

G. Representação fatorial da informação contida em uma Tabela de Contingência

1. Propriedades dos fatores de uma nuvem de pontos-perfis

a) Coordenadas dos pontos-perfis

$$F_{\alpha}(i) = \sum_{j=1}^J z_{ij} u_{\alpha j} \quad e \quad G_{\alpha}(j) = \sum_{i=1}^I k_{ij} w_{\alpha i}$$

b) Coordenadas do Centro de Gravidade

As coordenadas do Centro de Gravidade nos espaços fatoriais são todas nulas.

$$\forall i \in I; \forall \alpha, \sum_{i=1}^I f_i F_{\alpha}(G_L) = 0 = \sum_{j=1}^J f_j G_{\alpha}(G_C), \forall j \in J; \forall \alpha$$

Os fatores são p variáveis (quantitativas, a valor em \mathbb{R}) de média nula. São variáveis «centradas».

c) Variância dos fatores

Sabemos que : $\bar{x}_{\alpha} = 0$, de modo que : $\forall \alpha = 1, \dots, p$

$$s_{\alpha}^2 = \sum_{i=1}^I p_i (x_{\alpha i} - \bar{x}_{\alpha})^2 = \sum_{i=1}^I p_i (F_{\alpha}(i) - 0)^2 = \sum_{i=1}^I p_i (F_{\alpha}(i))^2$$

A variância do fator α é igual à inércia de $N(I)$ ao longo deste eixo.

G. Representação fatorial da informação contida em uma Tabela de Contingência

1. Propriedades dos fatores de uma nuvem de pontos-perfis

A inércia ao longo de um eixo \rightarrow é igual ao auto-valor associado ao eixo...

$$s_{\alpha}^2 = \lambda_{\alpha} = \sum_{i=1}^I p_i (F_{\alpha}(i))^2, \quad \forall \alpha = 1, 2, \dots, p$$

d) Significado dos auto-valores de uma matriz de inércia

O auto-valor associado a um eixo fatorial representa a parte da relação observada entre as variáveis da Tabela T que é assumida pelo eixo.

e) Que significa um auto-valor $I_a = 1$...?

f) Quais são os auto-valores (e os eixos fatoriais) pertinentes de uma análise...?

g) Taxa de inércia de um fator

$$\tau_{\alpha} = \frac{\lambda_{\alpha}}{\sum_{\alpha=1}^p \lambda_{\alpha}} \times 100 = \frac{\lambda_{\alpha}}{I_G} \times 100$$

de modo que : $\sum_{\alpha=1}^p \tau_{\alpha} = 100\%$

A taxa τ_{α} do eixo α é calculada por softwares especiais.

G. Representação fatorial da informação contida em uma Tabela de Contingência

1. Propriedades dos fatores de uma nuvem de pontos-perfis

(continuação)

h) Taxa de inércia de um plano fatorial

$$\tau_{\text{plano}(i,i+k)} = \frac{\left(\lambda_{\alpha_i} + \lambda_{\alpha_{i+k}}\right)}{\sum_{\alpha=1}^p \lambda_{\alpha}} \times 100 =$$

Que significa a taxa de inércia de um eixo fatorial...?

l) A representação fatorial das nuvens de pontos N(I) e N(J) conserva a informação da Tabela T

$$d_{(i,i')} = \sqrt{\sum_{j=1}^3 \left(f_{ij}^* - f_{i'j}^*\right)^2} = \sqrt{\sum_{j=1}^3 \frac{1}{f_{.j}} \left(\frac{f_{ij}}{f_{i.}} - \frac{f_{i'j}}{f_{i'.}}\right)^2}$$

e

$$d_{(i,i')} = \sqrt{\sum_{\alpha=1}^2 \left(F_{\alpha}(i) - F_{\alpha}(i')\right)^2}$$

G. Representação fatorial da informação contida em uma Tabela de Contingência

2. Componentes digitais da mensagem

- A simples «leitura» de um plano fatorial não permite compreender — sem ambiguidade — o sentido de informação representada no plano.

a) Contribuição dos pontos-perfis para a inércia ao longo de um eixo fatorial

A contribuição do ponto i da inércia do eixo α é a seguinte :

$$CTR_{\alpha}(i) = \frac{p_i [F_{\alpha}(i)]^2}{\lambda_{\alpha}} \times 100 \quad \forall i \in N(I)$$

$$CTR_{\alpha}(j) = \frac{p_j [G_{\alpha}(j)]^2}{\lambda_{\alpha}} \times 100 \quad \forall j \in N(J)$$

O coeficiente $CTR_{\alpha}(i)$ (o $CTR_{\alpha}(j)$) é chamado «contribuição absoluta» para a inércia do eixo α .

G. Representação fatorial da informação contida em uma Tabela de Contingência

2. Componentes digitais da mensagem

(continuação)

- b) Contribuição relativa de um fator para a distância de um ponto-perfil ao Centro de Gravidade

Distância de um ponto-perfil a origem dos eixos fatoriais :

$$d_{(i,G)}^2 = \sum_{\alpha=1}^p [F_{\alpha}(i)]^2 \quad \forall i \in N(I)$$

$$d_{(j,G)}^2 = \sum_{\alpha=1}^p [G_{\alpha}(j)]^2 \quad \forall j \in N(J)$$

De modo que :

$$\text{COR}_{\alpha}(i) = \frac{[F_{\alpha}(i)]^2}{d_{(i,G)}^2} \quad \forall \alpha = 1, 2, \dots, p$$

$$\text{COR}_{\alpha}(j) = \frac{[G_{\alpha}(j)]^2}{d_{(j,G)}^2} \quad \forall \alpha = 1, 2, \dots, p$$

Com :

$$\sum_{\alpha=1}^p \text{COR}_{\alpha}(i) = 1 \quad \forall i \in N(I)$$

$$\sum_{\alpha=1}^p \text{COR}_{\alpha}(j) = 1 \quad \forall j \in N(J)$$

O coeficiente $\text{COR}_{\alpha}(i)$ ($\text{COR}_{\alpha}(j)$) é também chamado

$\text{COS}^2_{\alpha}(i)$ ($\text{COS}^2_{\alpha}(j)$).

G. Representação fatorial da informação contida em uma Tabela de Contingência

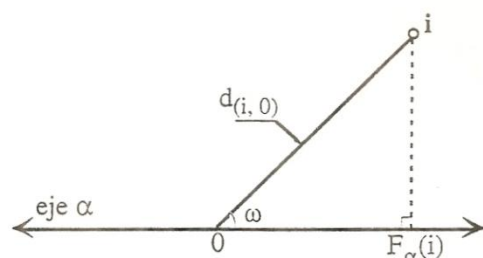
2. Componentes digitais da mensagem

(continuação)

b) Contribuição relativa de um fator a distância de um ponto-perfil ao Centro de Gravidade

- Que significa esse coeficiente...?

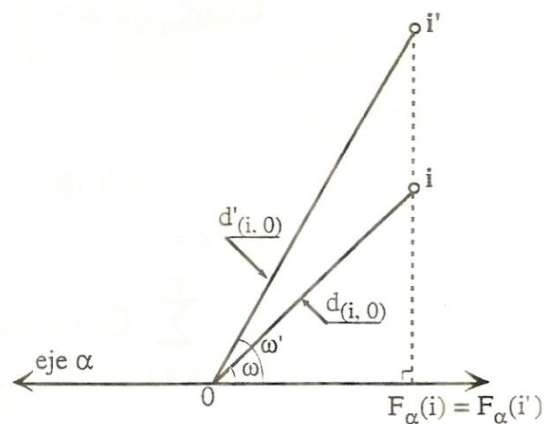
O quadrado do coseno do ângulo ω mede a «qualidade de representação» da distância do ponto i a origem do eixo α .



$$\cos^2(\omega) = \text{COR}_\alpha(i) = \frac{[F_\alpha(i)]^2}{d_{(i,0)}^2}$$

e

$$\cos^2(\omega') = \text{COR}_\alpha(i') = \frac{[F_\alpha(i')]^2}{d_{(i',0)}^2}$$



- Que significa «medir a qualidade de representação de um ponto» por um eixo fatorial...?

G. Representação fatorial da informação contida em uma Tabela de Contingência

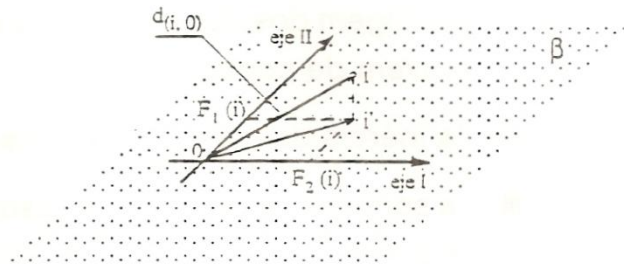
2. Componentes digitais da mensagem

(continuação)

- c) Contribuição relativa de um plano fatorial à distância de um ponto-perfil ao Centro de Gravidade :

Vemos que :

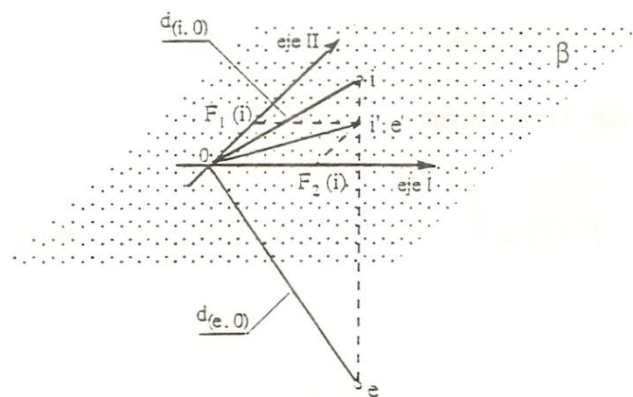
$$d_{(i,0)}^2 = \left(F_1^2(i) + F_2^2(i) \right)$$



de modo que a «qualidade de representação» sobre o plano fatorial é igual a :

$$\text{COR}_\beta(i) = \frac{\left(F_1^2(i) + F_2^2(i) \right)}{d_{(i,0)}^2}$$

Outro exemplo de controle de erros de projeção :



G. Representação fatorial da informação contida em uma Tabela de Contingência

3. Representação conjunta da inércia das nuvens de pontos $N(I)$ e $N(J)$

a) Justificativa da representação conjunta

- a origem dos dois espaços é o Centro de Gravidade de cada nuvem de pontos;
- a inércia das duas nuvens de pontos é a mesma;
- a soma dos auto-valores associados a cada matriz de inércia é a mesma;

$$I_{G_L}^{N(I)} = I_{G_C}^{N(J)} = \sum_{\alpha=1}^p \lambda_{\alpha}^{N(I)} = \sum_{\alpha=1}^q \lambda_{\alpha}^{N(J)} = \phi^2$$

- se pode demonstrar que os auto-valores são os mesmos e da mesma ordem;

$$\lambda_1^{N(I)} = \lambda_1^{N(J)} ; \lambda_2^{N(I)} = \lambda_2^{N(J)} ; \dots ; \lambda_{\alpha}^{N(I)} = \lambda_{\alpha}^{N(J)}$$

- ⊃ Estes espaços não são idênticos... mas são da mesma «natureza».
- ⊃ Eles podem ser superpostos ... porque existem fortes relações entre esses espaços.

b) Fórmulas de Transição :

$$F_{\alpha}(i) = \left(\frac{1}{\sqrt{\lambda_{\alpha}}} \right) \sum_{j=1}^J f_{i/C_j}^* G_{\alpha}(j) ; \quad \forall i \in I$$

$$G_{\alpha}(j) = \left(\frac{1}{\sqrt{\lambda_{\alpha}}} \right) \sum_{i=1}^I f_{j/L_i}^* F_{\alpha}(i) ; \quad \forall j \in J$$

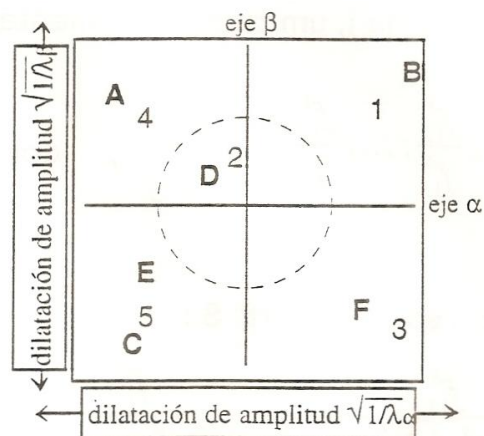
G. Representação fatorial da informação contida em uma Tabela de Contingência

3. Representação conjunta da inércia das nuvens de pontos $N(I)$ e $N(J)$ (continuação)

c) Significado das Fórmulas de Transição

- ⇒ é necessário compreender o significado destas relações porque a «leitura» dos planos fatoriais depende disto...

Exemplo :



- ⇒ As fórmulas de transição apresentam um grande interesse prático do ponto de vista do cálculo...

G. Representação fatorial da informação contida em uma Tabela de Contingência

4. Elementos suplementares ou ilustrativos de uma análise

a) Emprego de elementos suplementares...

(Ver documento anexo).

b) Coordenadas dos elementos suplementares

Seja S (em coluna), um perfil suplementar de J elementos :

$$S = \left[\frac{f_{11}^s}{f_{1.}^s}; \frac{f_{12}^s}{f_{1.}^s}; \dots; \frac{f_{1j}^s}{f_{1.}^s}; \dots; \frac{f_{1J}^s}{f_{1.}^s} \right], \quad \text{com } f_{1.}^s = \sum_{j=1}^J f_{1j}^s$$

Colocamos em escala o perfil S :

$$S = \left[\frac{f_{11}^s}{f_{1.}^s \sqrt{f_{.1}}}; \frac{f_{12}^s}{f_{1.}^s \sqrt{f_{.2}}}; \dots; \frac{f_{1j}^s}{f_{1.}^s \sqrt{f_{.j}}}; \dots; \frac{f_{1J}^s}{f_{1.}^s \sqrt{f_{.J}}} \right]$$

Centramos o perfil S em G_c :

$$S = \left[\frac{f_{11}^s}{f_{1.}^s \sqrt{f_{.1}}} - \sqrt{f_{.1}}; \dots; \frac{f_{1j}^s}{f_{1.}^s \sqrt{f_{.j}}} - \sqrt{f_{.j}}; \dots; \frac{f_{1J}^s}{f_{1.}^s \sqrt{f_{.J}}} - \sqrt{f_{.J}} \right]$$

G. Representação fatorial da informação contida em uma Tabela de Contingência

4. Elementos suplementares e ilustrativos de uma análise

(continuação)

c) Coordenadas do ponto-perfil S (ponderado e centrado em G_c) por projeção sobre os eixos fatoriais:

$$\text{Dado que : } G_\alpha(S) = \sum_{j=1}^J \left[\frac{f_{1j}^s}{f_{1.}^s \sqrt{f_{.j}}} - \sqrt{f_{.j}} \right] \cdot u_\alpha(j)$$

$$\text{sendo : } u_\alpha(j) = \sqrt{\frac{f_{.j}}{\lambda_\alpha}} G_\alpha(j)$$

$$\Rightarrow G_\alpha(S) = \sqrt{\frac{1}{\lambda_\alpha}} \sum_{j=1}^J \left[\frac{f_{1j}^s}{f_{1.}^s \sqrt{f_{.j}}} - f_{.j} \right] \cdot G_\alpha(j)$$

Fórmula muito semelhante às relações de transição .

G. Representação fatorial da informação contida em uma Tabela de Contingência

4. Elementos suplementares e ilustrativos de uma análise

d) Coeficiente de ajuda para interpretação dos elementos suplementares

- O coeficiente de «contribuição para inércia de um fator» não tem sentido no caso de um elemento suplementar, dado que são pontos de massa nula.

- Qualidade da representação sobre o eixo...(?)

- Valores teste... (ver documento anexo)

G. Representação fatorial da informação contida em uma Tabela de Contingência

5. Fórmula de reconstrução dos dados

$$f_{ij} = p_i p_j \left(1 + \sum_{\alpha} \frac{1}{\lambda_{\alpha}} F_{\alpha}(i) G_{\alpha}(j) \right) \quad \forall i \in I \text{ y } \forall j \in J$$

Desenvolvendo esta expressão para $\alpha = 1, 2, \dots, r \dots$

sendo : $f_{i.} = p_i$ e $f_{.j} = p_j \dots$

$$\alpha = 0; f_{ij} = p_i p_j$$

$$\alpha = 1; f_{ij} = p_i p_j + p_i p_j \frac{1}{\sqrt{\lambda_1}} F_1(i) G_1(j)$$

$$\alpha = 2; f_{ij} = p_i p_j + p_i p_j \frac{1}{\sqrt{\lambda_1}} F_1(i) G_1(j) + p_i p_j \frac{1}{\sqrt{\lambda_2}} F_2(i) G_2(j)$$

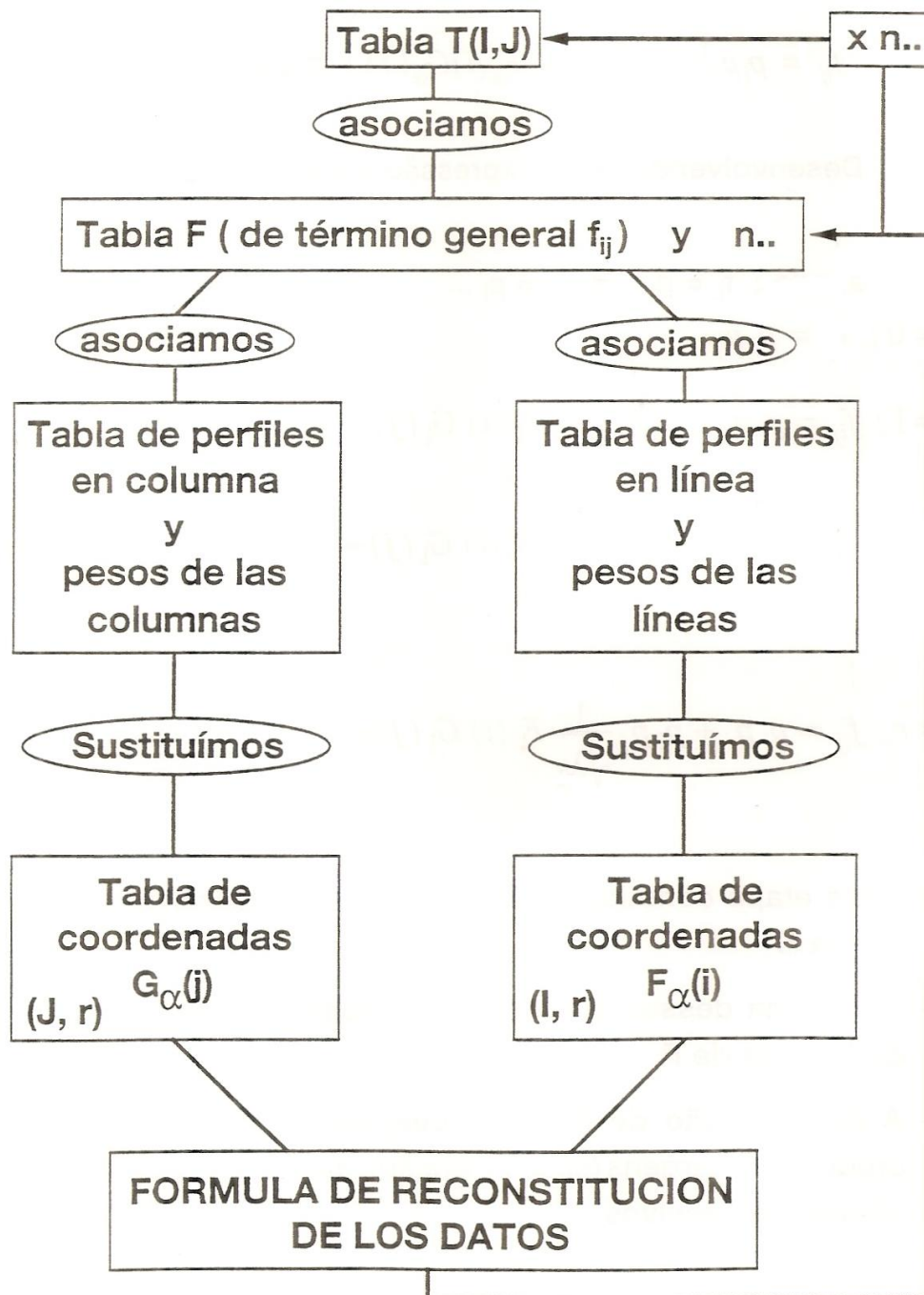
⋮

$$\alpha = r; f_{ij} = p_i p_j + p_i p_j \frac{1}{\sqrt{\lambda_1}} F_1(i) G_1(j) + \dots + p_i p_j \frac{1}{\sqrt{\lambda_r}} F_r(i) G_r(j)$$

- Cada etapa constitui uma aproximação sucessiva da tabela F.
- Cada uma dessas etapas é uma tabela das mesmas dimensões de F.
- A reconstrução de F se faz pela soma das tabelas produto de margens (uma dimensão), dadas pelos eixos fatoriais sucessivos.

G. Representação fatorial da informação contida em uma Tabela de Contingência

5. Fórmula de reconstrução de dados (continuação)



H. Exemplos de aplicação da Análise Fatorial de Correspondência Simples

1. Análise da mobilidade dos trabalhadores belgas «migrantes alternantes»

☞ A mobilidade geográfica da população ativa constitui um problema sócio-econômico de grande importância.

☞ A população ativa que exerce sua profissão fora do município de residência é chamada «migrante alternante».

Importância da população ativa belga
«migrante alternante» em 1981

Lugar de Trabajo	Año 1981	
		%
1. En el municipio de residencia	1.738.423	50,7
a) Trabajadores a domicilio	458.441	13,4%
b) Trabajadores que ejercen su profesión en el municipio de residencia	1.279.441	37,3%
2. Fuera del municipio de residencia	1.643.730	48,0%
a) En Belgica:		
- en el mismo distrito	783.812	22,9%
- en otro distrito	793.086	23,1%
b) En el extranjero	66.832	2,0%
3. Lugar de trabajo variable	44.371	1,3%
Total	3.426.524	100%
4. Lugar de trabajo desconocido	92.546	
Total general	3.519.000	

☞ Análise do meio de transporte até o local de trabalho, utilizado pela população ativa belga que se desloca cotidianamente.

H. Exemplos de aplicação da Análise Fatorial de Correspondência Simples

2. Estudo do meio de transporte empregado pela população ativa migrante alternante.

☞ Análise do meio de transporte empregado pela população ativa belga que exerce sua profissão fora do município de residência e se desloca pelo menos quatro dias por semana(*).

Distribuição da população segundo o meio de transporte empregado em 1981

Medio Principal de Transporte empleado por la población migrante alternante	% en 1981
Auto, como conductor	48,11%
Bicicleta	11,29%
Caminando	9,90%
Transp.Colectivo	9,68%
Auto, como acompañante	8,23%
Tren	7,89%
Transp. org. por la empresa	3,90%
Moto, motoneta	0,74%
Medio de transp. desc.	0,26%
Total:	100.00%

* Comentários :

- +/- 50 % da população ativa que trabalha fora do seu município de residência, empregava carro próprio, como meio de transporte até seu local de trabalho.
- um trabalhador sobre dez utilizava o transporte coletivo (ônibus, «bonde», metrô),
- mais de um quinto da população de migrantes alternantes se deslocava para o trabalho a pé ou de bicicleta.

(*) Se excluye las personas que, en 1981: trabajaban a domicilio; trabajaban al extranjero; realizaban un número desconocido de desplazamientos semanales; o aún aquellas personas cuyo lugar de trabajo es variable o desconocido.

H. Exemplos de aplicação da Análise Fatorial de Correspondência Simples

3. Objetivos do estudo e apresentação dos dados

População Belga Migrante Alternante por Grupo Profissional e Meio Principal de Transporte empregado para trabalhar (em 1981)

		T ₁	T ₂	T ₃	T ₄	T ₅	T ₆	T ₇	T ₈	T ₉	T
Arin	Arquitectos, Agnmensores, Ingenieros y asimilados	11%	2%	3%	72%	4%	0%	3%	5%	0%	100%
ChSc	Quimicos, físicos y otros científicos	8%	2%	7%	70%	5%	1%	4%	3%	0%	100%
BiAg	Biólogos, Agronomos, veterinarios y asimil.	6%	0%	4%	78%	3%	0%	2%	7%	0%	100%
Mein	Médicos, Dentistas, Personal paramédico y asimilados	3%	1%	8%	56%	9%	0%	10%	12%	0%	100%
Ense	Personal Docente	4%	1%	8%	64%	4%	0%	7%	13%	0%	100%
Micu	Ministros de Culto	1%	0%	2%	39%	3%	0%	7%	47%	1%	100%
ProJ	Profesiones Jurídicas	7%	1%	4%	66%	8%	0%	2%	12%	0%	100%
Arti	Artistas, Periodistas, Escritores y asimilados	10%	1%	6%	58%	12%	0%	2%	10%	1%	100%
DiTe	Dibujantes, Técnicos (de nivel secundario)	11%	3%	8%	61%	7%	1%	6%	5%	0%	100%
AuPr	Otras prof. liberales	12%	1%	7%	59%	10%	0%	4%	7%	0%	100%
JeEm	Jefes de Empresas del sect. privado	0%	1%	3%	76%	1%	0%	2%	17%	0%	100%
CuPr	Direct. y Cuadros Sup. sect. privado	6%	1%	5%	74%	5%	0%	3%	6%	0%	100%
CuFu	Direct. y Cuadros Sup. adm. publica	21%	1%	6%	50%	10%	1%	4%	3%	0%	100%
CaDa	Asist. Contables, Cajeros, Dactilógrafas	13%	1%	14%	36%	19%	0%	6%	10%	0%	100%
OtEm	Otros Empleados de Oficina	17%	1%	10%	43%	14%	0%	6%	8%	0%	100%
Coma	Comerciantes	1%	1%	10%	56%	3%	0%	3%	25%	0%	100%
Segu	Agentes de seguro, inmobiliarios y asimil.	5%	1%	5%	71%	6%	0%	2%	11%	1%	100%
Viaj	Viajantes de comercio	2%	2%	2%	87%	2%	0%	1%	4%	1%	100%
Vend	Vendedores y trabajadores asimilados	4%	1%	10%	35%	21%	0%	11%	16%	0%	100%
agri	Dirigentes y Jefes de expl. agrícola	0%	0%	0%	63%	0%	0%	19%	19%	0%	100%
Pays	Agricultores y asimilados	0%	0%	4%	40%	2%	0%	13%	40%	1%	100%
TrAg	Trabajadores agrícolas y asimilados	3%	2%	5%	41%	6%	2%	29%	12%	0%	100%
Pesc	Pescadores y asimilados	1%	3%	23%	35%	3%	1%	15%	20%	0%	100%
Fore	Trabajadores forestales	0%	4%	11%	65%	1%	2%	11%	6%	0%	100%
Mine	Mineros y asimilados	1%	18%	5%	38%	11%	1%	18%	8%	0%	100%
Coma	Comandantes y oficiales navales	6%	8%	5%	51%	7%	1%	9%	12%	1%	100%
Mari	Marineros y mecánicos navales	4%	7%	5%	38%	5%	1%	20%	20%	0%	100%
Pilo	Pilotos de avion y navegantes	4%	1%	3%	82%	2%	0%	1%	7%	1%	100%
Cond	Conductores y choferes (tren y transp. ruta)	4%	5%	2%	65%	4%	1%	10%	8%	0%	100%
Insp	Inspectores y Controladores (sect. del transporte)	30%	1%	1%	40%	8%	2%	8%	9%	0%	100%
Diri	Despachantes, Dirigentes de la circulación y asimilados	30%	1%	2%	39%	8%	2%	10%	7%	0%	100%
Tele	Telefonistas, telegrafistas y asimilados	17%	1%	5%	35%	12%	1%	21%	9%	0%	100%
CoMa	Contramaestres, Jefes de Equipo y asimil.	5%	7%	5%	63%	4%	1%	3%	6%	0%	100%
Teje	Tejedores, Tintoreros y asimilados	2%	5%	11%	29%	6%	1%	38%	8%	0%	100%
Sast	Sastres, Cortadores y asimilados	5%	4%	16%	21%	14%	0%	30%	9%	0%	100%
Cord	Obreros del Cuero	4%	4%	11%	30%	12%	1%	25%	13%	0%	100%
ObMe	Laminadores, Moideadores y asimilados	4%	8%	6%	48%	10%	2%	14%	7%	0%	100%
Orte	Mecánicos de precisión, relojeros y asimil.	7%	1%	11%	39%	9%	0%	20%	13%	0%	100%
Ajus	Soldadores, Plomeros, Ajustadores y asimil.	5%	7%	8%	51%	7%	1%	13%	6%	0%	100%
Elec	Electricistas y asimilados	10%	7%	9%	49%	9%	1%	9%	5%	0%	100%
TrMa	Trabajadores de la madera	4%	9%	8%	45%	4%	1%	22%	7%	0%	100%
Plnt	Pintores, Coladores y asimilados	5%	6%	7%	51%	8%	1%	14%	8%	0%	100%
Alba	Albañiles y obreros de la Construcción	7%	21%	8%	38%	4%	1%	14%	6%	0%	100%
Impr	Obreros de imprenta, de la edición y asimil.	10%	1%	7%	45%	12%	1%	12%	10%	0%	100%
TrVi	Trabajadores del vidrio y de la cerámica	4%	4%	9%	44%	9%	2%	19%	9%	0%	100%
TrAl	Trabajadores ind. alimentos y bebidas	3%	3%	7%	43%	7%	1%	21%	13%	0%	100%
TrOu	Obreros de la química y del papel	3%	9%	10%	46%	7%	1%	18%	6%	0%	100%
TrTa	Obreros del tabaco	5%	14%	19%	25%	12%	0%	20%	5%	0%	100%
ONa-	Obreros (sect. no especificado)	4%	8%	10%	44%	8%	1%	20%	7%	0%	100%
Etau	Embaladores y Etiquetadores	5%	7%	18%	22%	13%	1%	25%	9%	0%	100%
MaPe	Conductores maquinarias de excavación, gruas, etc	3%	16%	6%	52%	4%	1%	12%	5%	0%	100%
Dock	Dockers y asimilados	5%	4%	7%	42%	11%	2%	18%	11%	0%	100%
Polii	Policías, Bomberos y Guardianes	5%	1%	3%	58%	7%	1%	12%	13%	0%	100%
TrDo	Trabajadores Domesticos, de la Hostelería y asimilados	4%	1%	10%	22%	21%	1%	19%	21%	0%	100%
TrSP	Trabajadores Servicios Person.	4%	1%	11%	28%	19%	1%	15%	20%	0%	100%
Atle	Atletas y Deportistas	3%	1%	6%	63%	6%	1%	9%	12%	0%	100%
OtSe	Otras profesiones de servicios	8%	1%	9%	46%	18%	0%	5%	13%	0%	100%
Mili	Militares de carrera	14%	2%	6%	51%	6%	1%	6%	13%	0%	100%
PrDe	Profesion mal definida	7%	9%	9%	26%	24%	1%	9%	13%	1%	100%
Población activa migrante alternante		8%	4%	8%	48%	10%	1%	11%	10%	> 1%	100%

T ₁ : TREN	T ₂ : TRANSPORTE ORG. POR LA EMPRESA
T ₃ : AUTO-c/PASAJERO	T ₄ : AUTO-c/CONDUCTOR
T ₅ : OMNIBUS-TRANVIA-SUBTERRANEO	T ₆ : MOTO-MOTONETA
T ₇ : BICICLETA	T ₈ : UNICAMENTE CAMINANDO
T ₉ : MEDIO DE TRANSP. DESCON.	

Fuente : I.N.S., Etudes Statistiques, N°86, Bruxelles, 1989.

H. Exemplos de aplicação da Análise Fatorial de Correspondência Simples

4. Análise Fatorial de Correspondência da Tabela T(59 x 9)

a) Decomposição da inércia da Tabela T(59 x 9)

- ❶ Inércia global da Tabela T(59 x 9) = 0,2364.
- ❷ Decomposta ao longo de 8 direções principais de alargamento das nuvens N (I) e N (J).

Quais são os eixos fatoriais «interessantes» nesta análise?

VERIFICACION DE LA PRECISION DE LOS CALCULOS : TRAZA DE LA MATRIZ ANTES DE LA DIAGONALIZACION 0.2364
SUMA DE LOS VALORES PROPIOS 0.2364

HISTOGRAMA DE LOS 8 PRIMEROS VALORES PROPIOS

NUMERO	VALOR PROPIO	PORCENT.	PORCENT. ACUMULA.
1	0.0829	35.07	35.07
2	0.0770	32.57	67.64
3	0.0452	19.11	86.76
4	0.0169	7.17	93.92
5	0.0079	3.35	97.27
6	0.0053	2.24	99.51
7	0.0009	0.39	99.89
8	0.0003	0.11	100.00

- Os três primeiros auto-valores são superiores ao «valor médio» ($0,2364 / 8 = 0,0296$).
- A partir do quarto valor próprio : o decréscimo dos mesmos é regular; o valor absoluto dos 5 últimos auto-valores é muito pequeno.
- Retemos, para ser interpretados, os três primeiros eixos fatoriais.
- O primeiro e o segundo eixo fatorial apresentam taxas de inércia muito semelhantes \Rightarrow nos interessamos pelo significado do primeiro plano fatorial..

H. Exemplos de aplicação da Análise Fatorial de Correspondência Simples

4. Análise Fatorial de Correspondência da Tabela T(59 x 9)

b) Interpretação dos primeiros eixos fatoriais

- Estatísticas de ajuda para a interpretação do primeiro eixo fatorial: Coordenadas, Contribuições e Qualidade de Representação dos pontos-coluna da Tabela T(59 x 9)

FRECUENCIAS			COORDENADAS					CONTRIBUCIONES					CALIDAD DE REPR.				
IDEN	ETIQUETA	P.REL DISTO	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
TRAI	TREN	7.89 0.49	-0.28	-0.32	0.53	-0.11	0.14	7.6	10.6	48.4	5.6	18.4	0.16	0.21	0.56	0.02	0.04
TREM	TRANSP.ORG. POR EMP.	3.90 1.38	0.68	-0.76	0.48	0.31	0.05	21.9	29.6	19.6	22.6	1.2	0.34	0.42	0.17	0.07	0.00
AUPA	AUTO c/PASAJERO	8.23 0.12	0.13	-0.19	0.06	0.04	-0.15	1.7	4.1	0.6	0.8	22.0	0.15	0.32	0.03	0.01	0.18
AUCO	AUTO c/CONDUCTOR	48.11 0.08	-0.20	0.17	-0.06	-0.02	-0.02	24.1	19.0	3.8	1.0	2.5	0.54	0.40	0.05	0.00	0.01
OTRS	OMNIBUS-TRANVIA-SUBTE.	9.68 0.29	0.13	-0.47	0.08	0.12	-0.12	1.9	28.1	1.2	8.9	17.5	0.06	0.76	0.02	0.05	0.05
MOSC	MOTO-MOTONETA	0.74 0.36	0.35	0.19	0.16	-0.15	0.12	1.1	0.4	0.4	0.9	1.4	0.24	0.10	0.07	0.06	0.04
BICI	BICICLETA	11.29 0.38	0.55	-0.02	-0.11	-0.25	0.02	41.3	0.1	3.1	40.0	0.8	0.80	0.00	0.03	0.16	0.00
APIE	UNIC.CAMINANDO	9.90 0.23	0.06	-0.25	-0.32	0.18	0.17	0.4	8.1	22.9	19.3	36.0	0.02	0.27	0.44	0.14	0.12
SINI	MEDIO DE TRANSP.DESC.	0.26 0.20	-0.06	-0.14	-0.04	0.24	0.08	0.0	0.1	0.0	0.9	0.2	0.02	0.09	0.01	0.29	0.03

* No semi-eixo negativo :

☞ Os pontos «auto c/conductor» e «trem» apresentam :

- coordenadas importantes;
- boa contribuição da inércia ao longo do eixo;
- boa qualidade de representação.
- No semi-eixo positivo :

☞ salvo o ponto «unic. caminando», todos os outros meios de transporte apresentam :

- uma coordenada suficiente importante;
- apenas os pontos «transp. da empresa» e «bicicleta» são pontos-perfis fortemente contributivos para a inércia ao longo desse eixo;
- os pontos «transporte coletivo» e «unic.caminando» apresentam uma qualidade de representação muito baixa sobre esse eixo.

H. Exemplos de aplicação da Análise Fatorial de Correspondência Simples

4. Análise Fatorial de Correspondência da Tabela T(59 x 9)

b) Interpretação dos primeiros eixos fatoriais

Coordenadas, Contribuições e Qualidade de Representação dos pontos-linha da Tabela T(59 x 9)

INDIVIDUOS		COORDENADAS					CONTRIBUCIONES					CALIDAD DE REPR.						
IDEN	ETIQUETA	P.REL	DISTO	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
ArIn	Arquitectos. Ingeni	1.05	0.29	-0.48	0.23	0.04	-0.09	0.02	2.9	0.7	0.0	0.5	0.0	0.77	0.18	0.01	0.03	0.00
ChSc	Químicos. físicos y	0.06	0.22	-0.40	0.21	-0.01	-0.07	-0.08	0.1	0.0	0.0	0.0	0.0	0.75	0.20	0.00	0.02	0.03
BiAg	Biólogos. Agrónomos.	0.04	0.38	-0.51	0.29	-0.17	-0.05	-0.04	0.1	0.0	0.0	0.0	0.0	0.68	0.22	0.08	0.01	0.00
Meln	Médicos.Dent.. Pers	3.74	0.08	-0.13	0.00	-0.24	0.00	-0.07	0.7	0.0	4.7	0.0	2.1	0.21	0.00	0.72	0.00	0.06
Ense	Personal Docente	7.69	0.15	-0.26	0.13	-0.25	0.01	-0.01	6.1	1.6	11.0	0.1	0.0	0.43	0.11	0.42	0.00	0.00
Micu	Ministros de Culto	0.06	1.60	-0.03	-0.25	-0.81	0.50	0.78	0.0	0.0	0.9	0.9	4.8	0.00	0.04	0.41	0.16	0.38
ProJ	Profesiones Jurídic	0.30	0.21	-0.40	0.08	-0.16	0.08	0.03	0.6	0.0	0.2	0.1	0.0	0.78	0.03	0.12	0.03	0.00
Arti	Artistas. Period..	0.35	0.13	-0.33	-0.07	-0.01	0.10	-0.03	0.5	0.0	0.0	0.2	0.0	0.82	0.03	0.00	0.07	0.00
DiTe	Técnicos (nivel sec	2.48	0.11	-0.29	0.12	0.10	-0.07	-0.01	2.5	0.5	0.6	0.8	0.1	0.72	0.14	0.09	0.05	0.00
AuPr	Otras prof.liberales	2.14	0.13	-0.35	-0.03	0.08	-0.01	-0.04	3.1	0.0	0.3	0.0	0.4	0.93	0.01	0.05	0.00	0.01
JeEm	Jefes Empr. sect. p	0.43	0.50	-0.42	0.32	-0.43	0.14	0.11	0.9	0.6	1.8	0.5	0.7	0.36	0.20	0.37	0.04	0.03
CuFu	Direct./Cuadros Sup	3.51	0.29	-0.44	0.26	-0.11	-0.02	-0.09	8.3	3.1	0.9	0.1	3.5	0.69	0.23	0.04	0.00	0.03
CuFu	Direct./Cuadros Sup	1.07	0.30	-0.38	-0.20	0.29	-0.07	0.19	1.8	0.5	2.0	0.3	4.7	0.46	0.13	0.28	0.01	0.11
CaDa	Asist. Cont.. Cajer	2.20	0.22	-0.05	-0.40	0.16	0.12	-0.15	0.1	4.5	1.3	1.7	6.4	0.01	0.69	0.12	0.06	0.10
OtEm	Otros Empl. de Ofic	16.99	0.17	-0.18	-0.28	0.24	-0.01	0.00	6.9	17.1	21.3	0.1	0.1	0.20	0.46	0.34	0.00	0.00
Come	Comerciantes	1.75	0.45	-0.21	0.00	-0.49	0.29	0.18	0.9	0.0	9.3	8.6	7.1	0.10	0.00	0.53	0.18	0.07
Segu	Agentes Seguro. Inm	0.29	0.26	-0.43	0.17	-0.19	0.07	-0.02	0.7	0.1	0.2	0.1	0.0	0.72	0.11	0.14	0.02	0.00
Viaj	Viajantes de comerc	0.45	0.62	-0.54	0.49	-0.20	-0.02	-0.15	1.6	1.4	0.4	0.0	1.2	0.48	0.39	0.07	0.00	0.03
Vend	Vendedores y asimil	2.69	0.26	0.12	-0.40	-0.17	0.19	-0.12	0.5	5.5	1.7	5.6	4.5	0.06	0.60	0.11	0.14	0.05
Agri	Dirig./Jefes de exp	0.00	0.48	-0.04	0.21	-0.56	-0.18	0.27	0.0	0.0	0.0	0.0	0.0	0.00	0.09	0.65	0.07	0.15
Pays	Agricultores y asim	0.15	1.14	0.08	-0.17	-0.75	0.30	0.65	0.0	0.1	1.9	0.8	8.1	0.01	0.03	0.49	0.08	0.37
TrAg	Trabaj.agricolas y	0.39	0.35	0.37	0.01	-0.27	-0.35	0.13	0.7	0.0	0.6	2.9	0.8	0.39	0.00	0.20	0.35	0.04
Pesc	Pescadores y asimil	0.02	0.52	0.27	-0.11	-0.31	0.09	-0.03	0.0	0.0	0.1	0.0	0.0	0.14	0.02	0.19	0.02	0.00
Fore	Trabajadores forest	0.06	0.25	-0.07	0.37	-0.18	-0.09	-0.14	0.0	0.1	0.0	0.0	0.1	0.02	0.55	0.13	0.03	0.08
Mine	Mineros y asimilado	0.76	0.64	0.59	0.42	0.18	0.25	0.02	3.2	1.7	0.6	2.9	0.0	0.53	0.27	0.05	0.10	0.00
Coma	Comandantes/oficial	0.09	0.09	0.03	0.19	-0.01	0.15	0.10	0.0	0.0	0.0	0.1	0.1	0.01	0.40	0.00	0.26	0.11
Mari	Marineros/mecánicos	0.07	0.26	0.32	0.06	-0.21	0.04	0.32	0.1	0.0	0.1	0.0	0.9	0.41	0.02	0.17	0.01	0.39
Pilo	Pilotos de avion/na	0.02	0.49	-0.54	0.38	-0.22	-0.02	-0.06	0.1	0.0	0.0	0.0	0.0	0.59	0.29	0.10	0.00	0.01
Cond	Conduct./choferes(tr	3.35	0.16	-0.15	0.32	-0.13	-0.04	0.04	0.9	4.6	1.3	0.4	0.6	0.14	0.66	0.11	0.01	0.01
Insp	Inspect./Controlad.	0.17	0.74	-0.31	-0.29	0.51	-0.25	0.47	0.2	0.2	1.0	0.6	4.7	0.13	0.12	0.36	0.09	0.30
Diri	Despach.Dirig.cir	0.56	0.75	-0.26	-0.28	0.56	-0.31	0.43	0.5	0.6	3.9	3.1	12.8	0.09	0.10	0.42	0.13	0.24
Tele	Telefonistas. teleg	1.19	0.26	0.11	-0.28	0.16	-0.31	0.20	0.2	1.2	0.7	6.8	5.8	0.05	0.30	0.09	0.37	0.15
CoMa	Contramaestres. Jef	2.09	0.16	-0.09	0.39	0.00	0.00	-0.02	0.2	4.0	0.0	0.0	0.1	0.05	0.93	0.00	0.00	0.00
Teje	Tejedores. Tintorer	1.34	0.81	0.73	0.03	-0.20	-0.47	0.00	8.6	0.0	1.2	17.5	0.0	0.66	0.00	0.05	0.27	0.00
Sast	Sastres. Cortadores	1.73	0.57	0.64	-0.26	-0.04	-0.21	-0.15	8.7	1.5	0.1	4.6	5.0	0.73	0.11	0.00	0.08	0.04
Cord	Obreros del Cuero	0.14	0.26	0.45	-0.16	-0.13	-0.13	0.00	0.3	0.0	0.0	0.1	0.0	0.77	0.09	0.06	0.06	0.00
ObMe	Laminadores. Moldea	1.20	0.10	0.18	0.20	0.03	0.03	-0.02	0.5	0.6	0.0	0.1	0.1	0.34	0.40	0.01	0.01	0.01
Orfe	Mecánicos de precis	0.48	0.14	0.19	-0.17	-0.16	-0.17	0.04	0.2	0.2	0.3	0.8	0.1	0.27	0.22	0.19	0.22	0.01
Ajus	Soldad.. Plomeros.	9.04	0.07	0.10	0.22	0.05	-0.02	-0.06	1.1	5.6	0.5	0.3	4.2	0.16	0.72	0.04	0.01	0.06
Eiec	Electricistas y asi	3.11	0.07	0.00	0.14	0.21	0.03	-0.04	0.0	0.8	3.0	0.2	0.6	0.00	0.29	0.66	0.02	0.02
TrMa	Trabajadores de la	2.26	0.25	0.36	0.31	0.00	-0.15	0.04	3.6	2.7	0.0	2.8	0.4	0.53	0.37	0.00	0.09	0.01
Pint	Pintores. Coladores	0.64	0.04	0.09	0.15	-0.02	-0.04	-0.03	0.1	0.2	0.0	0.1	0.1	0.21	0.62	0.01	0.05	0.03
Alba	Albañiles/obreros	3.48	0.82	0.51	0.55	0.42	0.27	0.10	10.7	13.7	13.3	15.2	4.6	0.31	0.37	0.21	0.09	0.01
Impr	Obreros de imprenta	0.77	0.03	-0.03	-0.15	0.00	-0.07	0.01	0.0	0.2	0.0	0.3	0.0	0.02	0.68	0.00	0.16	0.01
TrVi	Trabaj. del vidrio	0.46	0.10	0.23	0.05	-0.10	-0.13	-0.04	0.3	0.0	0.1	0.4	0.1	0.55	0.03	0.09	0.17	0.01
TrAl	Trabajadores ind.ali	1.77	0.15	0.26	0.02	-0.24	-0.15	0.07	1.4	0.0	2.2	2.5	1.1	0.43	0.00	0.37	0.15	0.03
TrQu	Obreros de la quími	1.02	0.15	0.29	0.24	0.01	-0.05	-0.06	1.1	0.8	0.0	0.2	0.5	0.57	0.37	0.00	0.02	0.03
TrTa	Obreros del tabaco	0.14	0.63	0.64	0.09	0.28	0.14	-0.21	0.7	0.0	0.2	0.2	0.8	0.66	0.01	0.12	0.03	0.07
ONa	Obreros (sect. no e	0.59	0.13	0.29	0.13	-0.04	-0.13	-0.06	0.6	0.1	0.0	0.6	0.3	0.68	0.14	0.01	0.13	0.03
Etiqu	Embaladores y Etiqu	0.57	0.47	0.60	-0.18	0.03	-0.06	-0.17	2.5	0.2	0.0	0.1	2.2	0.76	0.07	0.00	0.01	0.06
MaPe	Conduct.maquín.exca	1.00	0.48	0.28	0.57	0.19	0.18	0.01	0.9	4.3	0.8	2.0	0.0	0.16	0.69	0.07	0.07	0.00
Dock	Dockers y asimilado	4.33	0.07	0.21	-0.02	-0.07	-0.08	0.02	2.4	0.0	0.5	1.6	0.3	0.69	0.00	0.08	0.10	0.01
Polí	Policías. Bomberos	1.61	0.11	-0.11	0.08	-0.24	-0.08	0.09	0.2	0.1	2.1	0.6	1.6	0.12	0.06	0.56	0.06	0.07
TrDo	Trabaj.Domesticos/H	5.25	0.50	0.39	-0.51	-0.25	0.13	0.05	9.5	17.6	7.2	5.0	1.6	0.30	0.52	0.13	0.03	0.01
TrSP	Trabajadores Servic	0.83	0.34	0.26	-0.44	-0.24	0.16	0.00	0.7	2.1	1.0	1.3	0.0	0.20	0.57	0.16	0.07	0.00
Atle	Atletas y Deportist	0.11	0.13	-0.19	0.16	-0.27	-0.03	-0.01	0.1	0.0	0.2	0.0	0.0	0.28	0.18	0.53	0.01	0.00
OtSe	Otras profesiones d	0.12	0.14	-0.13	-0.26	-0.05	0.18	-0.10	0.0	0.1	0.0	0.2	0.1	0.12	0.48	0.02	0.23	0.07
Mili	Militares de carrer	1.22	0.11	-0.25	-0.06	0.04	0.00	0.20	0.9	0.1	0.0	0.0	6.0	0.56	0.04	0.01	0.00	0.36
PrDe	Profesion mal defin	0.56	0.45	0.32	-0.27	0.16	0.42	-0.08	0.7	0.5	0.3	5.8	0.4	0.22	0.17	0.06	0.39	0.01

H. Exemplos de aplicação da Análise Fatorial de Correspondência Simples

4. Análise Fatorial de Correspondência da Tabela T(59x9)

b) Interpretação dos primeiros eixos fatoriais

No espaço de pontos-»grupos profesionales», o primeiro eixo fatorial opõe :

- as categorias profissionais que dispõem de maiores rendas, que necessitam uma maior autonomia ou horários mais flexíveis;
- os grupos profissionais com rendas reduzidas, que trabalham em grandes empresas industriais ou comerciais, sem que essas tarefas necessitem uma grande autonomia do trabalhador.

Este primeiro eixo fatorial constitui um eixo de «escala»...

Ordena os «meios de transporte» e as «categorias profissionais» em ordem decrescente entre dois polos extremos : recursos elevados-distância do local de moradia -flexibilidade de horários versus baixos recursos-proximidade do local de habitação-rigidez de horários.

Resultados complementares :

EDICIÓN DE LAS LINEAS SOBRE EL EJE FACTORIAL I				
NUMERO	IDENTIFICADOR DE LA LINEA	PESO	COORDENADA	
1	Viaj Viajantes de cOmercio	12648.00	-0.54	
2	Piló Pilotos de avión/navegantes	696.00	-0.54	
3	BiAg Biólogos. Agrónomos. Vet.	1236.00	-0.51	
4	Arln Arquitectos. Ingenieros	29253.00	-0.48	
5	CuPr Direct./Cuadros Sup. sect. priv.	98108.00	-0.44	
ZONA CENTRAL				
55	Mine Mineros y asimilados	21371.00	0.59	
56	Etqu Embaladores y Etiquetadores	16076.00	0.60	
57	Sast Sastres. Cortadores y asimil.	48419.00	0.64	
58	TrTa obreros del tabaco	3979.00	0.64	
59	Teje Tejedores. Tintoreros y asimil.	37436.00	0.73	
EDICIÓN DE LAS COLUMNAS SOBRE EL EJE I				
NUMERO	IDENTIFICADOR DE LA COLUMNA	PESO	COORDENADA	
1	TREN	220543.00	-0.28	
2	AUTO-c/CONDUCTOR	1345181.00	-0.20	
ZONA CENTRAL				
3	BICICLETA	315716.00	0.55	
4	TRANS. ORG. EMPRESA	109002.00	0.68	

H. Exemplos de aplicação da Análise Fatorial de Correspondência Simples

4. Análise Fatorial de Correspondência da Tabela T(59 x 9)

b) Interpretação dos primeiros eixos fatoriais

Estatísticas de ajuda para interpretação do segundo eixo fatorial

EDICIÓN DE LAS LINEAS SOBRE EL EJE FACTORIAL II				
NUMERO	IDENTIFICADOR DE LA LINEA		PESO	COORDENADA
1	TrDo	Trabaj.Domesticos/Hotelería	146728.00	-0.51
2	TrSP	Trabajadores Servicios Person.	23310.00	-0.44
3	Vend	Vendedores y asimilados	75217.00	-0.40
4	CaDa	Asist. Cont.. Cajeros. Dactil.	61387.00	-0.40
5	Insp	Inspect./Controlad.(sect.transp.)	4799.00	-0.29
ZONA CENTRAL				
55	CoMa	Contramaestres. Jefes de Equipo	58555.00	0.39
56	Mine	Mineros y asimilados	21371.00	0.42
57	Viaj	Viajantes de comercio	12648.00	0.49
58	Alba	Albañiles/obreros de la Constr.	97433.00	0.55
59	MaPe	Conduct.maquin.excavación/gruas	27966.00	0.57
EDICIÓN DE LAS COLUMNAS SOBRE EL EJE II				
NUMERO	IDENTIFICADOR DE LA COLUMNA		PESO	COORDENADA
1	OMNIBUS-TRANVIA-SUBTE		270726.00	-0.47
2	TREN		220543.00	-0.32
3	AUTO -c/PASAJERO		230057.00	-0.19
ZONA CENTRAL				
4	AUTO-c/CONDUCTOR		1345181.00	0.17
5	TRANS. ORG. EMPRESA		109002.00	0.76

No espaço de pontos- «meios de transporte» o segundo eixo opõe :

- um tipo de transporte até o lugar de trabalho, privado, limitado a áreas urbanas e a categorias profissionais de baixa renda;
- um tipo de transporte até o local de trabalho diretamente ligado ao exercício da profissão e organizado pelas empresas.

H. Exemplos de aplicação da Análise Fatorial de Correspondência Simples

4. Análise Fatorial de Correspondência da Tabela T(59x9)

b) Interpretação dos primeiros eixos fatoriais

Estatísticas de ajuda a interpretação ao terceiro eixo fatorial

EDICIÓN DE LAS LINEAS SOBRE EL EJE FACTORIAL III				
NUMERO	IDENTIFICADOR DE LA LINEA	PESO	COORDENADA	
1	Micu Ministros de Culto	1741.00	-0.81	
2	Pays Agricultores y asimilados	4271.00	-0.75	
3	Agri Dirig./Jefes de expl. agric.	16.00	-0.56	
4	Come Comerciantes	48795.00	-0.49	
5	JeEm Jefes Empr. sect. privado	12082.00	-0.43	
ZONA CENTRAL				
55	TrTa Obreros del tabaco	3979.00	0.28	
56	CuFu Direct./Cuadros Sup. adm. publ.	29838.00	0.29	
57	Alba Albañiles/obrerros de la Constr.	97433.00	0.42	
58	Insp Inspect./Controlad.(sect.transp.)	4799.00	0.51	
59	Diri Despach.Dirig.circulación	15583.00	0.56	
EDICIÓN DE LAS COLUMNAS SOBRE EL EJE III				
NUMERO	IDENTIFICADOR DE LA LINEA	PESO	COORDENADA	
1	UNICAMENTE CAMINANDO	276735.00	-0.32	
2	BICICLETA	315716.00	-0.11	
3	AUTO-c/CONDUCTOR	1345181.00	-0.06	
ZONA CENTRAL				
6	OMNIBUS-TRANVÍA-SUBTE	270726.00	0.08	
8	TRANS. ORG. EMPRESA	109002.00	0.48	
9	TREN	220543.00	0.53	

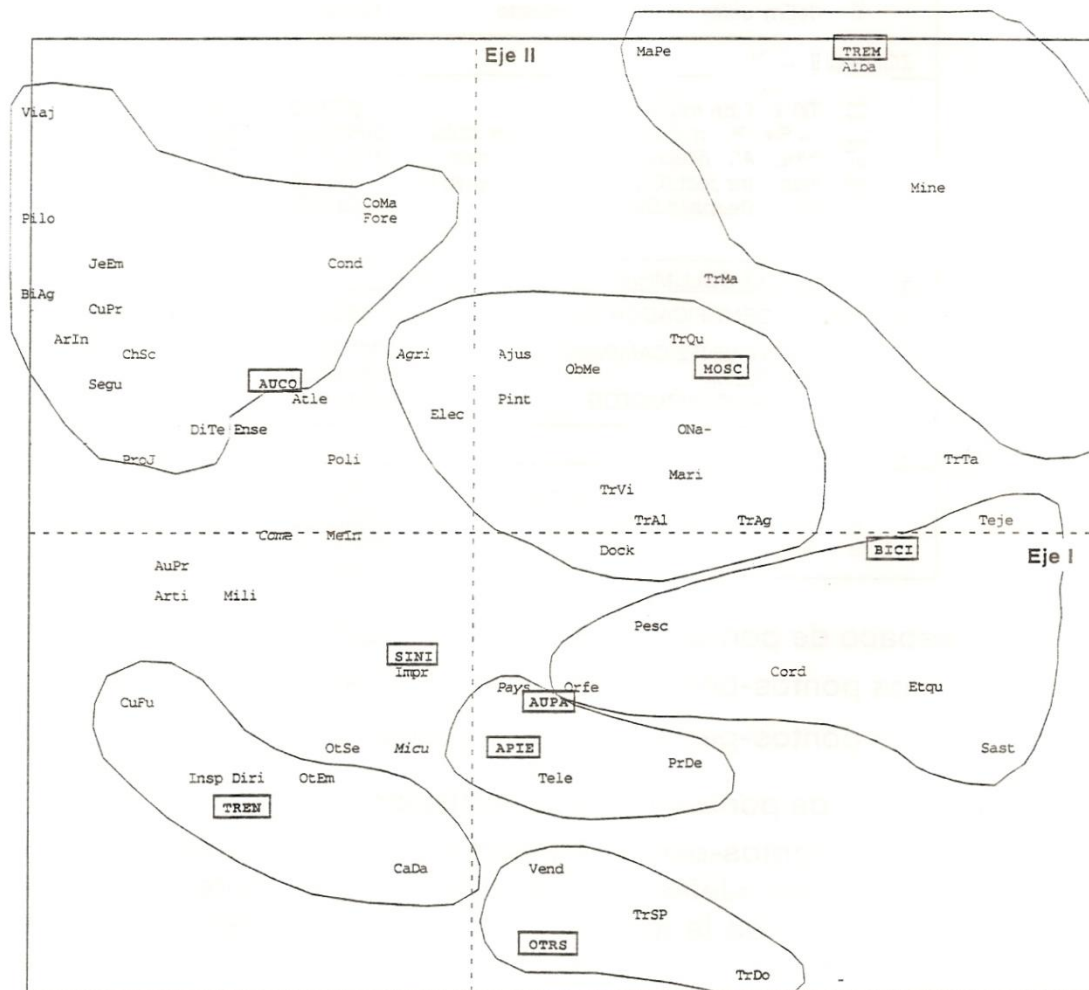
- o espaço de pontos - «meio de transporte», opõe :
 - os pontos-perfis «trans. org. pela empresa» e «trem»;
 - os pontos-perfis «unic.caminando» e «bicicleta».
- o espaço de pontos-»grupos profissionais», opõe :
 - os pontos-perfis «Ministros de Culto», personal sanit. e docente, «Jefes de empr.privadas», agricultores, pescadores, trabaj. de la aliment., «Policiais e Bomberos», «Atletas e deport.»;
 - os pontos-perfis «Cuadros sup. de la función pública», «Inspetores e controladores», «outros empleados», «Despachantes e dirigentes de circulación».

H. Exemplos de aplicação da Análise Fatorial de Correspondência Simples

4. Análise Fatorial de Correspondência da Tabela T(59x9)

- Primeiro plano fatorial :

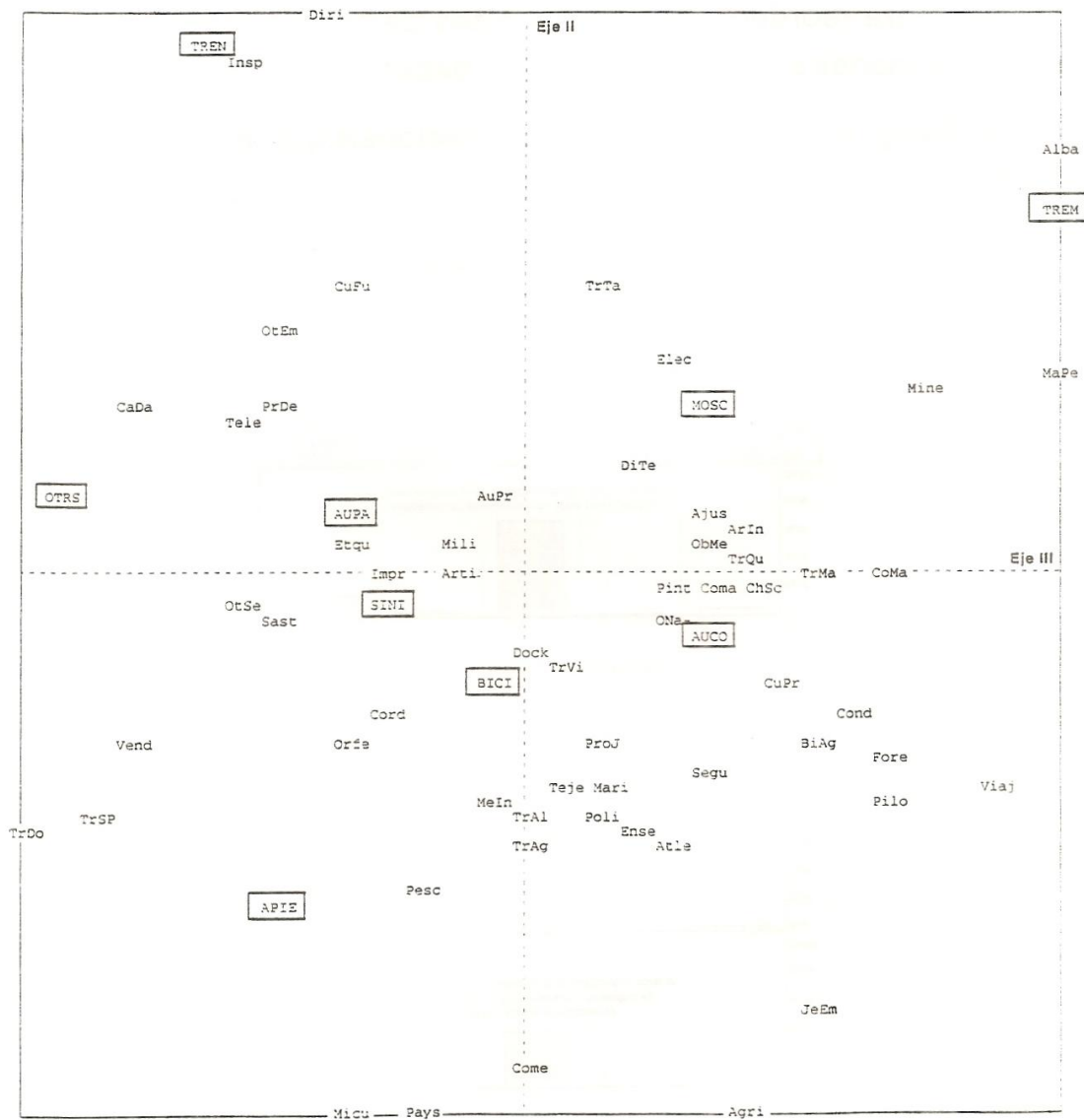
Reagrupando as categorias profissionais que apresentam perfis similares no que se refere aos meios de transporte....



H. Exemplos de aplicação da Análise Fatorial de Correspondência Simples

4. Análise Fatorial de Correspondência da Tabela T(59 x 9)

- Segundo plano fatorial :

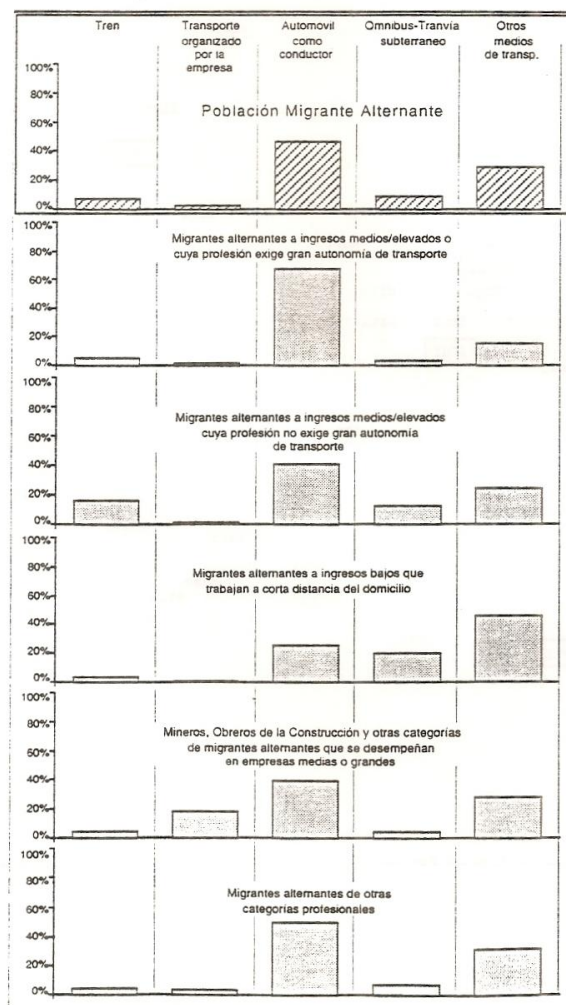


H. Exemplos de aplicação da Análise Fatorial de Correspondência Simples

5. Retorno aos dados e apresentação das conclusões da análise

➤ é necessário referir-se à interpretação dos resultados da análise, para reordenar os dados brutos de tal forma que possamos expor de maneira acessível o sentido dessa informação .

➤ Reagrupando as categorias profissionais que apresentam «perfis de meio de transporte» semelhantes...



H. Exemplos de aplicação da Análise Fatorial de Correspondência Simples

6. Estratégia de interpretação dos resultados da A.F.C. simples

- ☞ A interpretação de resultados de uma A.F.C. é uma etapa delicada destinada a tirar o máximo de proveito da capacidade de síntese que oferece essa técnica estatística.
- ☞ Uma análise eficaz começa com a definição pertinente da Tabela de Contingência que será estudada.
 - ❶ nomeclaturas...
 - ❷ informação suplementar disponível

Etapas do processo de interpretação

a) Estudo da inércia associada aos fatores

- ☞ Estudar a forma das nuvens de pontos e a importância relativa das relações postas assim em evidência.
- ☞ Estudar o decréscimo dos auto-valores não-nulos e as taxas de inércia dos mesmos... seleção de um certo número de eixos fatoriais sobre os quais se articula o trabalho de interpretação.
- ☞ Não tem sentido modificar os elementos ativos da análise tratando de melhorar, por exemplo, a taxa de inércia do primeiro ou do segundo eixo fatorial.

H. Exemplos de aplicação da Análise Fatorial de Correspondência Simples

6. Estratégia de interpretação dos resultados da A.F.C. simples

Etapas do processo de interpretação (continuação)

b) Interpretação dos eixos e planos fatoriais

- Grau de generalidade dos fatores :

☞ O interesse de um eixo fatorial está determinado, fundamentalmente, pelo número de elementos que o configuram: «grau de generalidade» do fator.

- Coordenadas, contribuições e qualidade de representação dos elementos ativos :

☞ Para interpretar a informação que um eixo possa trazer :

- uma forte contribuição da inércia do fator. Esses elementos «constrõem» o fator.

- uma coordenada extrema, conjuntamente com uma qualidade de representação sobre o eixo. Estes elementos permitem qualificar o fator: tendo um perfil muito diferente ao perfil médio (coordenada extrema), sua qualidade de representação mostra que essa diferença está bem «traduzida» pelo fator.

- uma coordenada extrema, conjuntamente com uma qualidade média de representação sobre o eixo. Estes elementos apresentam as características que estão associadas ao fator. Mas estas se somam a outras características colocadas em evidência por outros eixos, o que torna difícil colocá-las em evidência sobre esse único eixo considerado.

- Interpretação dos planos fatoriais

☞ Quando consideramos um plano fatorial tomamos em conta simultaneamente duas dimensões do problema estudado o que fornece uma imagem mais fiel dos dados brutos analisados (síntese).

H. Exemplos de aplicação da Análise Fatorial de Correspondência Simples

6. Estratégia de interpretação dos resultados da A.F.C. simples

Etapas do processo de interpretação
(continuação)

b) Interpretação dos eixos e planos fatoriais

- Interpretação dos elementos suplementares

O uso de elementos suplementares se justifica:

- quando dispomos de uma informação exterior, mas relacionada com o problema estudado,
- da qual esperamos que traga uma modificação interessante, pertinente, do contexto de interpretação.

Na interpretação desses elementos, devemos guiar-nos por suas coordenadas e suas qualidades de representação nos eixos ou nos planos fatoriais.

Os elementos suplementares que tem uma coordenada extrema e uma boa qualidade de representação sobre um eixo permitem:

- explicar eventualmente esse fator;
- melhorar a interpretação proposta do mesmo a partir das modalidades ativas;
- sugerir que reconsideremos a interpretação de um fator.

H. Exemplos de aplicação da Análise Fatorial de Correspondência Simples

6. Estratégia de interpretação dos resultados da A.F.C. simples

Etapas do processo de interpretação
(continuação)

c) Integrar os resultados em seu contexto

➤ Interpretar os resultados, é sobretudo uma operação de integração dos mesmos em três contextos :

- ❶ o primeiro contexto no qual devemos situar os resultados é o da própria Tabela de Contingência .
- ❷ o segundo contexto no qual devemos situar os resultados é constituído pelos elementos suplementares.
- ❸ o terceiro contexto no qual devemos situar os resultados é exterior aos dados brutos analisados.

➤ Os elementos suplementares desempenham um papel fundamental neste terceiro contexto de referência dos resultados.

➤ Desta forma se põe em ação um modo de raciocínio associativo que é muito eficaz... quando se dispõe de instrumentos que permitem a verificação objetiva das diversas etapas.

(cf. ESCOFIER, B. y PAGES, J, Analyses factorielles simples et multiples - Objectifs, méthodes et interprétation, Ed. Dunod, Paris, 1990)

I. Generalização da Análise fatorial de Correspondências : Análise Fatorial de Correspondências Múltiplas

1 - Introdução :

- Por meio de «enquetes», elaboram-se tabelas que resumem as p características observadas em n unidades de observação.
 - Procedendo à redução conceitual das dimensões do observado, os investigadores se interessam pelas análises simultâneas das relações que podem ser estabelecidas entre as diversas unidades temáticas em estudo.
 - As unidades temáticas que orientam o trabalho de redução das múltiplas dimensões do observado resultam do quadro conceitual do estudo, o qual já serviu ao investigador para definir a estratégia de observação colocada em prática.
- A exploração das relações observadas entre essas unidades temáticas e/ou entre elas e certas características observadas, constitui a etapa final do trabalho de construção do objeto do estudo.
- A Análise Fatorial de Correspondências Múltiplas é um instrumento particularmente adaptado ao tratamento estatístico dos dados produzidos por esse tipo de observações.
- Este método de análise estatística responde à dupla exigência de objetividade nesse processo de redução e de exploração do observado e do tratamento da informação com o nível de síntese adequado ao quadro conceitual utilizado.

Porque dizemos que este instrumento de análise estatística responde as exigências metodológicas da investigação em ciências sociais...?

I. Generalização da Análise Fatorial de Correspondências : Análise Fatorial de Correspondências Múltiplas

1 - Introdução (continuação):

- Primeiro erro (muito comum):
 - Ignorando o quadro conceitual com o qual se organizou o estudo, trata-se de responder à demanda do investigador elaborando as $p \times q$ tabelas de contingência que cruzam p características com outras q características observadas...
- Segundo erro (muito comum):
 - Selecionam-se as «Tabelas interessantes» em função do grau de significância estatística do teste de independência feito automaticamente sobre todas essas Tabelas de Contingência calculadas...
 - A informação que interessa ao investigador é dada por uma «tipologia» das n unidades observadas.
- Terceiro erro (muito comum):
 - Essas tipologias são construídas «subjetivamente»...ã
- A A.F.C.M. dos dados de uma pesquisa permite a descrição e a exploração do observado para construir descritores objetivos com o mesmo nível de síntese com o qual se pensou o objeto em estudo.
- A A.F.C.M. pode ser apresentada como uma simples generalização do A.F.C. simples.

I. Generalização da Análise Fatorial de Correspondências : Análise Fatorial de Correspondências Múltiplas

2 - Tabelas estatísticas de resumo do observado

- O investigador dispõe :
 - de um protocolo normalizado de pesquisa.
 - de uma lista padrão de características observáveis. Se no protocolo de inquérito é definido corretamente as categorias utilizadas em cada um dos indicadores operacionais são mutuamente exclusivas.

Primera Característica Observada	Código
1º modalidad 1	1
2º modalidad 2	2
3º modalidad 3	3

Segunda Característica Observada	Código
1º modalidad 1	1
2º modalidad 2	2
3º modalidad 3	3
4º modalidad 4	4
(.....)	

j-ésima Característica Observada	Código
1º modalidad 1	1
2º modalidad 2	2
(.....)	

p-ésima Característica Observada	Código
1º modalidad 1	1
2º modalidad 2	2
3º modalidad 3	3
4º modalidad 4	4

A j-ésima característica utilizada na observação é composta de K_j modalidades mutuamente exclusivas.

Quando se observa a j-ésima característica apresentada pelo i-ésimo indivíduo, pode-se atribuir — sem ambigüidade nenhuma — a k-ésima modalidade dessa j-ésima característica ao i-ésimo indivíduo.

I. Generalização da Análise Fatorial de Correspondências : Análise Fatorial de Correspondências Múltiplas

2 - Tabelas estatísticas de resumo do observado : Tabela de Códigos Condensados (TCC)

- A partir dos documentos de codificação, faz-se um primeiro resumo do observado construindo uma tabela composta de uma linha por indivíduo e de uma coluna por característica observada.

Tabela de Códigos Condensados TCC ($n \times p$)

Ind.	Código de la Característica					
	1°	2°	(...)	j-ésima	(...)	p-ésima
1	1	2	...	2	...	4
2	2	1	...	1	...	3
⋮	⋮	⋮	⋮	⋮	⋮	⋮
i	k_{i1}	k_{i2}	...	k_{ij}	...	k_{ip}
⋮	⋮	⋮	⋮	⋮	⋮	⋮
n	k_{n1}	k_{n2}	...	k_{nj}	...	k_{np}

- Cada linha concontribuições todos os códigos correspondentes às modalidades atribuídas a um indivíduo, para cada uma das características observadas.

- Na interseção da i-ésima linha e da j-ésima coluna figura o valor k_{ij} = o código numérico que foi dado à modalidade «atribuída» ao i-ésimo indivíduo a propósito da j-ésima característica observada.

Tratando-se de uma tabela de códigos, a tabela TCC($n \times p$) não possui propriedades numéricas.

I. Generalização da Análise Fatorial de Correspondências : Análise Fatorial de Correspondências Múltiplas

2 - Tabelas estatísticas de resumo do observado : Tabelas Lógicas (TL)

- ➔ A partir de uma Tabela de Códigos Condensados pode-se gerar uma tabela numérica capaz de resumir a mesma informação.

Considerando a primeira coluna da tabela TCC(n x p)...

Tabla TL(n x 3)
asociada a la Primera Columna
de la tabla TCC(n x p)

Tabla TCC(n x 1)
correspondiente a la Primera
Columna de la tabla TCC(n x p)

Individuo	Código de la Primera Característica Observ.
1	1
2	2
⋮	⋮
i	k _{i1}
⋮	⋮
n	k _{n1}

Individuo	Modalidades de la Primera Característica Observ.			$\sum_{j=1}^3 x_{ij} =$
	Código: 1	Código: 2	Código: 3	
1	1	0	0	1
2	0	1	0	1
⋮	⋮	⋮	⋮	⋮
i	x _{i1}	x _{i2}	x _{i3}	1
⋮	⋮	⋮	⋮	⋮
n	0	0	x _{n1}	1
$\sum_{i=1}^n x_{ij} =$	n _{.1}	n _{.2}	n _{.3}	n

Na tabela TL(n x 3) associada à 1ª coluna de TCC (n x p) :

$$x_{i1} = 1, \text{ si } k_{i1} = 1; \text{ si } k_{i1} = 2 \text{ o } k_{i1} = 3 \Rightarrow x_{i1} = 0$$

$$x_{i2} = 1, \text{ si } k_{i2} = 2; \text{ si } k_{i2} = 1 \text{ o } k_{i2} = 3 \Rightarrow x_{i2} = 0$$

$$x_{i3} = 1, \text{ si } k_{i3} = 3; \text{ si } k_{i3} = 1 \text{ o } k_{i3} = 2 \Rightarrow x_{i3} = 0$$

À 1ª coluna da Tabela TCC(n x p) associamos uma Tabela TL (n x 3), porque essa coluna de TCC(n x p) comporta três modalidades mutuamente exclusivas.

A tabela TL(n x 3) contém as três «variáveis indicadoras» das modalidades da 1ª característica observada («variáveis de presença-ausência»).

I. Generalização da Análise Fatorial de Correspondências : Análise Fatorial de Correspondências Múltiplas

2 - Tabelas estatísticas de resumo do observado : Tabelas Lógicas (TL)

➤ Repetindo essa operação para cada coluna de TCC ($n \times p$), podemos associar a esta uma Tabela TL ($n \times k$).

Tabela Lógica ou Tabela Disjuntiva Completa correspondente à
Tabela TCC ($n \times p$)

Indiví- duos	Primera Característica Modalidades			(***)	j-ésima Característica Modalidades			(***)	p-ésima Característica Modalidades			margen
	1	(...)	k_1		$k_{j-1}+1$	(...)	$k_{j-1}+k_j$		$k_{p-1}+1$	(...)	K	
1	0	...	1	...	0	...	0	...	0	...	0	p
2	1	...	0	...	0	...	0	...	1	...	0	p
3	0	...	1	...	0	...	1	...	0	...	0	p
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
i	0	...	1	...	0	...	x_{ij}	...	0	...	0	p
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
n	1	...	0	...	0	...	x_{nj}	...	0	...	1	p
margen	n_1	...	n_{k_1}	(***)	$n_{k_{j-1}+1}$...	$n_{k_{j-1}+k_j}$	(***)	$n_{k_{p-1}+1}$...	n_K	np

$$K = \sum_{m=1}^p k_m : \text{total de modalidades das } p \text{ características observadas.}$$

A soma em linha dessa tabela = margem constante, igual a p.

Margem inferior da Tabela TL($n \times k$) = distribuição de frequências brutas de todas as características observadas.

- A Tabela TL($n \times K$) resume o observado nos mesmos termos que a Tabela de Códigos Condensados correspondente.
- A Tabela Lógica apresenta propriedades numéricas.
- A Tabela TL($n \times K$) = tabela de Correspondências de tipo particular.

I. Generalização da Análise Fatorial de Correspondências : Análise Fatorial de Correspondências Múltiplas

3 - Análise Fatorial de Correspondências de uma Tabela Lógica TL (n x K)

a) Objetivos da análise

① Facilitar a construção dessas tipologias de indivíduos permitindo a comparação de todas as unidades de observação através de todas as modalidades das características observadas.

② Estudar a relação existente entre as características observadas.

③ Resumir o conjunto de características observadas em um pequeno número de variáveis quantitativas relacionadas com o conjunto de variáveis qualitativas estudadas.

④ Permitir a comparação de modalidades das características observadas.

I. Generalização da Análise Fatorial de Correspondências : Análise Fatorial de Correspondências Múltiplas

3 - Análise Fatorial de Correspondências de uma Tabela Lógica TL (n x K)

b) Distância entre «indivíduos» (linhas) de uma Tabela Lógica

➤ j-ésima coordenada do i-ésimo indivíduo:

$$f_{ij} = \frac{x_{ij}}{np} ; f_{i.} = \frac{1}{n} ; f_{.j} = \frac{n_j}{np} \frac{\frac{x_{ij}}{np}}{\frac{1}{n} \sqrt{\frac{n_j}{np}}} = \frac{x_{ij}}{p \sqrt{\frac{n_j}{np}}}$$

➤ j-ésima coordenada do i'-ésimo indivíduo:

$$f_{i'j} = \frac{x_{i'j}}{np} ; f_{i'.} = \frac{1}{n} ; f_{.j} = \frac{n_j}{np} \frac{\frac{x_{i'j}}{np}}{\frac{1}{n} \sqrt{\frac{n_j}{np}}} = \frac{x_{i'j}}{p \sqrt{\frac{n_j}{np}}}$$

➤ Distância do Qui² entre os pontos-indivíduos i e i':

$$d_{(i,i')}^2 = \sum_{j=1}^K \left(\frac{x_{ij}}{p \sqrt{\frac{n_j}{np}}} - \frac{x_{i'j}}{p \sqrt{\frac{n_j}{np}}} \right)^2 = \sum_{j=1}^K \frac{np}{n_j} \left(\frac{x_{ij}}{p} - \frac{x_{i'j}}{p} \right)^2$$

$$d_{(i,i')}^2 = \frac{1}{p} \sum_{j=1}^K \frac{n}{n_j} (x_{ij} - x_{i'j})^2$$

[1]

I. Generalização da Análise Fatorial de Correspondências : Análise Fatorial de Correspondências Múltiplas

3 - Análise Fatorial de Correspondências de uma Tabela Lógica TL(n x K)

b) Distância entre «indivíduos» (linhas) de uma Tabela Lógica (continuação):

➤ Em [1]...

cada termo dessa somatória $(X_{ij} - X_{i'j})^2$ não pode valer mais que 1 ou 0.

⇒ esse fator vale 1 somente se os dois indivíduos considerados não apresentam simultaneamente a j-ésima modalidade,

⇒ vale 0 no caso contrário (presença simultânea ou ausência simultânea da j-ésima modalidade).

➤ A distância entre os mesmos cresce à medida que aumentam as diferenças de modalidades apresentadas pelos indivíduos i e i' .

➤ Cada modalidade intervém no cálculo da distância entre dois indivíduos com o fator n / nk , ou seja, o inverso do peso da modalidade.

A distância entre os indivíduos será maior, se estes apresentarem o mesmo número de divergências a propósito de modalidades pouco frequentes (de pequeno peso).

➤ A distância do Qui² entre indivíduos de uma Tabela Lógica respeita o critério de comparação de indivíduos que adotamos.

I. Generalização da Análise Fatorial de Correspondências : Análise Fatorial de Correspondências Múltiplas

3 - Análise Fatorial de Correspondências de uma Tabela Lógica TL(n x K)

c) Distância entre modalidades (colunas) de uma Tabela Lógica

Cada coluna de uma Tabela TL(n x K) é uma variável indicadora de uma modalidade de uma das características observadas.

- Coordenadas de modalidades j e k da Tabela Lógica TL(n x K):

$$\text{modalidade j : } \left[\frac{\frac{x_{1j}}{np}}{\frac{n_j}{np} \sqrt{\frac{p}{np}}}; \frac{\frac{x_{2j}}{np}}{\frac{n_j}{np} \sqrt{\frac{p}{np}}}; \dots; \frac{\frac{x_{ij}}{np}}{\frac{n_j}{np} \sqrt{\frac{p}{np}}}; \dots; \frac{\frac{x_{nj}}{np}}{\frac{n_j}{np} \sqrt{\frac{p}{np}}} \right]$$

$$\text{modalidade k : } \left[\frac{\frac{x_{1k}}{np}}{\frac{n_k}{np} \sqrt{\frac{p}{np}}}; \frac{\frac{x_{2k}}{np}}{\frac{n_k}{np} \sqrt{\frac{p}{np}}}; \dots; \frac{\frac{x_{ik}}{np}}{\frac{n_k}{np} \sqrt{\frac{p}{np}}}; \dots; \frac{\frac{x_{nk}}{np}}{\frac{n_k}{np} \sqrt{\frac{p}{np}}} \right]$$

- A distância entre essas duas modalidades:

$$d_{(j,k)}^2 = \sum_{i=1}^n \left[\frac{\frac{x_{ij}}{np}}{\frac{n_j}{np} \sqrt{\frac{1}{n}}} - \frac{\frac{x_{ik}}{np}}{\frac{n_k}{np} \sqrt{\frac{1}{n}}} \right]^2 = \sum_{i=1}^n \left[\frac{x_{ij}}{n_j \sqrt{\frac{1}{n}}} - \frac{x_{ik}}{n_k \sqrt{\frac{1}{n}}} \right]^2$$

- A distância do Qui² entre duas modalidades:

$$d_{(j,k)} = \sqrt{\sum_{i=1}^n n \left[\frac{x_{ij}}{n_j} - \frac{x_{ik}}{n_k} \right]^2} \quad [2]$$

I. Generalização da Análise Fatorial de Correspondências : Análise Fatorial de Correspondências Múltiplas

3 - Análise Fatorial de Correspondências de uma Tabela Lógica TL(n x K)

c) Distância entre modalidades (colunas) de uma Tabela Lógica (continuação):

Desenvolvendo a expressão [2] :

$$d_{(j,k)}^2 = \sum_{i=1}^n n \cdot \left[\frac{x_{ij}^2}{n_j^2} + \frac{x_{ik}^2}{n_k^2} - 2 \frac{x_{ij} x_{ik}}{n_j n_k} \right] \quad (3)$$

Chamando :

n_{jk} : a frequência bruta de indivíduos que apresentam simultaneamente a modalidade j e a modalidade k,

n_j : a frequência bruta de indivíduos que apresentam somente a modalidade j,

n_k : a frequência bruta de indivíduos que apresentam somente a modalidade k,

Na expressão (3), o fator entre colchetes vale :

- $\left(\frac{1}{n_j^2}\right)$ para $(n_j - n_{jk})$ indivíduos que apresentam somente a modalidade j. Ou seja, os indivíduos para os quais $x_{ij} = 1$ enquanto que $x_{ik} = 0$.
- $\left(\frac{1}{n_k^2}\right)$ para $(n_k - n_{jk})$ indivíduos que apresentam somente a modalidade k. Ou seja, os indivíduos para os quais $x_{ik} = 1$ enquanto que $x_{ij} = 0$.
- $\left(\frac{1}{n_j^2} + \frac{1}{n_k^2} - \frac{2}{n_j n_k}\right)$ para los n_{jk} indivíduos que apresentam simultaneamente a modalidades j e a modalidade k. Ou seja, os indivíduos para os quais $x_{ik} = 1$ e $x_{ij} = 1$.
- Esse fator vale 0 para todos os indivíduos que não apresentam nem a modalidade j nem a modalidade k.

I. Generalização da Análise Fatorial de Correspondências : Análise Fatorial de Correspondências Múltiplas

3 - Análise Fatorial de Correspondências de uma Tabela Lógica TL(n x K)

c) Distância entre modalidades (colunas) de uma Tabela Lógica (continuação):

Substituindo em [2] esses valores, obtemos :

$$d_{(j,k)}^2 = n \cdot \left[\left(\frac{n_j - n_{jk}}{n_j n_k} \right) + \left(\frac{n_k - n_{jk}}{n_j n_k} \right) \right]$$

O quadrado da distância entre as modalidades j e k é igual à proporção de indivíduos que apresentam a modalidade j, porém não apresentam a modalidade k, mais a proporção de indivíduos que apresentam a modalidade k, porém não apresentam a modalidade j.

A distância entre duas modalidades de uma Tabela Lógica cresce em função da proporção de indivíduos que apresentam uma e somente uma das duas modalidades. Por outra parte, esta distância é inversamente proporcional à importância relativa de cada uma das modalidades.

De modo que:

- duas modalidades de uma mesma característica (mutuamente exclusivas) estarão obrigatoriamente muito distantes no espaço de representação;

- duas modalidades comuns à maioria dos indivíduos observados estarão representadas em um mesmo ponto do espaço;

- duas modalidades muito raras estarão representadas muito distantes de todas as outras.

➡ A distância do Qui² atende corretamente ao critério de comparação de duas modalidades que definimos.

I. Generalização da Análise Fatorial de Correspondências : Análise Fatorial de Correspondências Múltiplas

4 - Definição de um novo referencial de representação das nuvens de pontos N(I) e N(J): os eixos fatoriais.

➤ Dispondo-se:

- das coordenadas dos pontos-perfis ponderados das modalidades da Tabela Lógica, no referencial de indivíduos,
- das coordenadas dos pontos-perfis ponderados dos indivíduos da Tabela Lógica, no referencial das modalidades,
- das coordenadas do Centro de Gravidade G_c da nuvem N(J) de pontos-perfis das modalidades,
- das coordenadas do Centro de Gravidade G_L da nuvem N(I) de pontos-perfis dos indivíduos;

$$\text{Coordenadas de } G_c : \left[\sqrt{\frac{1}{n}}; \sqrt{\frac{1}{n}}; \dots; \sqrt{\frac{1}{n}}; \dots; \sqrt{\frac{1}{n}} \right] \quad \forall i = 1; \dots; n$$

$$\text{Coordenadas de } G_L : \left[\sqrt{\frac{n_1}{np}}; \dots; \sqrt{\frac{n_j}{np}}; \dots; \sqrt{\frac{n_K}{np}} \right] \quad \forall j = 1, \dots, K$$

- podemos centrar essas nuvens de pontos-perfis ponderados no ponto G correspondente.
- dos pesos associados a cada ponto-perfil ponderado da Tabela Lógica.
- é possível calcular as matrizes de inércia das duas nuvens de pontos-perfis centradas nos Centros de Gravidade respectivos.
- A busca das direções principais de alargamento das nuvens de pontos N(I) e N(J) se faz por meio da operação chamada de «diagonalização» dessas matrizes de inércia.

➤ Determinam-se assim os eixos fatoriais (novo referencial).

I. Generalização da Análise Fatorial de Correspondências : Análise Fatorial de Correspondências Múltiplas

5 - Inércia total relativa ao Centro de Gravidade da nuvem de pontos N(J)

- A contribuição da j-ésima modalidade à inércia global relativa ao Centro de Gravidade G_c é assim definida:

$$Contr_{.(j)} a I_{G_c} = p_j d_{(j,G_c)}^2 = \frac{n_j}{np} d_{(j,G_c)}^2 \quad [1]$$

- O quadrado da distância da j-ésima modalidade ao Centro de Gravidade G_c :

$$d_{(j,G_c)}^2 = \sum_{i=1}^n \left(\frac{nx_{ij}^2}{n_j} + \frac{1}{n} - 2 \frac{x_{ij}}{n_j} \right)^2 \quad [2]$$

- Sabendo-se que:

x_{ij} : 1 para os n_j indivíduos que apresentam a j-ésima modalidade,

x_{ij} : 0 no caso dos $(n - n_j)$ indivíduos que apresentam a j-ésima modalidade,

$$d_{(j,G_c)}^2 = n_j \left(\frac{n}{n_j} + \frac{1}{n} - \frac{2}{n_j} \right) + (n - n_j) \frac{1}{n} = \frac{n}{n_j} + \frac{n_j}{n} - 2 + 1 - \frac{n_j}{n}$$

$$d_{(j,G_c)}^2 = \frac{n}{n_j} - 1$$

$$Contr_{.(j)} a I_{G_c} = \frac{n_j}{np} d_{(j,G_c)}^2 = \frac{n_j}{np} \left(\frac{n}{n_j} - 1 \right) = \frac{1}{p} \left(1 - \frac{n_j}{n} \right) \quad [3]$$

A expressão [3] indica que : uma modalidade rara \Rightarrow forte contribuição à inércia global da nuvem de pontos N(J).

- Consequência imediata...

I. Generalização da Análise Fatorial de Correspondências : Análise Fatorial de Correspondências Múltiplas

5 - Inércia total relativa ao Centro de Gravidade da nuvem de pontos N(J) (continuação):

➤ A inércia global com respeito a G_c da nuvem de pontos N(J) :

$$\sum_{j=1}^K \text{Contr.}_{(j)} a I_{G_c} = \sum_{j=1}^K \frac{1}{p} \left(1 - \frac{n_j}{n} \right) = \sum_{j=1}^K \frac{1}{p} - \sum_{j=1}^K \frac{n_j}{np}$$

Ou seja : $I_{G_c}^{N(J)} = \frac{K}{p} - 1$

➤ Consequência imediata...

I. Generalização da Análise Fatorial de Correspondências : Análise Fatorial de Correspondências Múltiplas

6 - Relações de Transição

➤ Tratando-se da análise de uma Tabela Lógica, as relações de transição são definidas da seguinte maneira:

$$F_{\alpha}(i) = \frac{1}{\sqrt{\lambda_{\alpha}}} \sum_{j=1}^K \frac{x_{ij}}{p} \times G_{\alpha}(j) \quad \forall i = 1, \dots, n$$

$$G_{\alpha}(j) = \frac{1}{\sqrt{\lambda_{\alpha}}} \sum_{i=1}^n \frac{x_{ij}}{n_j} \times F_{\alpha}(i) \quad \forall j = 1, \dots, K$$

Como x_{ij} vale 0 ou 1, a interpretação dessas relações de transição é muito simples :

- 1 a primeira relação mostra que a projeção de um indivíduo i (i -ésima linha da Tabela Lógica) no eixo fatorial α corresponde (com um coeficiente de dilatação) ao baricentro das modalidades que foram observadas no i -ésimo indivíduo.
- 2 a segunda relação mostra que a j -ésima modalidade da Tabela Lógica é representada no eixo fatorial α (com um coeficiente de dilatação) pela coordenada média dos indivíduos observados que apresentaram a modalidade j .

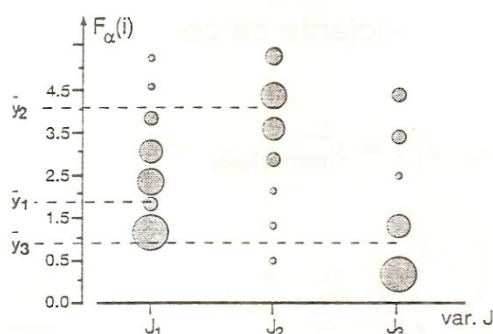
A proximidade, no eixo fatorial α , de duas modalidades k e m traduz graficamente uma forte associação entre essas variáveis indicadoras.

A proximidade, no eixo fatorial α , entre dois indivíduos observados traduz graficamente a semelhança das características apresentadas por esses indivíduos.

I. Generalização da Análise Fatorial de Correspondências : Análise Fatorial de Correspondências Múltiplas

7 - Estudo das variáveis na análise de uma Tabela Lógica

Relação entre uma variável qualitativa observada e um fator resultante da análise de uma Tabela Lógica.



Nas abcissas estão representadas as três modalidades J_1 , J_2 e J_3 da variável qualitativa J que compõe uma Tabela Lógica $TL(n \times K)$.

A variável J é uma das p variáveis da Tabela, de modo que as modalidades J_1 , J_2 e J_3 são três das K colunas da mesma.

Nas ordenadas estão representadas as coordenadas $F_\alpha(i)$ dos n indivíduos que conidos na tabela.

Sendo:

$$\bar{y}_1 = \frac{1}{n_{J_1}} \sum_{i=1}^{n_{J_1}} F_\alpha(i) ; \quad \bar{y}_2 = \frac{1}{n_{J_2}} \sum_{i=1}^{n_{J_2}} F_\alpha(i) ; \quad \bar{y}_3 = \frac{1}{n_{J_3}} \sum_{i=1}^{n_{J_3}} F_\alpha(i)$$

da segunda relação de transição:

$$\sqrt{\lambda_\alpha} G_\alpha(J_1) = \bar{y}_1 ; \quad \sqrt{\lambda_\alpha} G_\alpha(J_2) = \bar{y}_2 ; \quad \sqrt{\lambda_\alpha} G_\alpha(J_3) = \bar{y}_3$$

I. Generalização da Análise Fatorial de Correspondências : a Análise Fatorial de Correspondências Múltiplas

7 - Estudo das variáveis na análise de uma Tabela Lógica (continuação)

Como avaliar o grau de associação entre a variável J e o eixo α ...?

A intensidade da associação entre essa variável qualitativa e o eixo fatorial é medida pelo coeficiente de correlação η^2 :

$$\sqrt{\lambda_\alpha} G_\alpha(J_1) = s_{\text{inter-clases}}^2 = \sum_{k=1}^3 p_k \lambda_\alpha G_\alpha^2(J_k) =$$

$$s_{\text{inter-clases}}^2 = \sum_{k=1}^3 \lambda_\alpha \frac{n_k}{n} G_\alpha^2(J_k) = p \cdot \lambda_\alpha \sum_{k=1}^3 \frac{n_k}{np} G_\alpha^2(J_k)$$

coeficiente de correlação η^2 entre a variável qualitativa J e o eixo α :

$$\eta^2 = \frac{s_{\text{inter-clases}}^2}{\lambda_\alpha} = p \cdot \sum_{k=1}^3 \frac{n_k}{np} G_\alpha^2(J_k) = p \cdot \sum_{k=1}^3 p_k G_\alpha^2(J_k)$$

Generalizando :

$$\eta_{(F_\alpha; J)}^2 = p \times \lambda_\alpha \times \left[\sum_{k=k_{j-1}+1}^{k=k_{j-1}+k_j} \text{Inercia}_{//\alpha}^{(J_k)} \right] \quad \forall J = 1, 2, \dots, p$$

$$\eta_{(F_\alpha; J)}^2 = p \times \lambda_\alpha \times \left[\sum_{k=k_{j-1}+1}^{k=k_{j-1}+k_j} \tau(J_k) \right] \quad \forall J = 1, 2, \dots, p$$

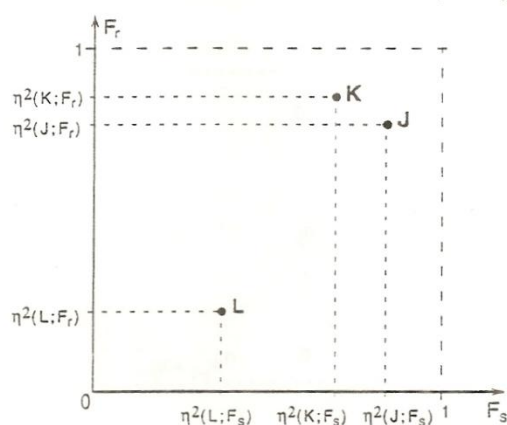
I. Generalização da Análise Fatorial de Correspondências : a Análise Fatorial de Correspondências Múltiplas

7 - Estudo das variáveis na análise de uma Tabela Lógica (continuação)

Os programas informáticos especializados calculam e editam os valores $\eta^2_{(jk)}$ para o conjunto de variáveis.

Na prática da análise de uma Tabela Lógica...

Gráfico de variáveis qualitativas J, K e L no referencial dos fatores F_r e F_s



Nesses espaços, a proximidade entre dois pontos-variáveis expressa a semelhança das partições dos n indivíduos observados produzidas por essas duas variáveis.

Desta maneira se satisfaz o terceiro objeto específico da análise fatorial de uma Tabela Lógica.

I. Generalização da Análise Fatorial de Correspondências : a Análise Fatorial de Correspondências Múltiplas

8 - Outro resumo (mais prático) do observado : a Tabela de Burt (TB)

Tabela de Burt TB ($K \times K$)

		Primera Característica			j-ésima Característica				p-ésima Característica				margen	
		Modalidades			Modalidades				Modalidades					
		1	(...)	k_1	$k_{j-1}+1$	(...)	j	(...)	$k_{j-1}+k_j$	$k_{p-1}+1$	(...)	p	(...)	K
Primera Característica Modalidades	1	n_{11}	0	0	n_{1j}	n_{1p}	...	$p \times n_{1\cdot}$
	(...)	0	...	0
	k_1	0	0	n_{k_1}	$p \times n_{\cdot k_1}$
j-ésima Característica Modalidades	$k_{j-1}+1$	$n_{(k_{j-1}+1)\cdot}$	0	$n_{(k_{j-1}+1)p}$...	$p \times n_{\cdot (k_{j-1}+1)}$
	(...)
	j	n_{j1}	0	...	n_{jj}	...	0	n_{jp}	...	$p \times n_{\cdot j}$
	(...)
p-ésima Característica Modalidades	$k_{j-1}+k_j$	0	...	$n_{(k_{j-1}+k_j)\cdot}$	$n_{(k_{j-1}+k_j)p}$...	$p \times n_{\cdot (k_{j-1}+k_j)}$
	(...)
	$k_{p-1}+1$	$n_{(k_{p-1}+1)\cdot}$	0	$p \times n_{\cdot (k_{p-1}+1)}$	
	(...)
p-ésima Característica Modalidades	p	n_{p1}	n_{pj}	0	...	n_{pp}	...	$p \times n_{\cdot p}$
	(...)
	K	0	n_{K}	...	$p \times n_{\cdot K}$
margen		$p \times n_{\cdot 1}$...	$p \times n_{\cdot k_1}$	$p \times n_{\cdot j}$	$p \times n_{\cdot p}$	$p \times n_{\cdot K}$	$p^2 \times n$

A tabela de Burt TB($K \times K$) é uma tabela simétrica que apresenta o conjunto das Tabelas de Contingência. Estas podem ser construídas cruzando duas a duas as p características observadas.

I. Generalização da Análise Fatorial de Correspondências : a Análise Fatorial de Correspondências Múltiplas

8 - Outro resumo (mais prático) do observado: a Tabela de Burt (TB) (continuação)

Chamando-se: K a soma das modalidades das p características observadas nos n indivíduos, a Tabela de Burt $TB(K \times K)$ apresenta :

- > na interseção da j -ésima linha e da j -ésima coluna, o valor n_{jj} , isto é, o número de indivíduos que apresentam a j -ésima modalidade de uma característica dada,
- > na interseção da j -ésima linha e da k -ésima coluna, o valor $n_{jk} = 0$ se a j -ésima modalidade e a k -ésima modalidade pertencem a uma mesma característica observada,
- > na interseção da q -ésima linha e da k -ésima coluna, o valor n_{qk} , indica o número de indivíduos que apresentam, simultaneamente, a q -ésima modalidade de uma característica dada e a k -ésima modalidade da outra característica observada.

A diagonal da Tabela $TB(K \times K)$ é constituída pela distribuição de frequências brutas de todas as modalidades de todas as variáveis qualitativas.

I. Generalização da Análise Fatorial de Correspondências : a Análise Fatorial de Correspondências Múltiplas

9 - A.F.C. de uma Tabela de Burt

Considerando a Tabela TB(K x K) como Tabela de Correspondências, qual é a informação fornecida por essa Tabela...?

Cada variável qualitativa realiza uma partição dos n indivíduos observados em tantas classes quanto sejam as modalidades que constituam a variável. Duas classes de indivíduos são semelhantes quando apresentam perfis equivalentes de distribuição do conjunto de variáveis qualitativas consideradas.

- A j-ésima coordenada da i-ésima linha de uma Tabela TB(K x K)

$$\left[\frac{\frac{n_{ij}}{p^2 n}}{\frac{n_{i.}}{p^2 n} \sqrt{\frac{n_{.j}}{p^2 n}}} \right] = \sqrt{\frac{n}{p}} \left(\frac{n_{ij}}{n_{i.} n_{.j}} \right)$$

- A distância do Qui² entre duas modalidades i e i' :

$$d_{(i,i')}^2 = \sum_{j=1}^K \left(\sqrt{\frac{n}{p}} \frac{n_{ij}}{n_{i.} n_{.j}} - \sqrt{\frac{n}{p}} \frac{n_{i'j}}{n_{i'.} n_{.j}} \right)^2 = \frac{1}{p} \sum_{j=1}^K \frac{n}{n_{.j}} \left(\frac{n_{ij}}{n_{i.}} - \frac{n_{i'j}}{n_{i'.}} \right)^2$$

I. Generalização da Análise Fatorial de Correspondências : a Análise Fatorial de Correspondências Múltiplas

10 - Análise Fatorial de Correspondência de uma Tabela TB($K \times K$) e de uma Tabela TL($n \times K$)

A análise fatorial de uma Tabela TB($K \times K$) permite representar a nuvem de pontos-modalidades não como baricentros da nuvem de pontos-indivíduos (Tabela Lógica), mas diretamente, a partir da semelhança entre as classes de indivíduos.

- Que relações existem entre a Tabela TL e a Tabela TB, associadas a uma Tabela de Códigos Condensados...?
 - ❑ o k -ésimo ponto-perfil ponderado (da Tabela TB) está localizado no baricentro dos perfis dos n_k indivíduos (da Tabela TL) que apresentam a k -ésima modalidade,
 - ❑ as modalidades (colunas) da Tabela TL e da Tabela TB têm o mesmo peso : os pontos-indivíduos (definidos a partir de uma Tabela TL) e seus baricentros (definidos a partir da Tabela TB) estão situados no mesmo espaço euclidiano, no mesmo referencial de modalidades.
 - ❑ os pontos-indivíduos na representação da Tabela TL têm todos o mesmo peso, enquanto que os pontos-modalidades na representação da Tabela TB estão afetados de um peso que é proporcional à importância da classe.
 - ❑ A análise de uma nuvem de pontos $N(J)$ (baricentros dos indivíduos de uma Tabela Lógica) se obtém mediante a análise fatorial de Correspondência simples de uma Tabela de Burt.
- Pode-se demonstrar que a análise fatorial de uma Tabela Lógica e de uma Tabela de Burt levam a resultados equivalentes.
- O que indica que esses resultados são equivalentes...?

I. Generalização da Análise Fatorial de Correspondências : a Análise Fatorial de Correspondências Múltiplas

10 - Análise Fatorial de Correspondência de uma Tabela TB($K \times K$) e de uma Tabela TL($n \times K$)

- Os eixos fatoriais associados à nuvem de pontos-indivíduos (linhas de uma T.L.) são similares aos eixos fatoriais associados à nuvem de pontos de seus baricentros (linhas de uma T.B.).
- Obtém-se a melhor representação simultânea de ambas nuvens de pontos-perfis de duas maneiras :
 - ❑ tendo-se determinado os eixos fatoriais no espaço de pontos-baricentro (linhas da T.B.) pode-se projetar sobre esses eixos as linhas da Tabela Lógica como elementos suplementares.
 - ❑ tendo-se determinado os eixos fatoriais no espaço de pontos-indivíduos (linhas da T.L) pode-se projetar sobre esses eixos as linhas da T.B. como elementos suplementares.
- Na representação fatorial da informação fornecida por uma Tabela Lógica, interpreta-se a proximidade entre pontos-colunas :
 - ❑ em termos de associação entre variáveis indicadoras,
 - ❑ em termos de semelhança de perfis da distribuição de variáveis qualitativas entre classes de indivíduos.
- Os eixos fatoriais da representação de variáveis indicadoras (colunas da T.L.) e dos eixos fatoriais de representação dos baricentros de indivíduos da mesma (linhas ou colunas da T.B.) não são idênticos. Trata-se de duas representações homotéticas.
 - ❑ Os eixos fatoriais (referencial) são proporcionais, com uma relação $\sqrt{\lambda_\alpha}$ para todo eixo fatorial de ordem α .
- Os valores próprios λ_α associados aos eixos fatoriais de uma T.B. são iguais ao quadrado de seus homólogos em uma T.L.
- A análise da informação produzida por meio de pesquisas é feita mediante a análise fatorial da Tabela de Burt.

I. Generalização da Análise Fatorial de Correspondências : a Análise Fatorial de Correspondências Múltiplas

11 - Componentes digitais de representação da informação de uma T.L. (ou de uma T.B.)

a) Coeficientes de ajuda á seleção dos eixos fatoriais

▲ Auto-Valores :

- Para cada eixo fatorial, a soma das contribuições à inércia ao longo do eixo α de todas as modalidades das p variáveis = inércia da nuvem de pontos-modalidades projetada ao longo do referido eixo (λ_α).
- A inércia ao longo do eixo α de todas as modalidades das p variáveis = valor médio do índice de correlação η^2 , das p variáveis qualitativas com o eixo α .

$$\text{Se } \lambda_\alpha = \frac{1}{p} \sum_{j=1}^p \eta_{(G_\alpha; j)}^2, \text{ quando } \eta_{(G_\alpha; j)}^2 = 1 \quad \forall j = 1, \dots, p \Rightarrow \lambda_\alpha = 1$$

Os auto-valores de uma A.F.C.M. estão no intervalo $[0,1]$.

- Na prática da sondagem, os p auto-valores λ_α resultantes de uma Análise Fatorial de Correspondências Múltiplas são : $0 < \lambda_\alpha < 1$.

I. Generalização da Análise Fatorial de Correspondências : a Análise Fatorial de Correspondências Múltiplas

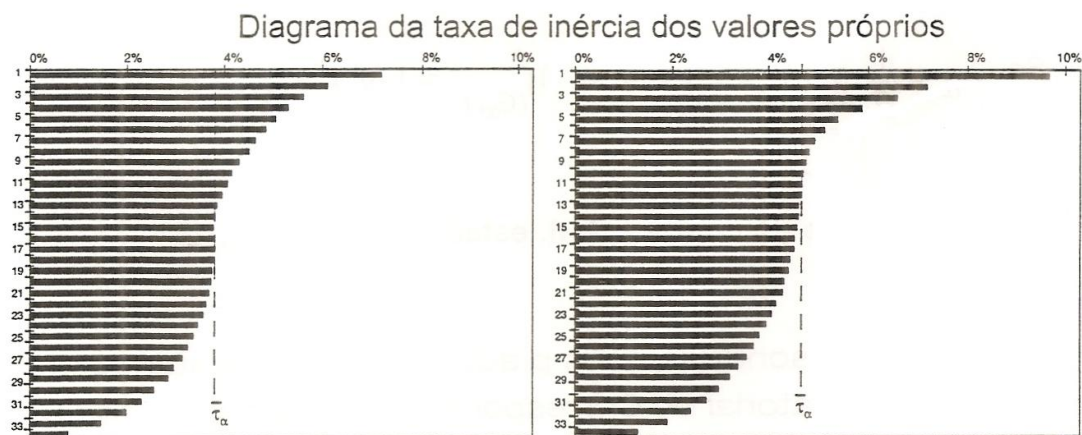
11 - Componentes digitais de representação da informação de uma T.L. (ou de uma T.B.)

a) Coeficientes de ajuda à seleção dos eixos fatoriais

▲ Taxa de inércia de um fator :

$$\tau_{\alpha} = \frac{\lambda_{\alpha}}{\sum_{\alpha=1}^p \lambda_{\alpha}} \cdot 100 = \frac{\lambda_{\alpha}}{I_G} \cdot 100$$

A taxa de inércia de um fator não pode ser interpretada em termos do índice do grau de importância ou de interesse que apresenta um fator.



I. Generalização da Análise Fatorial de Correspondências : a Análise Fatorial de Correspondências Múltiplas

11 - Componentes digitais de representação da informação de uma T.L. (ou de uma T.B.)

b) Coeficientes de ajuda à interpretação dos eixos fatoriais

▲ Estudo dos eixos fatoriais com relação às variáveis :

Os programas de aplicação editam:

- as contribuições à inércia ao longo de cada eixo fatorial do conjunto de modalidades da Tabela de Burt,

- a soma dessas contribuições para o conjunto de modalidades de uma mesma variável qualitativa.

Os programas de aplicação fornecem os elementos necessários para o cálculo do índice de correlação η^2 entre cada variável qualitativa da tabela e cada eixo fatorial.

O utilizador dispõe assim de um índice que permite analisar a importância de cada variável qualitativa na configuração de cada eixo fatorial.

▲ Contribuição dos indivíduos e das modalidades à inércia ao longo de um eixo :

$$CTR_{\alpha}(i) = \frac{\frac{1}{n} [F_a^2(i)]}{\lambda_{\alpha}} \times 100 \forall i \in N(I)$$

$$CTR_{\alpha}(j) = \frac{\frac{n_{.j}}{n} [G_a^2(j)]}{\lambda_{\alpha}} \times 100 \forall j \in N(J)$$

I. Generalização da Análise Fatorial de Correspondências : a Análise Fatorial de Correspondências Múltiplas

11 - Componentes digitais de representação da informação de uma T.L. (ou de uma T.B.)

b) Coeficientes de ajuda à interpretação dos eixos fatoriais

- ▲ «grau de generalidade» de cada eixo fatorial :

Um eixo fatorial apresenta um alto grau de generalidade quando a inércia projetada ao longo do mesmo resulta da contribuição de uma grande quantidade de elementos (indivíduos ou modalidades).

- ▲ «qualidade de representação» das modalidades sobre um eixo fatorial.

Não é possível utilizar esse índices como instrumento de ajuda na interpretação de resultados de um A.F.C.M..

As modalidades de uma mesma variável —tratando-se de atributos mutuamente exclusivos— constituem um conjunto de variáveis indicadoras mutuamente independentes (ortogonais entre elas).

As modalidades de uma mesma variável qualitativa, não podem estar simultaneamente bem representadas sobre um mesmo eixo fatorial.

I. Generalização da Análise Fatorial de Correspondências : a Análise Fatorial de Correspondências Múltiplas

12 - Elementos suplementares numa A.F.C.M.

▲ Indivíduos suplementares :

$$F_{\alpha}(i'_s) = \sqrt{\frac{1}{\lambda_{\alpha}}} \sum_{j=1}^K \frac{x'_{sj}}{p} \cdot G_{\alpha}(j) = \frac{1}{p} \sqrt{\frac{1}{\lambda_{\alpha}}} \sum_{j=1}^K x'_{sj} G_{\alpha}(j)$$

O indivíduo suplementar i'_s será localizado no baricentro das modalidades apresentadas por ele.

		modalidades de p variables cualitativas									
		1	j				K				
1											
i		1	0	1	0	x_{ij}	0	1	0	1	p
n											
		$n_{.1}$				$n_{.j}$				$n_{.K}$	n_p
1											
i'		1	0	1	0	x'_{ij}	1	0	1	p	
n'											

I. Generalização da Análise Fatorial de Correspondências : a Análise Fatorial de Correspondências Múltiplas

12 - Elementos suplementares numa A.F.M.C.

▲ Atributos suplementares :

Podemos justapor à Tabela Lógica, as colunas das variáveis indicadoras das diferentes categorias ilustrativas.

Coordenadas fatoriais da variável indicadora C_q :

$$G_\alpha(C_q) = \frac{1}{P} \sqrt{\frac{1}{\lambda_\alpha}} \sum_{i=1}^n \left(\frac{x_{iq}}{n_{.q}} \times F_\alpha(i) \right)$$

	modalidades de p variables cualitativas								Atributos Ilustrativos			
	1					j				K		
1												
i	1	0	1	0	x_{ij}	1	0	1	p			
n												
	$n_{.1}$				$n_{.j}$				$n_{.K}$	n_p		

Atributos Ilustrativos		
	q	
1	0	1
0	1	0
1	0	0
0	1	1
1	x_{iq}	0
0	1	1
0	1	1
0	1	1
...	$n_{q.}$...

I. Generalização da Análise Fatorial de Correspondências : a Análise Fatorial de Correspondências Múltiplas

12 - Elementos suplementares numa A.F.C.M.

▲ Modalidades suplementares :

Podemos obter a representação de pontos-modalidades de uma Tabela Lógica fazendo a análise — mais econômica— da Tabela de Burt correspondente.

Coordenadas do perfil-coluna ilustrativo C_q :

$$F_{\alpha}(C_q) = \frac{1}{p} \sqrt{\frac{1}{\lambda_{\alpha}}} \sum_{j=1}^K \left(\frac{x_{qj}}{n} \times G_{\alpha}(j) \right)$$

		modalidades de p variables cualitativas			Mod.	Var. Ilustrativas
		1	j	K		C_q
modalidades de p variables cualitativas	1	n_{11}	n_{1j}	n_{1K}		n_{1q} TC ($K_1 \times Q$)
	j	n_{j1}	n_{jj}	n_{jK}		n_{jq} TC ($K_2 \times Q$)
	K	n_{K1}	n_{Kj}	n_{KK}		n_{Kq} TC ($K_3 \times Q$)
		$p \times n_{1.}$	$p \times n_{.j}$	$p \times n_{.K}$		$p \times n_{.q}$

➔ Como utilizar a técnica de elementos suplementares...?

I. Generalização da Análise Fatorial de Correspondências : a Análise Fatorial de Correspondências Múltiplas

13 - Utilização de elementos suplementares

▲ Exemplo...

i n d	Apreciación de diferentes tipos de programas de radio						Características socio-culturales				
	P ₁	P ₂	P ₃	P ₄	P ₅	P ₆	...	Edad	Sexo	Instr.	Ingresos
1	1	2	1	4	4	3		5	2	6	7
2	2	1	3	2	2	2		3	1	4	3
...											
i	x _{p1}	x _{p2}	x _{p3}	x _{p4}	x _{p5}	x _{p6}		x _E	x _S	x _{Ins}	x _{In}
...											
n	x _{p1}	x _{p2}	x _{p3}	x _{p4}	x _{p5}	x _{p6}		x _E	x _S	x _{Ins}	x _{Ing}

- o que significa «explicar» a estrutura da opinião da população observada em função de suas características sociais...?
- Podemos dar diferentes (boas e más) respostas a esta questão.

I. Generalização da Análise Fatorial de Correspondências : a Análise Fatorial de Correspondências Múltiplas

14 - Um instrumento de caracterização dos eixos fatoriais : os «valores de prova»

a) Os «valores de prova» : um instrumento de exploração de tabelas de grande tamanho

Seja o exemplo que, em n indivíduos, se tenha observado p variáveis quantitativas e q variáveis qualitativas.

A p -ésima coluna da tabela têm os n valores da p -ésima variável quantitativa observada.

A k -ésima coluna da tabela descreve uma partição dos n indivíduos em tantas classes como modalidades têm la k -ésima variável qualitativa : n_{1k} indivíduos possuem a modalidade «1» da k -ésima variável qualitativa, n_{2k} indivíduos possuem a modalidade «2», etc.

■ A exploração da tabela de resumo do observado deve responder a dois questionamentos:

- a) Como fazer para ordenar as p variáveis quantitativas observadas segundo a importância de cada uma delas para caracterizar cada um desses grupos de indivíduos?,
- b) Como fazer para determinar, nessas classes, quais são seus atributos mais típicos (modalidades) das outras $q-1$ variáveis qualitativas..?

I. Generalização da Análise Fatorial de Correspondências : a Análise Fatorial de Correspondências Múltiplas

14 - Um instrumento de caracterização dos eixos fatoriais : os «valores de prova»

b) Variáveis quantitativas características de uma classe

Para caracterizar a relação existente entre a p-ésima variável quantitativa e a k-ésima partição, podemos proceder de maneira muito semelhante ao caso de uma prova estatística clássica.

Faz-se a seguinte hipótese estatística:

« H_0 : os n_{jk} valores da p-ésima variável quantitativa são extraídos ao acaso entre os n indivíduos». Ou seja, $H_0: \mu = \bar{x}_{jk} = \bar{x}$;

« H_1 : os n_{jk} valores da p-ésima variável quantitativa não são extraídos ao acaso entre os n indivíduos». Ou seja, $H_1: \mu \neq \bar{x}_{jk} \neq \bar{x}$;

e se consideramos que esses valores constituem uma «amostra» aleatória simples, sem reposição, de n_{jk} observações em n ;

podemos então calcular a probabilidade de observar um valor «média da amostra» que seja tão extremo como o observado numa classe de indivíduos.

Denominamos:

- \bar{x} : a média geral da p-ésima variável quantitativa observada nos n indivíduos,
- σ^2 : a variância dessa variável,
- \bar{x}_{jk} : é a média intra-classe da k-ésima partição dos n_{jk} indivíduos que pertencem à classe j da k-ésima variável qualitativa.

I. Generalização da Análise Fatorial de Correspondências : a Análise Fatorial de Correspondências Múltiplas

14 - Um instrumento de caracterização dos eixos fatoriais : os «valores de prova»

b) Variáveis quantitativas características de uma classe (continuação)

Sob a hipótese que os valores da classe que nos interessa tenham sido selecionados ao acaso :

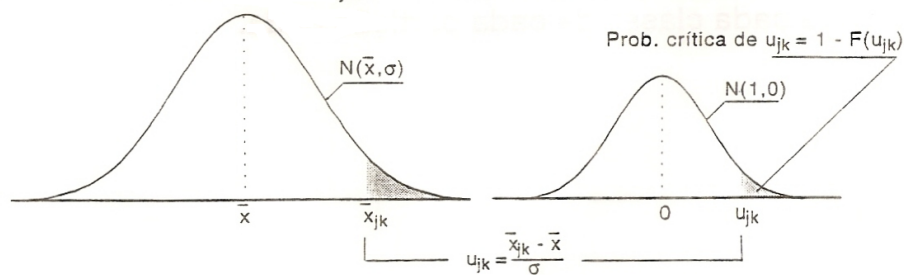
a média do «estimador de \bar{x} » : tende ao valor de \bar{x} : $E_{H_0}(M_k) = \bar{x}$,

- a variância do «estimador de \bar{x} » : $Var_{H_0}(M_k) = \frac{s^2}{n_{jk}} \times \frac{n - n_{jk}}{n - 1} = \sigma^2$

A variável aleatória M_k se ajusta a uma variável normal de parâmetros : $M_k \equiv N(\bar{x}, s)$.

Centrando e reduzindo a variável aleatória M_k podemos definir uma variável U , que será uma variável normal reduzida : $U_k = \frac{M_k - \bar{x}}{\sigma} \equiv N(0, 1)$.

Se n e n_{jk} são suficientemente grandes, pode-se demonstrar que a probabilidade associada à ocorrência de um valor dado de M_k , é igual à probabilidade que uma lei normal ultrapasse o valor U_{jk} calculado a partir da « média da amostra » dos n_{jk} indivíduos da classe que nos interessa.



O valor $u_{jk} = \frac{\bar{x}_{jk} - \bar{x}}{\sigma}$ é chamado « valor de prova »

I. Generalização da Análise Fatorial de Correspondências : a Análise Fatorial de Correspondências Múltiplas

14 - Um instrumento de caracterização dos eixos fatoriais : os «valores de prova»

b) Variáveis quantitativas características de uma classe (continuação)

Se o «valor de prova» é $u_{jk} = \frac{\bar{X}_{jk} - \bar{X}}{\sigma}$.

A distância entre a média da p-ésima variável quantitativa e a média da mesma na classe j da k-ésima partição, é avaliada em termos da «quantidade de desvios padrões de uma variável normal» :

Se o «valor de prova» é positivo, a classe considerada apresenta valores elevados da variável quantitativa levada em conta.

Se o «valor de prova» é negativo, a classe considerada se caracteriza por apresentar valores muito baixos da variável quantitativa levada em conta.

Dispomos, assim, de um índice que permite as comparações entre as p variáveis quantitativas consideradas.

Podemos ordenar as p variáveis segundo a importância do «valor de prova» para cada classe de cada partição dada.

I. Generalização da Análise Fatorial de Correspondências : a Análise Fatorial de Correspondências Múltiplas

14 - Um instrumento de caracterização dos eixos fatoriais : os «valores de prova»

c) Atributos típicos de uma classe

Seja uma classe de indivíduos resultante de uma partição definida pela k-ésima variável qualitativa da tabela.

Estabelecemos a hipótese estatística seguinte :

H_0 : os n_{jq} indivíduos foram selecionados ao acaso, sem reposição, entre os n indivíduos observados, tirando uma amostra aleatória de n_{jk} indivíduos. Ou seja, que $H_0: (n_{jq} / n_{jk}) = (n_{.q} / n)$;

H_1 : os n_{jq} indivíduos foram selecionados ao acaso, sem reposição, entre os n indivíduos observados, tirando uma amostra aleatória de n_{jk} indivíduos. Ou seja, $H_1: (n_{jq} / n_{jk}) = (n_{.q} / n)$;

Quanto maior for (n_{jq} / n_{jk}) com respeito a $(n_{.q} / n)$ mais inadmissível será a hipótese que os n_{jq} indivíduos tenham sido extraídos ao acaso dentre os n indivíduos.

Pode-se demonstrar que :

- se se selecionam n_{jk} indivíduos em uma população de n indivíduos de maneira tal que todos eles tenham a mesma probabilidade de serem escolhidos ;

- se essa seleção é feita sem reposição (já que cada indivíduo recebe uma e somente uma modalidade da q-ésima variável qualitativa);

- e se a observação de um indivíduo consiste em registrar o fato que este possui ou não possui o atributo q da q-ésima variável qualitativa;

A quantidade (n_{jq}) de indivíduos que possuem o atributo q numa amostra de n_{jk} indivíduos constitui uma variável aleatória H_k que admite uma distribuição de probabilidades hipergeométrica especificada por três parâmetros: n , n_{jk} , $p = (n_{.q} / n)$.

I. Generalização da Análise Fatorial de Correspondências : a Análise Fatorial de Correspondências Múltiplas

14 - Um instrumento de caracterização dos eixos fatoriais : os «valores de prova»

c) Atributos típicos de uma classe (continuação) :

$$\text{Pode-se demonstrar que : } P(H_k = n_{jq}) = \frac{\binom{n_{.q}}{n_{jq}} \binom{n_{jk} - n_{.q}}{n - n_{jq}}}{\binom{n_{jk}}{n}}$$

A probabilidade de observar n_{jq} indivíduos, na classe de n_{jk} indivíduos que nos interessa, é : $P_{jk}(n_{jq}) = \text{Prob}_{H_0}(H_k \geq n_{jq})$.

Quanto maior for a quantidade n_{jq} observada, menor será a probabilidade hipergeométrica associada a ela e mais duvidosa será a hipótese H_0 , pela qual a amostra seria produzida de maneira aleatória.

Pode-se utilizar a probabilidade hipergeométrica associada a cada uma das modalidades das variáveis qualitativas da tabela para ordenar todos os atributos observados em ordem crescente de suas probabilidades críticas para uma classe determinada de indivíduos.

Os atributos que apresentam as probabilidades hipergeométricas menores, são os atributos mais típicos da classe considerada.

A fim de dispor de um índice homogêneo de comparação das variáveis quantitativas e qualitativas, podemos utilizar a probabilidade hipergeométrica crítica associada a um atributo dado — para uma classe de indivíduos — como valor de referência que pode ser utilizado em relação com a distribuição de probabilidades da lei normal.

Para uma classe dada de indivíduos, o «valor de prova» de cada modalidade, de cada uma das q -variáveis qualitativas da tabela, é igual ao valor de uma variável normal que tenha a mesma probabilidade de ser ultrapassado que a probabilidade hipergeométrica associada a cada modalidade.

Os «valores de prova» permitem classificar por ordem de importância os atributos típicos de cada classe.

I. Generalização da Análise Fatorial de Correspondências : a Análise Fatorial de Correspondências Múltiplas

14 - Um instrumento de caracterização dos eixos fatoriais : os «valores de prova»

d) Exploração da tabela de coordenadas fatoriais em relação com a tabela TCC

- O cálculo dos «valores de prova» supõe que se adote a hipótese de uma seleção aleatória das n_j coordenadas dos indivíduos ao longo de cada eixo fatorial.

Esta hipótese não é respeitada, por construção dos eixos fatoriais, no caso das variáveis ativas da análise.

- Os « valores de prova » devem ser considerados como índices de ajuda à interpretação das modalidades ilustrativas (suplementares) da análise.

Tratando-se de uma modalidade ativa, porém pouco contributiva a um eixo dado, poder-se-á consultar o «valor de prova» correspondente a essa modalidade em um eixo dado.

I. Generalização da Análise Fatorial de Correspondências : a Análise Fatorial de Correspondências Múltiplas

Exemplo de aplicação da A.F.C.M. ao tratamento de dados de pesquisas

Modalidades da escuta de emissoras de radiofonia pela população belga de expressão francesa

1 - Objetivos do estudo

Estudo permanente (R.T.B.F./R.M.B. e B.R.T.N./V.A.R.), realizado por SOBEMAP Marketing, destinado a produzir uma «imagem» fiel de diferentes facetas do modo de escuta da rádio pelas pessoas maiores de 12 anos, residentes na Bélgica.

- Objetivos gerais do estudo:

- controlar o número de pessoas que ouvem a rádio, as estações que escutam e as horas de escuta ;
- descrever os «ouvintes» segundo suas características sócio-demográficas ;
- analisar os comportamentos de escuta durante um período relativamente amplo a fim de descrever os «hábitos» de escuta ;

Observação permanente dos ouvintes de rádio por meio de um painel e de um registro de escuta que é distribuído às pessoas selecionadas no painel.

O registro de escuta permite registrar, a cada 15 minutos, a rádio ouvida por cada indivíduo. Dessa maneira, é possível calcular a audiência média, por um quarto de hora, das diferentes emissoras de rádio.

O painel é constituído de 2000 pessoas selecionadas em Flandres e outras 2000 pessoas da Wallonie.

O modo de observação escolhido permite uma análise fina da evolução anual, mensal ou semanal do comportamento de escuta da população adulta, residente na Bélgica.