TABELA DE CONTINGÊNCIAS

Quadro que cruza duas variáveis qualitativas. O elemento genérico $n(i, j)$ dá o número de ocorrências simultâneas das modalidades de duas variáveis.

	1	…	j	…	p	*Soma*
1						
⋮						
i			$n(i, j)$			$n(i)$
⋮						
k						
Soma			$n(j)$			n

$$n = \sum_i \sum_j n(i, j) \qquad n(i) = \sum_j n(i, j) \qquad n(j) = \sum_i n(i, j)$$

Variáveis independentes

$$\frac{n(i, j)}{n(i)} = \frac{n(j)}{n} \xrightarrow{independencia} n'(i, j) = \frac{n(j)\, n(i)}{n}$$

Medida da independência

$$\chi^2 = \sum \frac{\left[n(i, j) - n'(i, j)\right]^2}{n'(i, j)}$$

ANÁLISE FACTORIAL DAS CORRESPONDÊNCIAS

OBJECTIVO

Método factorial de Análise de Dados desenvolvido no início da década de 70 para a análise e descrição de tabelas de contingências.

TABELA DE CONTINGÊNCIAS

Quadro que cruza duas variáveis qualitativas. O elemento genérico $n(i, j)$ dá o número de ocorrências simultâneas das modalidades de duas variáveis.

	1	$\ldots$	j	$\ldots$	p	*Soma*	
1							
$\vdots$							
i			$n(i, j)$			$n(i) = \sum_j n(i, j)$	*Frequências marginais*
$\vdots$							
k							
Soma			$n(j) = \sum_i n(i, j)$			$n = \sum_i \sum_j n(i, j)$	

Frequências marginais

ANÁLISE FACTORIAL DAS CORRESPONDÊNCIAS

OBJECTIVO

Método factorial de Análise de Dados desenvolvido no início da década de 70 para a análise e descrição de tabelas de contingências.

TABELA DE CONTINGÊNCIAS

Quadro que cruza duas variáveis qualitativas. O elemento genérico $n(i, j)$ dá o número de ocorrências simultâneas das modalidades de duas variáveis.

	1 … j … p	*Soma*	
1 ⋮ i ⋮ k	$f_{ij} = \dfrac{n(i,j)}{n}$	$f_j = \sum_i f_{ij} = \dfrac{n(j)}{n}$	*Frequências marginais*
Soma	$f_i = \sum_j f_{ij} = \dfrac{n(i)}{n}$	$1 = \sum_i \sum_j f_{ij}$	

Frequências marginais

ANÁLISE FACTORIAL DAS CORRESPONDÊNCIAS

NUVEM EM R^p

- Coordenada j da modalidade i (perfil f^i)

$$f_j^i = \frac{f_{ij}}{f_i} = \frac{n(i,j)}{n(i)} \qquad \forall\, j = 1,\, p$$

É a frequência condicional de ocorrência de j sabendo que ocorreu i.

$$\begin{array}{cc|ccccc|}
 & & 1 & \dots & j & \dots & p \\
\hline
 & 1 & & & & & \\
 & \vdots & & & & & \\
\textit{perfil } f^i & i & \dfrac{f_{i1}}{f_i} & \dots & \dfrac{f_{ij}}{f_i} & \dots & \dfrac{f_{ip}}{f_i} \\
 & \vdots & & & & & \\
 & k & & & & & \\
\hline
\end{array}$$

- Peso da modalidade i

$$f_i = \sum_j f_{ij} = \frac{n(i)}{n} \qquad \text{(frequência marginal de } i\text{)}$$

$$\sum_j \frac{f_{ij}}{f_i} = 1$$

ANÁLISE FACTORIAL DAS CORRESPONDÊNCIAS

NUVEM EM R^k

- Coordenada i da modalidade j (perfil f^j)

$$f_i^{\,j} = \frac{f_{ij}}{f_j} = \frac{n(i,j)}{n(j)} \qquad \forall\ i = 1,\ k$$

É a frequência condicional de ocorrência de i sabendo que ocorreu j.

perfil f^j

	1	…	j	…	p
1			$\frac{f_{1j}}{f_j}$		
⋮			⋮		
i			$\frac{f_{ij}}{f_j}$		
⋮			⋮		
k			$\frac{f_{nj}}{f_j}$		

- Peso da modalidade j

$$f_j = \sum_i f_{ij} = \frac{n(j)}{n} \qquad \text{(frequência marginal de } j\text{)}$$

$$\sum_i \frac{f_{ij}}{f_j} = 1$$

ANÁLISE FACTORIAL DAS CORRESPONDÊNCIAS

DISTÂNCIA DO χ^2

- Definição

$$d^2(i,i') = \sum_j \frac{1}{f_j}\left(\frac{f_{ij}}{f_i} - \frac{f_{i'j}}{f_{i'}}\right)^2 \qquad \text{em } R^p$$

$$d^2(j,j') = \sum_i \frac{1}{f_i}\left(\frac{f_{ij}}{f_j} - \frac{f_{ij'}}{f_{j'}}\right)^2 \qquad \text{em } R^k$$

- Propriedades

  - É equivalente a uma distância euclideana ponderada pelos pesos dados pelas frequências marginais.

  - Equivalência distribucional

    Dois pontos i_1 e i_2 com as mesmas coordenadas, podem ser substituídos por um único ponto i_0 com massa igual à soma das massas sem que se altere a estrutura de distâncias em R^p e R^k .

ANÁLISE FACTORIAL DAS CORRESPONDÊNCIAS

1ª TRANSFORMAÇÃO DE COORDENADAS

$$\frac{f_{ij}}{f_i} \Rightarrow \frac{f_{ij}}{f_i\sqrt{f_j}}$$

DISTÂNCIAS

$$d^2(i,i') = \overbrace{\sum_j \left(\frac{f_{ij}}{f_i\sqrt{f_j}} - \frac{f_{i'j}}{f_{i'}\sqrt{f_j}} \right)^2}^{\text{Distância euclideana}} = \overbrace{\sum_j \frac{1}{f_j} \left(\frac{f_{ij}}{f_i} - \frac{f_{i'j}}{f_{i'}} \right)^2}^{\text{Distância do } \chi^2}$$

CENTRO DE GRAVIDADE

$$g_j = \sum_i f_i \frac{f_{ij}}{f_i\sqrt{f_j}} = \frac{f_j}{\sqrt{f_j}} = \sqrt{f_j}$$

ANÁLISE FACTORIAL DAS CORRESPONDÊNCIAS

2ª TRANSFORMAÇÃO DE COORDENADAS

$$\frac{f_{ij}}{f_i\sqrt{f_j}} \quad \Rightarrow \quad \frac{f_{ij}}{f_i\sqrt{f_j}} - \sqrt{f_j}$$

DADOS TRANSFORMADOS

$$\mathbf{X}: \; x_{ij} = \frac{f_{ij}}{f_i\sqrt{f_j}} - \sqrt{f_j}$$

MATRIZ DE INÉRCIA

$$\mathbf{V} = \mathbf{X}' \; \mathbf{X}$$

$$v_{jj'} = \sum_i f_i \left(\frac{f_{ij}}{f_i\sqrt{f_j}} - \sqrt{f_j} \right) \left(\frac{f_{ij}}{f_i\sqrt{f_{j'}}} - \sqrt{f_{j'}} \right)$$

ANÁLISE FACTORIAL DAS CORRESPONDÊNCIAS

EIXOS FACTORIAIS

λ_α — *Valores próprios (inércia explicada)*

u_α — *Vectores próprios (direcção dos eixos factoriais)*

NOTAS:

- 1º valor próprio desta matriz de inércia é nulo.
- Existem *min* (p-1; k-1) valores próprios não nulos.

ANÁLISE FACTORIAL DAS CORRESPONDÊNCIAS

COORDENADAS DAS MODALIDADES NOS EIXOS

$$\psi_{i\alpha} = \sum_j \frac{f_{ij}}{f_i \sqrt{f_j}} u_{\alpha j}$$

$$\varphi_{j\alpha} = \sum_i \frac{f_{ij}}{f_j \sqrt{f_i}} v_{\alpha i}$$

RELAÇÕES DE TRANSIÇÃO

$$\varphi_{j\alpha} = \frac{1}{\sqrt{\lambda_\alpha}} \sum_i \frac{f_{ij}}{f_j} \psi_{\alpha i}$$

$$\psi_{i\alpha} = \frac{1}{\sqrt{\lambda_\alpha}} \sum_j \frac{f_{ij}}{f_i} \varphi_{\alpha j}$$

ANÁLISE FACTORIAL DAS CORRESPONDÊNCIAS

PROJECÇÃO EM SUPLEMENTAR

- Coluna suplementar

Perfil da coluna suplementar

$$f_i^{j^+} = \frac{f_{ij^+}}{f_{j^+}} = \frac{n(i, j^+)}{n(j^+)}$$

$$\varphi_{j^+\alpha} = \frac{1}{\sqrt{\lambda_\alpha}} \sum_i \frac{f_{ij^+}}{f_{j^+}} \psi_{\alpha i}$$

- Linha suplementar

Perfil da linha suplementar

$$f_j^{i^+} = \frac{f_{i^+j}}{f_{i^+}} = \frac{n(i^+, j)}{n(i^+)}$$

$$\psi_{i^+\alpha} = \frac{1}{\lambda_\alpha} \sum_j \frac{f_{i^+j}}{f_{i^+}} \varphi_{\alpha j}$$

AJUDAS À INTERPRETAÇÃO

Inércia explicada pelo eixo α

$$\lambda_\alpha = \sum_i f_i \ \psi_{i\alpha}^2 = \sum_j f_j \ \varphi_{j\alpha}^2$$

Contribuição absoluta

Medida da contribuição de cada modalidade para a construção de cada um dos eixos:

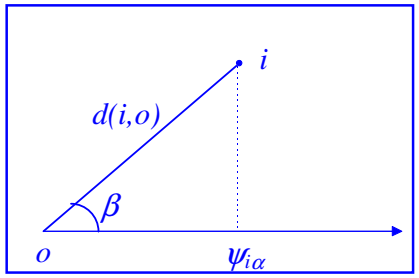
$$C_{i\alpha}^a = \frac{1}{\lambda_\alpha} f_i \ \psi_{i\alpha}^2 \qquad C_{j\alpha}^a = \frac{1}{\lambda_\alpha} f_j \ \varphi_{j\alpha}^2$$

Contribuição relativa

Medida da contribuição de cada eixo para a explicação de cada uma das modalidades.

$$d^2(i,o) = \sum_\alpha \psi_{i\alpha}^2 \qquad d^2(j,o) = \sum_\alpha \varphi_{j\alpha}^2$$

$$C_{i\alpha}^r = \frac{\psi_{i\alpha}^2}{d^2(i,0)} \qquad C_{j\alpha}^r = \frac{\varphi_{j\alpha}^2}{d^2(j,0)}$$



$$C_{i\alpha}^r = \cos^2(\beta)$$

$$\sum_\alpha C_{i\alpha}^r = 1$$

ANÁLISE FACTORIAL DAS CORRESPONDÊNCIAS

TIPOS DE QUADROS

- Tabelas de contingências
- Tabelas de medidas homogéneas
- Tabelas de notas
- Tabelas de preferência
- Tabelas de contingências justapostas

Análise Factorial das Correspondências Binárias

- Tabelas de variáveis qualitativas
  - Quadros disjuntivos completos
  - Quadros de Burt

Análise Factorial das Correspondências Múltiplas

ANÁLISE FACTORIAL DAS CORRESPONDÊNCIAS

QUADRO DE DADOS QUALITATIVOS

Variáveis / Amostras	Var 1 (3 modalidades)	Var 2 (2 modalidades)	Var 3 (3 modalidades)
1	2	1	1
2	3	1	2
3	3	2	1
4	2	1	3
5	1	1	2
6	3	2	3

QUADRO DISJUNTIVO COMPLETO

VARIÁVEIS	Var 1			Var 2		Var 3		
Modalidades / Amostras	1	2	3	1	2	1	2	3
1	0	1	0	1	0	1	0	0
2	0	0	1	1	0	0	1	0
3	0	0	1	0	1	1	0	0
4	0	1	0	1	0	0	0	1
5	1	0	0	1	0	0	1	0
6	0	0	1	0	1	0	0	1

ANÁLISE FACTORIAL DAS CORRESPONDÊNCIAS

QUADRO DISJUNTIVO COMPLETO

VARIÁVEIS	Var 1			Var 2		Var 3		
Modalidades / Amostras	1	2	3	1	2	1	2	3
1	0	1	0	1	0	1	0	0
2	0	0	1	1	0	0	1	0
3	0	0	1	0	1	1	0	0
4	0	1	0	1	0	0	0	1
5	1	0	0	1	0	0	1	0
6	0	0	1	0	1	0	0	1

QUADRO DE BURT

$B = Z' Z$

Variáveis		Var 1			Var 2		Var 3		
	Modalidades	1	2	3	1	2	1	2	3
Var 1	1	1	0	0	1	0	0	1	0
	2	0	2	0	2	0	1	0	1
	3	0	0	3	1	2	1	1	1
Var 2	1	1	2	1	4	0	1	2	1
	2	0	0	2	0	2	1	0	1
Var 3	1	0	1	1	1	1	2	0	0
	2	1	0	1	2	0	0	2	0
	3	0	1	1	1	1	0	0	2

COMPARAÇÃO DOS MÉTODOS FACTORIAIS

X

	I	II	III
1	4.00	16.00	16.00
2	1.00	4.00	13.00
3	2.00	8.00	14.00
4	1.00	4.00	13.00
	I	II	III

Y

	1	2	3	4
I	4.00	1.00	2.00	1.00
II	16.00	4.00	8.00	4.00
III	16.00	13.00	14.00	13.00
	1	2	3	4

ANÁLISE EM COMPONENTES PRINCIPAIS DE **X**

MATRIZ DAS CORRELACOES

	I	II	III
I	1.00	1.00	1.00
II	1.00	1.00	1.00
III	1.00	1.00	1.00
	I	II	III

COORDENADAS

	EIXO 1
I	1.000
II	1.000
III	1.000
	EIXO 1

ANÁLISE EM COMPONENTES PRINCIPAIS DE **Y**

MATRIZ DAS CORRELACOES

	1	2	3	4
1	1.00	0.69	0.87	0.69
2	0.69	1.00	0.96	1.00
3	0.87	0.96	1.00	0.96
4	0.69	1.00	0.96	1.00
	1	2	3	4

COORDENADAS

	EIXO 1	EIXO 2
I	-1.1402	0.2358
II	-0.0419	-0.4474
III	1.1822	0.2116
	EIXO 1	EIXO 2

ANÁLISE DAS CORRESPONDÊNCIAS DE **X**

MATRIZ DAS DISTÂNCIAS DO χ^2

	I	II	III
I	0.00	0.00	0.50
II	0.00	0.00	0.50
III	0.50	0.50	0.00
	I	II	III

COORDENADAS

	EIXO 1
I	0.2900
II	0.2900
III	-0.2100
	EIXO 1

ANÁLISE DAS CORRESPONDÊNCIAS DE **Y**

COORDENADAS

	EIXO 1
I	0.2900
II	0.2900
III	-0.2100
	EIXO 1