\cdot\binom{10}{10} = 45 \qquad p_1-p_2=-0,8 \qquad \theta = 0$$

(10,0, 1,9)

10	0	10
1	9	10
11	9	

$$\binom{10}{10}\cdot\binom{10}{1} = 10 \qquad p_1-p_2= 0,9 \qquad \theta = \infty$$

(9,1, 2,8)

9	1	10
2	8	10
11	9	

$$\binom{10}{9}\cdot\binom{10}{2} = 450 \qquad p_1-p_2= 0,7 \qquad \theta = 36$$

(8,2, 3,7)

8	2	10
3	7	10
11	9	

$$\binom{10}{8}\cdot\binom{10}{3} = 5400 \qquad p_1-p_2= 0,5 \qquad \theta = \frac{28}{3}$$

(7,3, 4,6)

7	3	10
4	6	10
11	9	

$p_1 - p_2 = 0,3$

$$\binom{10}{7}\cdot\binom{10}{4} = 25200$$

$\theta = \frac{7}{2}$

(6,4, 5,5)

6	4	10
5	5	10
11	9	

$p_1 - p_2 = 0,1$

$$\binom{10}{6}\cdot\binom{10}{5} = 52920$$

$\theta = \frac{3}{2}$

(5,5, 6,4)

5	5	10
6	4	10
11	9	

$p_1 - p_2 = -0,1$

$$\binom{10}{5}\cdot\binom{10}{6} = 52920$$

$\theta = \frac{2}{3}$

(4,6, 7,3)

4	6	10
7	3	10
11	9	

$p_1 - p_2 = -0,3$

$$\binom{10}{4}\cdot\binom{10}{7} = 25200$$

$\theta = \frac{2}{7}$

(3,7, 8,2)

3	7	10
8	2	10
11	9	

$p_1 - p_2 = -0,5$

$$\binom{10}{3}\cdot\binom{10}{8} = 5400$$

$\theta = \frac{3}{28}$

(2,8, 9,1)

2	8	10
9	1	10
11	9	

$$\binom{10}{2}\cdot\binom{10}{9} = 450 \qquad p_1 - p_2 = -0{,}7 \qquad \theta = \frac{1}{36}$$

(1,9,10,0)

1	9	10
10	0	10
11	9	

$$\binom{10}{1}\cdot\binom{10}{10} = 10 \qquad p_1 - p_2 = -0{,}9 \qquad \theta = 0$$

(10,0,0,10)

10	0	10
0	10	10
10	10	

$$\binom{10}{10}\cdot\binom{10}{0} = 1 \qquad p_1 - p_2 = 1{,}0 \qquad \theta = \infty$$

(9,1, 1,9)

9	1	10
1	9	10
10	10	

$$\binom{10}{9}\cdot\binom{10}{1} = 100 \qquad p_1 - p_2 = 0{,}8 \qquad \theta = 81$$

(8,2, 2,8)

8	2	10
2	8	10
10	10	

$$\binom{10}{8}\cdot\binom{10}{2} = 2025 \qquad p_1 - p_2 = 0{,}6 \qquad \theta = 16$$

(7,3, 3,7)

7	3	10
3	7	10
10	10	

$$p_1 - p_2 = 0{,}4$$
$$\binom{10}{7}\cdot\binom{10}{3} = 14400$$
$$\theta = \frac{49}{9}$$

(6,4, 4,6)

6	4	10
4	6	10
10	10	

$$p_1 - p_2 = 0{,}2$$
$$\binom{10}{6}\cdot\binom{10}{4} = 44100$$
$$\theta = \frac{9}{4}$$

(5,5, 5,5)

5	5	10
5	5	10
10	10	

$$p_1 - p_2 = 0{,}0$$
$$\binom{10}{5}\cdot\binom{10}{5} = 63504$$
$$\theta = 1$$

(4,6, 6,4)

4	6	10
6	4	10
10	10	

$$p_1 - p_2 = -0{,}2$$
$$\binom{10}{4}\cdot\binom{10}{6} = 44100$$
$$\theta = \frac{4}{9}$$

(3,7, 7,3)

3	7	10
7	3	10
10	10	

$$p_1 - p_2 = -0{,}4$$
$$\binom{10}{3}\cdot\binom{10}{7} = 14400$$
$$\theta = \frac{9}{49}$$

(2,8, 8,2)

2	8	10
8	2	10
10	10	

$$\binom{10}{2}\cdot\binom{10}{8} = 2025 \qquad p_1 - p_2 = -0{,}6 \qquad \theta = \frac{1}{16}$$

(1,9, 9,1)

1	9	10
9	1	10
10	10	

$$\binom{10}{1}\cdot\binom{10}{9} = 100 \qquad p_1 - p_2 = -0{,}8 \qquad \theta = \frac{1}{81}$$

(0,10,10,0)

0	10	10
10	0	10
10	10	

$$\binom{10}{0}\cdot\binom{10}{10} = 1 \qquad p_1 - p_2 = -1{,}0 \qquad \theta = 0$$

REFERENCIAS BIBLIOGRAFICAS

1. BARNARD, G. A. In contradiction to J. Berkson's dispraise: Condictional tests can be more efficient. *Journal of Statistical Planning and Inference*, v.3, p.181-7, 1979.
2. BERKSON, J. In dispraise of the exact test. *Journal of Statistical Planning and Inference*, v.2, p.27-42, 1978.
3. BERKSON, J. Do the marginal totals of the 2x2 table contain relevant information respecting the table proportions? *Journal of Statistical Planning and Inference*, v.2, p.43-4, 1978.-4, 1978.
4. BESKIN, N. M. *Dividing a segment in a given ratio*. Moscow: Mir Publishers, 1975.
5. COX, D. R. *The analysis of binary data*. London: Methuen, 1970.
6. CORSTEN, L. C. A. , KROON, J. P. M. Comment on J. Berkson's paper "In dispraise of the exact test". *Journal of Statistical Planning and Inference*, v.3, p.193-7, 1979.
7. EDWARDS, A. W. F. The measure of association in a 2x2 table. *Journal of the Royal Statistical Society*, Série A, v.126, p.109-14, 1963.
8. EDWARDS, A. W. F. Distances between populations on the basis of gene frequencies. *Biometrics*, v.27, p.873-81, 1971.
9. FIENBERG, S. E. , GILBERT, J. P. The geometry of a two by two

contingency table. *Journal of the American Statistical Association*, v.65, p.694-701, 1970.

10. FISHER, R. A. The logic of inductive inference. *Journal of the Royal Statistical Society* , Série A, v.98, p.39-54, 1935.

11. GART, J. J. , ZWEIFEL, J. R. On the bias of various estimators of the logit and its variance with application to quantal bioassay. *Biometrika*, v.54, p.181-7, 1967.

12. GOOD, I. J. , MITTAL, Y. The amalgamation and geometry of two-by-two contingency tables. *The Annals of Statistics*, v.15, p.694-711, 1987.

13. KALBFLEISCH, J. D. , SPROTT, D. A. Application of likelihood methods to models involving large numbers of parameters. *Journal of the Royal Statistical Society*, Série B, v.32, p.175-208, 1970.

14. KALBFLEISCH, J. D. , SPROTT, D. A. Marginal and conditional likelihoods. *Sankhyã*, Série A, v.35, p.311-28, 1973.

15. KEMPTHORNE, O. In dispraise of the exact test: Reactions. *Journal of Statistical Planning and Inference*, v.3, p.199-213, 1979.

16. MOSTELLER, F. Association and estimation in contingency tables. *Journal of the American Statistical Association*, v.63, p.1-28, 1968.

17. PLACKETT, R. L. The marginal totals of a 2x2 table. *Biometrika*, v.64, p.37-42, 1977.

18. THOMPSON, E. A. The likelihood for multinomial proportions under stereographic projection. *Biometrics*, p.618-20, 1972.

19. YATES, F. Tests of significance for 2x2 contingency tables. *Journal of the Royal Statistical Society*, Série A, v.147, p.426-63, 1984.