



CIMPA-UCR

Análisis de Correspondencias Múltiples

(ACM)



CIMPA-UCR

ANÁLISIS DE CORRESPONDENCIAS MÚLTIPLES

Lebart, 1974

Hacer análisis de correspondencias sobre:

- Tabla disyuntiva completa
- Tabla de Burt

Representación simultánea de todas las modalidades.

Ojo: # de modalidades

Bibliografía:

Lebart et al. (inglés)

Lebart et al. (español)

Greenacre (inglés)

Escofier & Pagès (francés)



ANÁLISIS DE CORRESPONDENCIAS MÚLTIPLES

La tabla de datos:

ind	SEXO	EDAD	INGRESO
1	F	5	Medio
2	F	3	Alto
3	M	4	Bajo
4	F	1	Bajo
5	F	2	Medio
6	M	5	Alto
7	F	2	Medio
8	M	3	Bajo
9	M	1	Alto
10	F	4	Medio

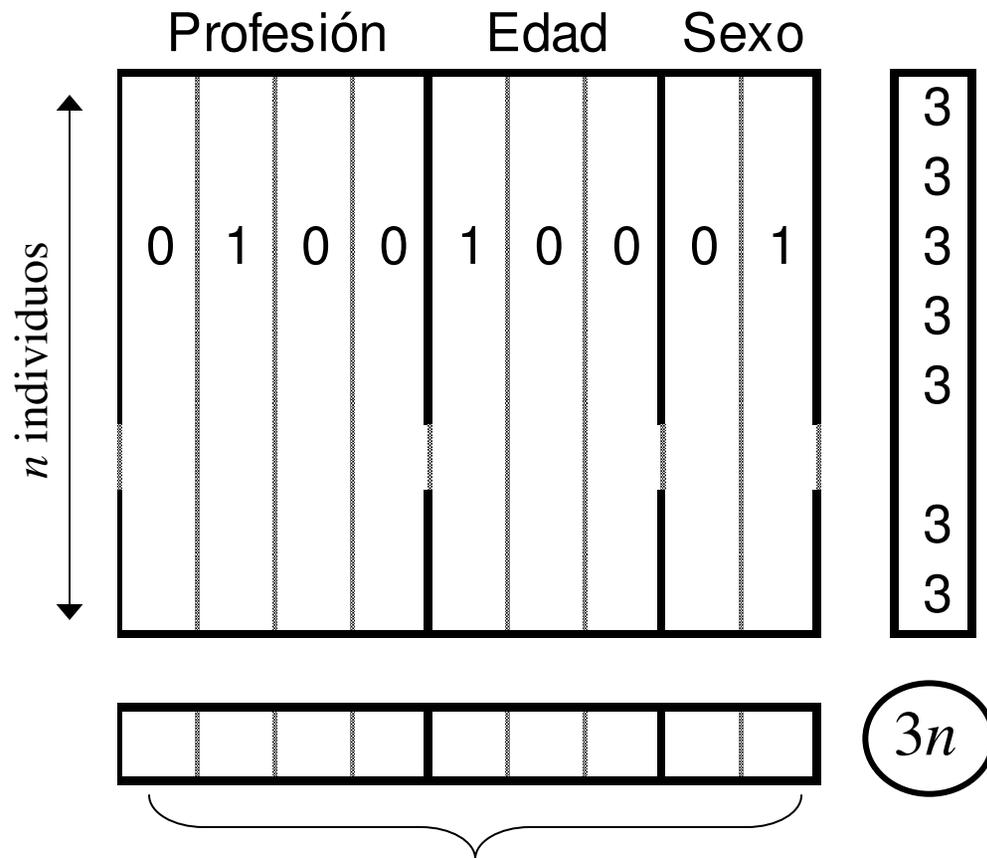
Tabla de Burt

		Sexo		Edad					Ingresos		
		F	M	1	2	3	4	5	B	M	A
Sexo	F	6		1	2	1	1	1	1	4	1
	M		4	1	0	1	1	1	2	0	2
Edad	1	1	1	2					1	0	1
	2	2	0		2				0	2	0
	3	1	1			2			1	0	1
	4	1	1				2		1	1	0
	5	1	1					2	0	1	1
Ingresos	B	1	2						3		
	M	4	0							4	
	A	1	2								3
		18	12								



ANÁLISIS DE CORRESPONDENCIAS MÚLTIPLES

ACM \Leftrightarrow Análisis de Correspondencias de una tabla disyuntiva completa



Estructura particular de la tabla

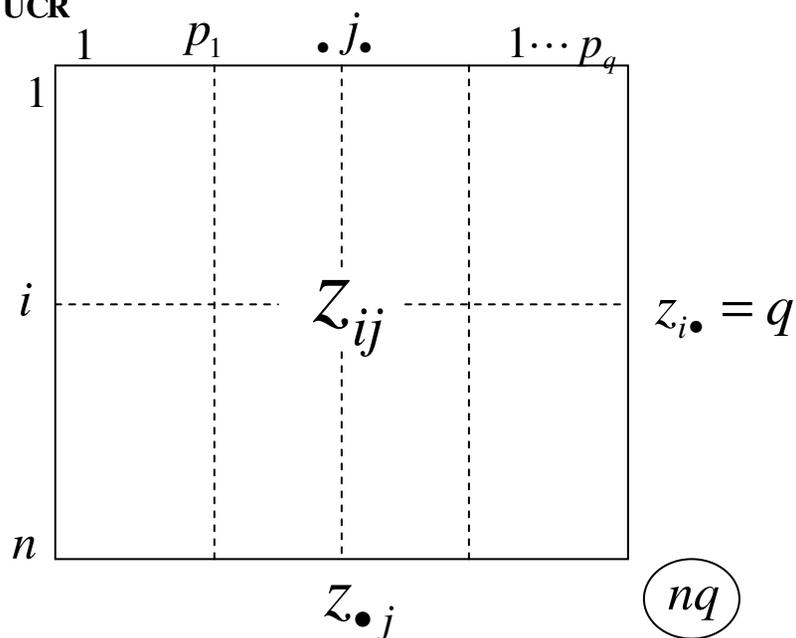


Propiedades particulares del análisis



CIMPA-UCR

ACM: AFC DE UNA TABLA DISYUNTIVA



$$Z = (z_{ij}) \quad z_{ij} = \begin{cases} 1 \\ 0 \end{cases}$$

Márgenes:

$z_{i\bullet} = q$ pues hay q unos en la fila i

$z_{\bullet j}$ # individuos que tienen modalidad j

$$\sum_j \sum_i z_{ij} = nq$$

$n = \#$ individuos

$q = \#$ variables cualitativas

$p_k = \#$ modalidades de la variable k

$$p = \sum_{k=1}^q p_k = \text{total de columnas}$$

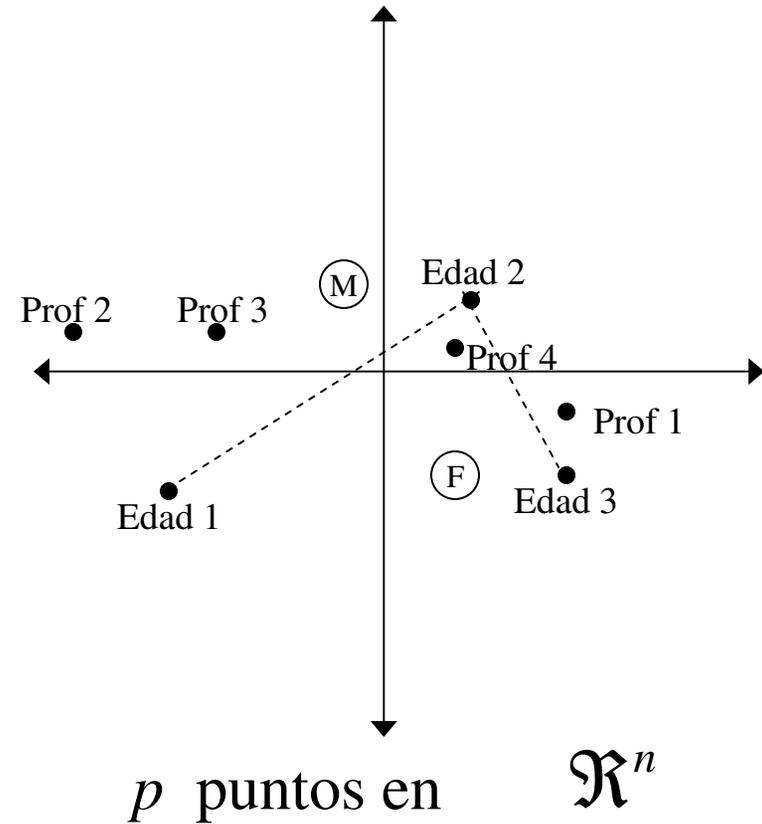
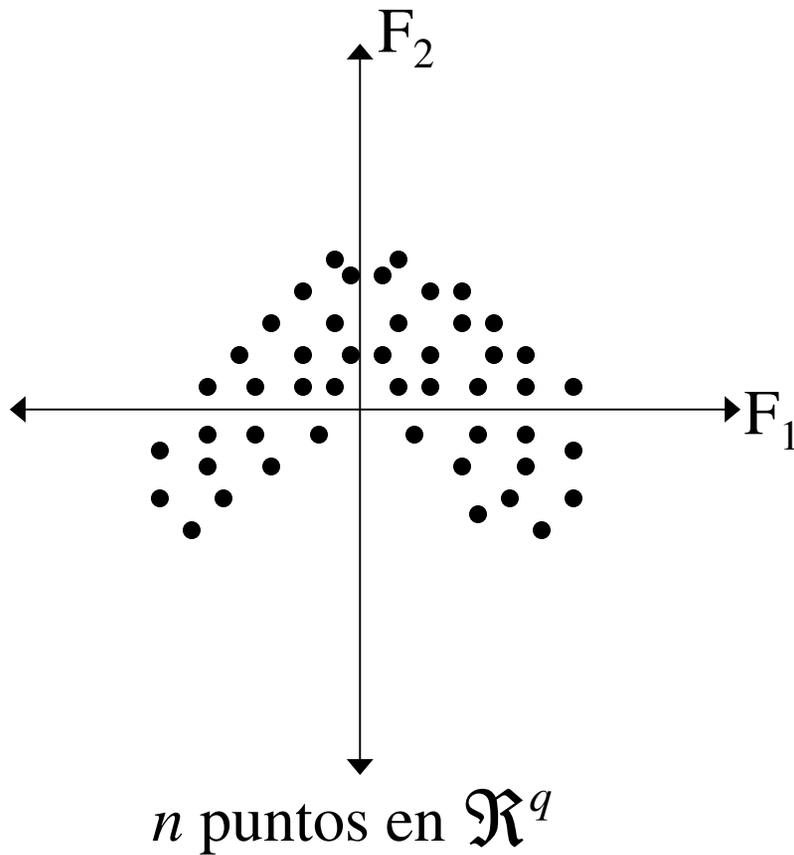
Matriz a diagonalizar:

$$S = \frac{1}{q} Z^t Z D^{-1} = \frac{1}{q} B D^{-1}$$

con $D = \text{diag}(z_{\bullet j})$



ANÁLISIS DE CORRESPONDENCIAS MÚLTIPLES





CIMPA-UCR

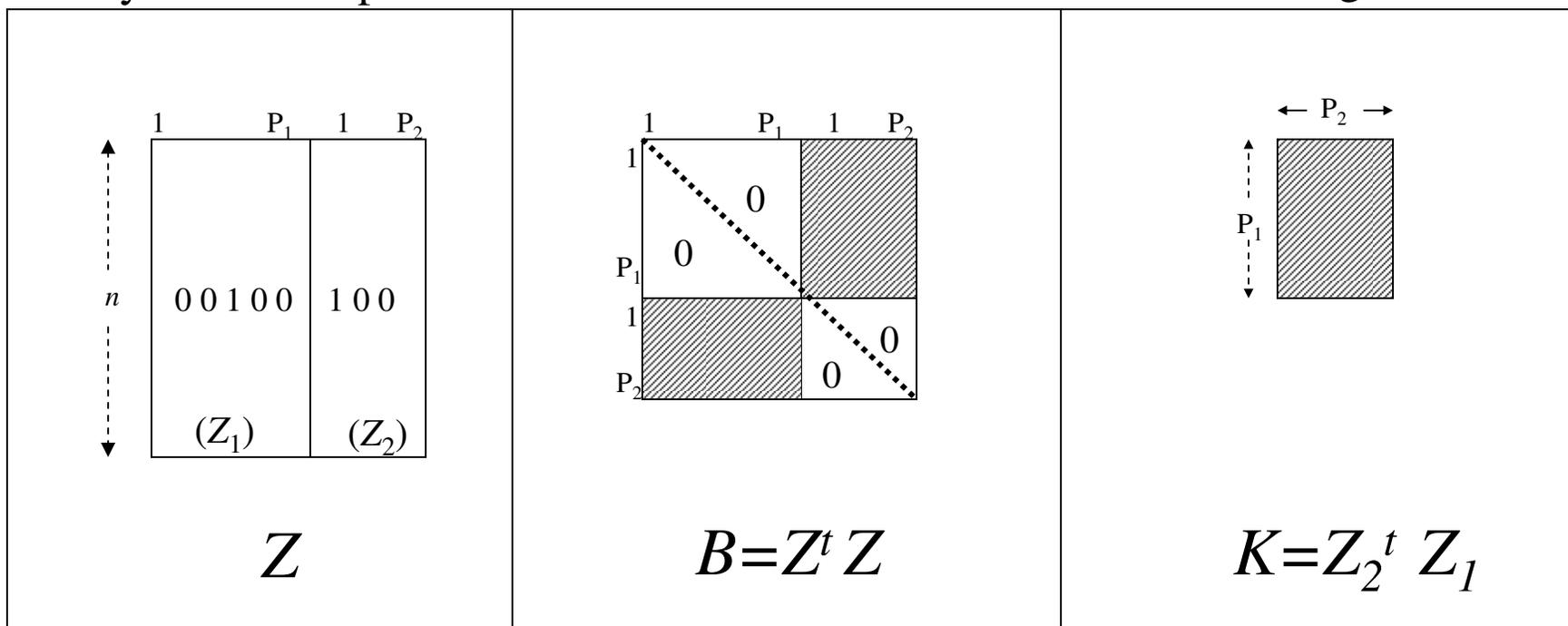
ACM: CASO DE 2 VARIABLES

→ Análisis factorial de correspondencias

Disyuntiva completa

Burt

Contingencia



$$\lambda_z$$

$$\lambda_B = \lambda_z^2$$

$$\lambda_K = (2\lambda_z - 1)$$



CIMPA-UCR

ACM: NUBES DE PUNTOS - PERFILES

Perfiles- fila (individuos)

Puntos (coordenadas): $\frac{Z_{ij}}{Z_{i\bullet}} = \frac{Z_{ij}}{q}$ Tabla: $\frac{1}{q}Z$

Peso: $\frac{Z_{i\bullet}}{nq} = \frac{q}{nq} = \frac{1}{n}$

Métrica: $diag\left(\frac{nq}{Z_{\bullet j}}\right)$

Distancia χ^2 : $d^2(i, i') = \frac{n}{q} \sum_{j=1}^p \frac{1}{Z_{\bullet j}} (Z_{ij} - Z_{i'j})^2$

Perfiles- columna (modalidades)

Puntos (coordenadas): $\frac{Z_{ij}}{Z_{\bullet j}}$

Peso: $\frac{Z_{\bullet j}}{nq}$

Métrica: $diag\left(\frac{nq}{Z_{i\bullet}}\right) = diag\left(\frac{nq}{q}\right) = diag(n)$

Distancia χ^2 : $d^2(j, j') = n \sum_{i=1}^n \left(\frac{Z_{ij}}{Z_{\bullet j}} - \frac{Z_{ij'}}{Z_{\bullet j'}}\right)^2$



CIMPA-UCR

ACM: DISTANCIA DE χ^2

Perfiles de fila:

$$d^2(i, i') = \frac{n}{q} \sum_{j=1}^p \frac{1}{Z_{\bullet j}} (Z_{ij} - Z_{i'j})^2 = \frac{n}{q} \sum_{j \in M_{ii'}} \frac{1}{Z_{\bullet j}}$$

con $M_{ii'}$: modalidades que tiene solo un individuo i o i'
Más parecidos si tienen más modalidades en común.

Perfiles de columna:

$$d^2(j, j') = n \sum_{i=1}^n \left(\frac{Z_{ij}}{Z_{\bullet j}} - \frac{Z_{ij'}}{Z_{\bullet j'}} \right)^2 = n \frac{\text{card}(\text{ind}[j, \text{no } j']) + \text{card}(\text{ind}[j', \text{no } j])}{Z_{\bullet j} Z_{\bullet j'}}$$

Entre más objetos tengan sólo una de j o j' mayor es la distancia



INTERPRETACIÓN

- Dos modalidades escogidas por los mismos individuos coinciden
- Dos individuos son cercanos si escogen las mismas modalidades
- Modalidades con poco efectivo están alejadas del centro de gravedad



CIMPA-UCR

ACM: INERCIA

Centro de gravedad de nube de modalidades $G = \left(\frac{1}{n}, \dots, \frac{1}{n} \right)$

$$d^2(j, G) = n \sum_{i=1}^n \left(\frac{Z_{ij}}{Z_{\bullet j}} - \frac{1}{n} \right)^2 = n \sum_{i=1}^n \left(\frac{Z_{ij}}{Z_{\bullet j}^2} - \frac{2Z_{ij}}{nZ_{\bullet j}} + \frac{1}{n^2} \right) = \frac{n}{Z_{\bullet j}} - 1$$

Distancia mayor si el efectivo es pequeño

Inercia de la modalidad j:

$$I(j) = \frac{Z_{\bullet j}}{nq} d^2(j, G) = \frac{Z_{\bullet j}}{nq} \left(\frac{n}{Z_{\bullet j}} - 1 \right) = \frac{1}{q} \left(1 - \frac{Z_{\bullet j}}{n} \right)$$

Mayor inercia si el efectivo es pequeño



CIMPA-UCR

ACM: INERCIA

Inercia de la variable k :

$$I(k) = \sum_{j=1}^{p_k} I(j) = \sum_{j=1}^{p_k} \frac{1}{q} \left(1 - \frac{Z_{\bullet j}}{n} \right) = \frac{1}{q} (p_k - 1)$$

Crece con el número de modalidades

Si $p_k = 2 \longrightarrow$ mínimo $\frac{1}{q}$

Inercia total:

$$I = \sum_k I(k) = \sum_k \frac{1}{q} (p_k - 1) = \frac{1}{q} (p - q) = \frac{p}{q} - 1$$

No tiene significado estadístico



CIMPA-UCR

ACM: SOLUCIÓN

Diagonalización: En \mathcal{R}^p :
$$\frac{1}{q} Z^t Z D^{-1} u_\alpha = \lambda_\alpha u_\alpha$$

factor $\varphi_\alpha = D^{-1} u_\alpha \Rightarrow \frac{1}{q} D^{-1} Z^t Z \varphi_\alpha = \lambda_\alpha \varphi_\alpha$

En \mathcal{R}^n :
$$\frac{1}{q} Z D^{-1} Z^t \psi_\alpha = \lambda_\alpha \psi_\alpha$$

Relaciones de transición:
$$\varphi_\alpha = \frac{1}{\sqrt{\lambda_\alpha}} D^{-1} Z^t \psi_\alpha$$

$$\psi_\alpha = \frac{1}{q \sqrt{\lambda_\alpha}} Z \varphi_\alpha$$



CIMPA-UCR

ACM: SOLUCIÓN

Relaciones baricéntricas:

$$\psi_{\alpha i} = \frac{1}{\sqrt{\lambda_{\alpha}}} \sum_{j=1}^p \frac{Z_{ij}}{Z_{i\bullet}} \varphi_{\alpha j} = \frac{1}{q\sqrt{\lambda_{\alpha}}} \sum_{j \in P(i)} \varphi_{\alpha j}$$

$$\varphi_{\alpha j} = \frac{1}{\sqrt{\lambda_{\alpha}}} \sum_{i=1}^n \frac{Z_{ij}}{Z_{\bullet j}} \psi_{\alpha i} = \frac{1}{Z_{\bullet j} \sqrt{\lambda_{\alpha}}} \sum_{i \in I(j)} \psi_{\alpha i}$$

$P(i)$ modalidades que tiene i

$I(j)$ modalidades que tiene j



CIMPA-UCR

ACM: INTERPRETACIÓN

- Proximidad entre individuos en términos de parecido:
 - Dos individuos se parecen si tienen casi las mismas modalidades:
- Proximidad entre modalidades de variables diferentes en términos de asociación:
 - Son cercanos puesto que globalmente están presentes en los mismos individuos
- Proximidad entre modalidades de una misma variable en términos de parecido:
 - Son excluyentes por construcción
 - Si son cercanas es porque los individuos que las poseen presentan casi el mismo comportamiento en las otras variables



CONCEPTOS SOCIOLÓGICOS EN LA EDUCACIÓN C.R.

Opinión: var 1, ... , var 5/var 6, ... , var 10

Señalización: sexo, tipo colegio, ingr, reli, poli, edpa, edma

Libertad y empresa privada

- Var 1. La empresa privada es necesaria para la libertad
- Var 2. Se puede entrar en un sindicato sin ser perseguido
- Var 3. Si se trabaja en una e.p. no hay tiempo para ocuparse de la familia
- Var 4. Los trabajadores deben respetar las órdenes de su patrono sin criticarlas
- Var 5. Mi padre debe dedicarle más tiempo a la empresa o institución donde trabaja

Democracia y Ejército

- Var 6. Un país democrático no necesita ejército
- Var 7. Un país con ejército es un país totalitario
- Var 8. Es necesario que C.R. tenga ejército para defenderse de amenazas externas
- Var 9. La existencia de un ejército en C.R. podría llevar a una dictadura militar a corto plazo
- Var 10. E.E. U.U. no es un país democrático porque tiene ejército



CIMPA-UCR

ACM Ejemplo

- *edpa*: ninguna educación y educación primaria (1+33).
- *poli*: partido Vanguardia Popular, partido del Pueblo Costarricense y otros (3+1+5).
- *var6*: *des6* y *fde6* (31+9).
- *var2*: *dac2* y *fac2* (19+5).
- *var3*: *fde3* y *des3* (1+9).
- *var4*: *fde4* y *des4* (7+51).
- *var5*: *fde5* y *des5* (12+29).
- *var9*: *des9* y *fde9* (9+25).



CIMPA-UCR

Análisis de Correspondencias Múltiples

ACM Ejemplo

Variable	Modalidades	código	frec.	%
Tipo de colegio	público	publ	124	53
	privado	priv	108	47
Sexo	masculino	masc	101	44
	femenino	feme	131	56
Nivel de ingreso de la familia (en miles de colones)	$I < 7$	I1	24	10
	$7 \leq I < 14$	I2	44	19
	$14 \leq I < 21$	I3	35	15
	$21 \leq I < 28$	I4	30	13
	$28 \leq I < 35$	I5	34	15
	$35 \leq I$	I6	47	20
	no responde	I7	18	8
Grado de educación del padre	primaria incompleta	P1	35	15
	primaria completa	P2	30	13
	secundaria incompleta	P3	34	15
	secundaria completa	P4	35	16
	universit. incompleta	P5	29	12
	universit. completa	P6	68	29
Grado de educación de la madre	primaria incompleta	M1	37	16
	primaria completa	M2	41	18
	secundaria incompleta	M3	43	18
	secundaria completa	M4	41	17
	universit. incompleta	M5	18	8
	universit. completa	M6	52	23
Religión	católica	cato	127	55
	evangélica	evan	75	32
	no creyente	nocr	8	3
	otra	otra	22	10
Partido político	PUSC	pusc	82	35
	PLN	pln	129	56
	otro	otro	8	4
	no responde	nore	12	5



CIMPA-U

Análisis de Correspondencias Múltiples

ACM Ejemplo

Cód.	Variable	Mod.	frec.	%
<i>var1</i>	La empresa privada es necesaria para que exista libertad	<i>fde1</i>	60	26
		<i>des1</i>	74	32
		<i>ind1</i>	39	17
		<i>dac1</i>	32	14
		<i>fac1</i>	27	11
<i>var2</i>	Se puede entrar en un sindicato sin ser perseguido o despedido	<i>fde2</i>	60	26
		<i>des2</i>	119	51
		<i>ind2</i>	29	13
		<i>dac2</i>	24	10
<i>var3</i>	Si se trabaja en la empresa privada no hay tiempo para ocuparse de sus propias cosas	<i>des3</i>	10	4
		<i>ind3</i>	36	16
		<i>dac3</i>	145	62
		<i>fac3</i>	41	18
<i>var4</i>	Los trabajadores deben respetar las órdenes de su patrono sin criticarlas	<i>des4</i>	58	25
		<i>ind4</i>	22	9
		<i>dac4</i>	123	53
		<i>fac4</i>	29	13
<i>var5</i>	Mi padre debe dedicarle más tiempo a la empresa o institución donde trabaja	<i>des5</i>	32	18
		<i>ind5</i>	93	40
		<i>dac5</i>	60	26
		<i>fac5</i>	37	16



CIMPA-U

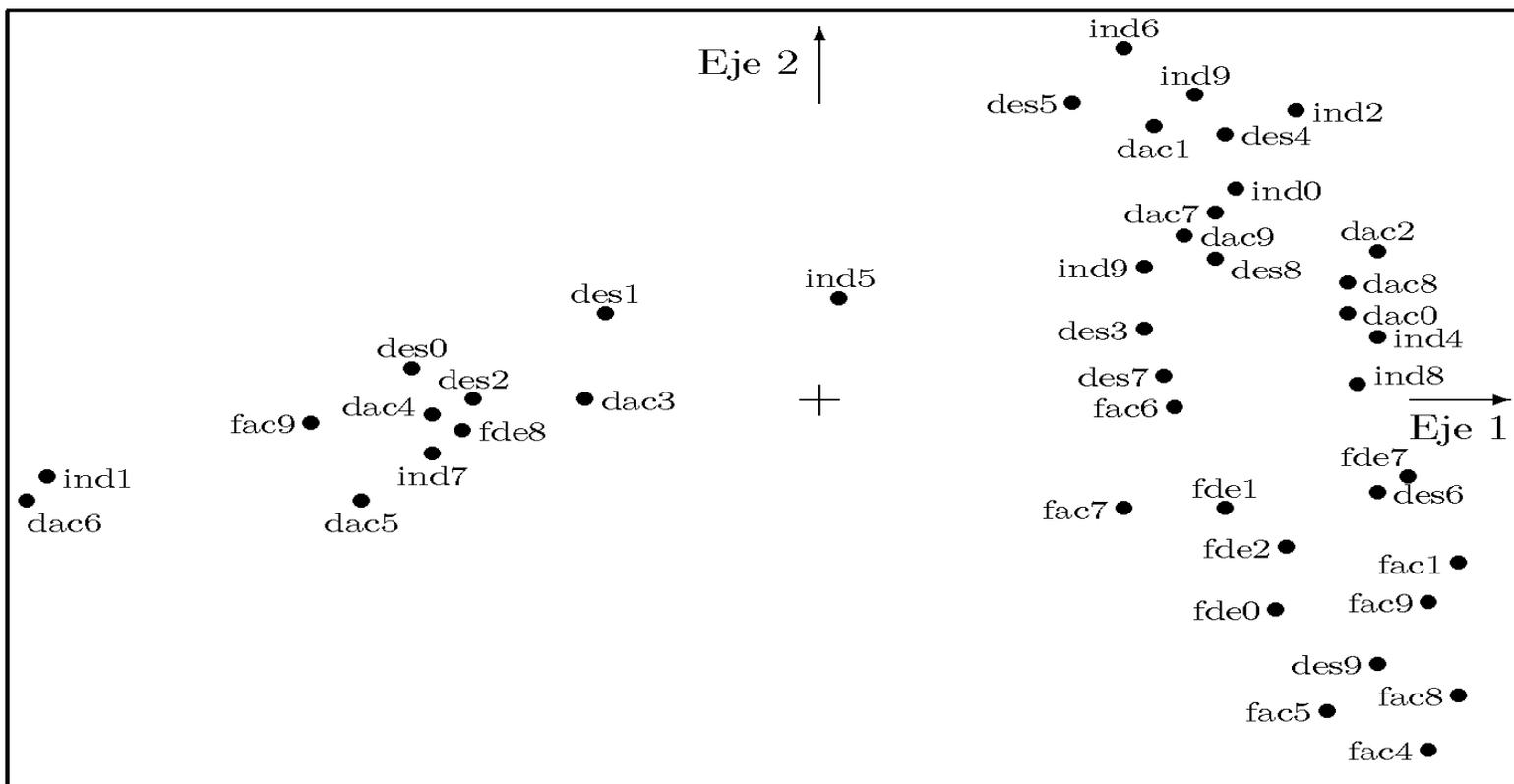
ACM Ejemplo

<i>var6</i>	Un país democrático no necesita ejército	<i>des6</i> <i>ind6</i> <i>dac6</i> <i>fac6</i>	40 24 56 112	17 11 24 48
<i>var7</i>	Un país con ejército es un país totalitario	<i>fde7</i> <i>des7</i> <i>ind7</i> <i>dac7</i> <i>fac7</i>	19 45 106 44 18	8 19 46 19 8
<i>var8</i>	Es necesario que Costa Rica tenga ejército para defenderse de amenazas externas	<i>fde8</i> <i>des8</i> <i>ind8</i> <i>dac8</i> <i>fac8</i>	15 23 36 41 117	6 10 16 18 50
<i>var9</i>	La existencia de un ejército en Costa Rica podría llevar a una dictadura...	<i>des9</i> <i>ind9</i> <i>dac9</i> <i>fac9</i>	97 67 34 34	41 29 15 15
<i>var10</i>	Estados Unidos no es un país democrático porque tiene ejército	<i>fde0</i> <i>des0</i> <i>ind0</i> <i>dac0</i>	23 41 118 50	10 18 51 21



CIMPA-UCR

ACM Ejemplo





CIMPA-UCR

ACM Ejemplo

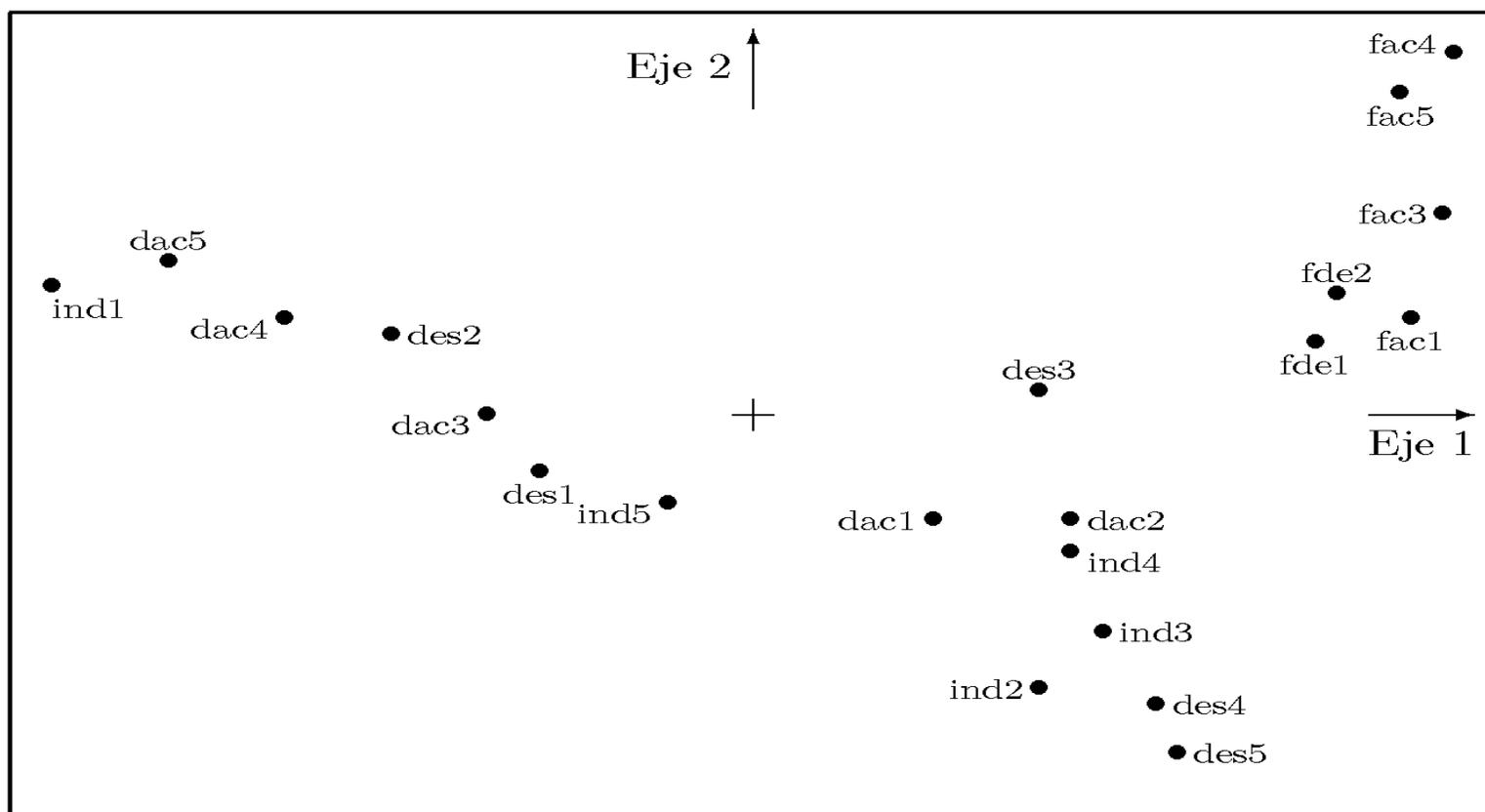
3.2 ACM de las preguntas “empresa privada–libertad”

Al realizar el ACM entre las modalidades de las variables 1 a 5 se obtiene como plano principal el mostrado en la Figura 2. En ésta se ve claramente la fuerte relación directa que hay entre var_4 y var_5 , observándose también la trayectoria ordenada de la que antes hablábamos. En el plano generado por el primer y tercer ejes principales se confirma esta relación (sin observar la trayectoria), y como aquí, no se ve ninguna otra relación importante.



CIMPA-UCR

ACM Ejemplo





CIMPA-UCR

ACM Ejemplo

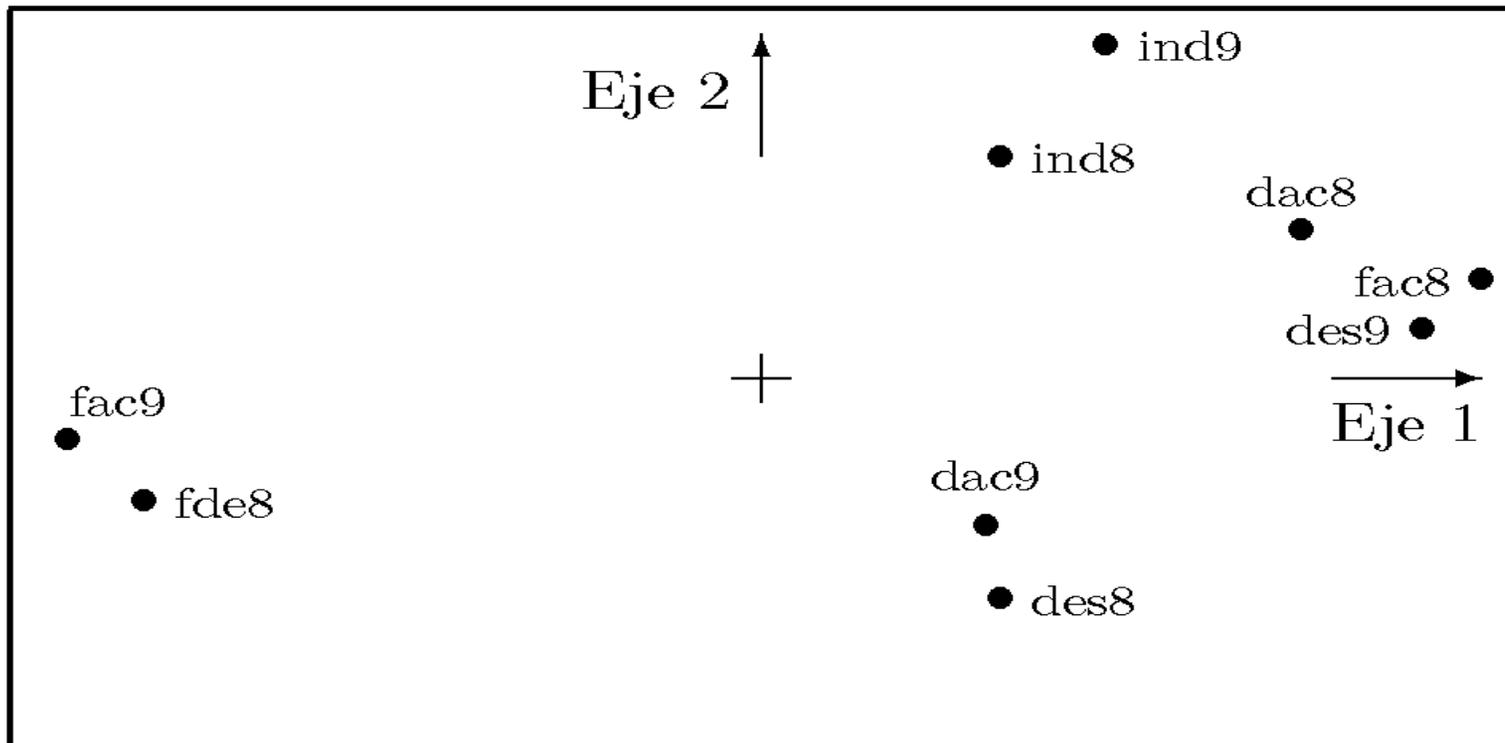
3.3 ACM de las preguntas “democracia–ejército”

Al hacer el análisis entre las variables $var6$ a $var10$, se confirma el agrupamiento de algunas modalidades de $var8$ y $var9$, pero en cierto desorden, y no es sino en el gráfico que cruza los ejes principales 1 y 3 en que se puede ver una cierta trayectoria, con $dac8$ mejor ubicado para que ésta tenga un sentido, como se ilustra en la Figura 3.



CIMPA-UCR

ACM Ejemplo





CIMPA-UCR

ACM Ejemplo

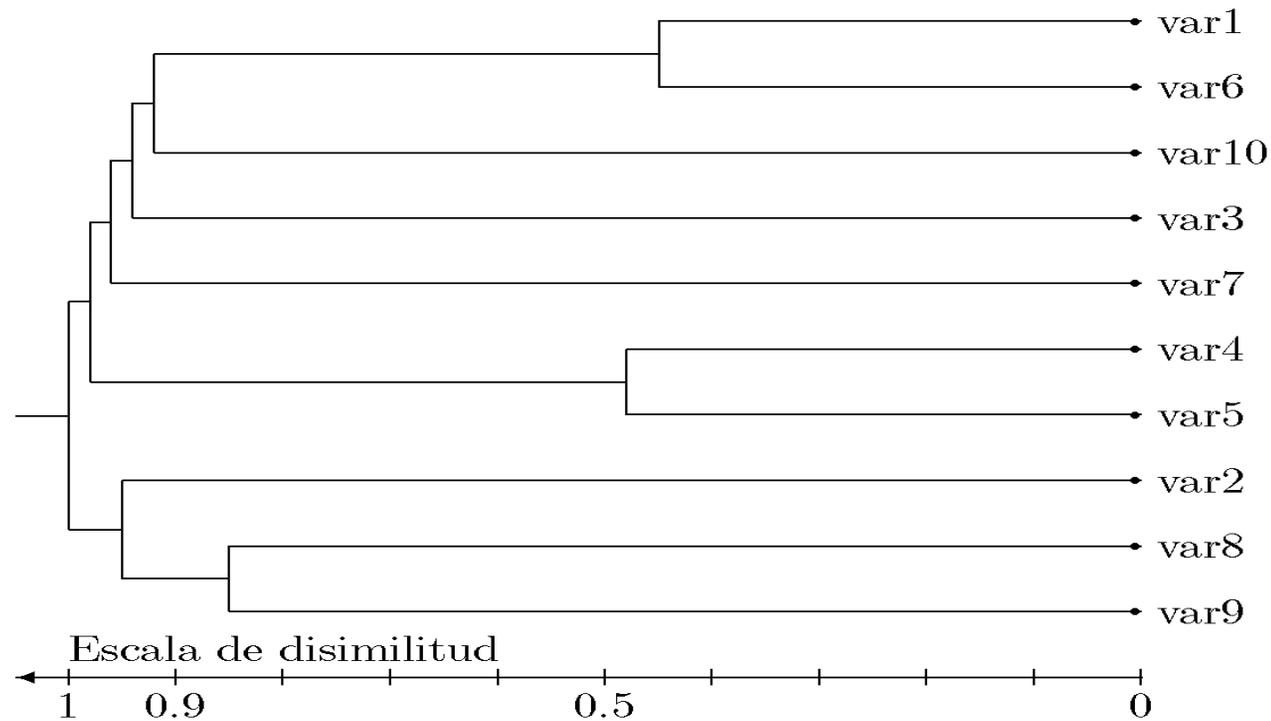


Figure 4: Arbol de clasificación para las variables de opinión.



CIMPA-UCR

ACM Ejemplo

- Clasificación de variables cuya asociación es medida con el T de Chuprov

$$T(a, b) = \frac{\chi^2}{(p - 1)(q - 1)},$$



CIMPA-UCR

ACM Ejemplo

1. La relación directa entre “los trabajadores deben respetar las órdenes de su patrono sin criticarlas” (*var4*) y “mi padre debe dedicarle más tiempo a la empresa o institución donde trabaja” (*var5*) denota una fuerte concordancia sumisión-explotación.
2. La relación inversa entre “la empresa privada es necesaria para que exista libertad” (*var1*) y “un país democrático no necesita ejército” (*var6*) concuerda con la fuerte propaganda que salía en la prensa en la época anterior en que se pasó la encuesta, que hacía énfasis en que la empresa privada produce libertad, y que insistía a la vez en que Costa Rica debía armarse para preservar su democracia. Aquí se muestra la gran permeabilidad de los estudiantes de secundaria a la propaganda masiva e ideológica.



CIMPA-UCR

ACM Ejemplo

3. La relación inversa entre “Costa Rica necesita ejército para defenderse de amenazas externas” (*var8*) y “un ejército en Costa Rica podría llevar a una dictadura militar a corto plazo” (*var9*) parece muy lógica: quienes quieren la implantación de un ejército en Costa Rica hacen creer que en nuestro país no se correría el riesgo de una dictadura militar.
4. La ausencia de otras relaciones entre las otras preguntas puede ser debida a la falta de consistencia en la opinión de los estudiantes de educación media.
5. Respecto de la metodología estadística empleada, que dicho sea de paso tiene un carácter descriptivo por excelencia, cabe mencionar la importancia de los gráficos obtenidos a partir del ACM, en la medida que nos permitieron explorar, no sólo las asociaciones entre variables, sino también su carácter (directo o inverso), propiedad que no es posible observar mediante el clásico chi-cuadrado.