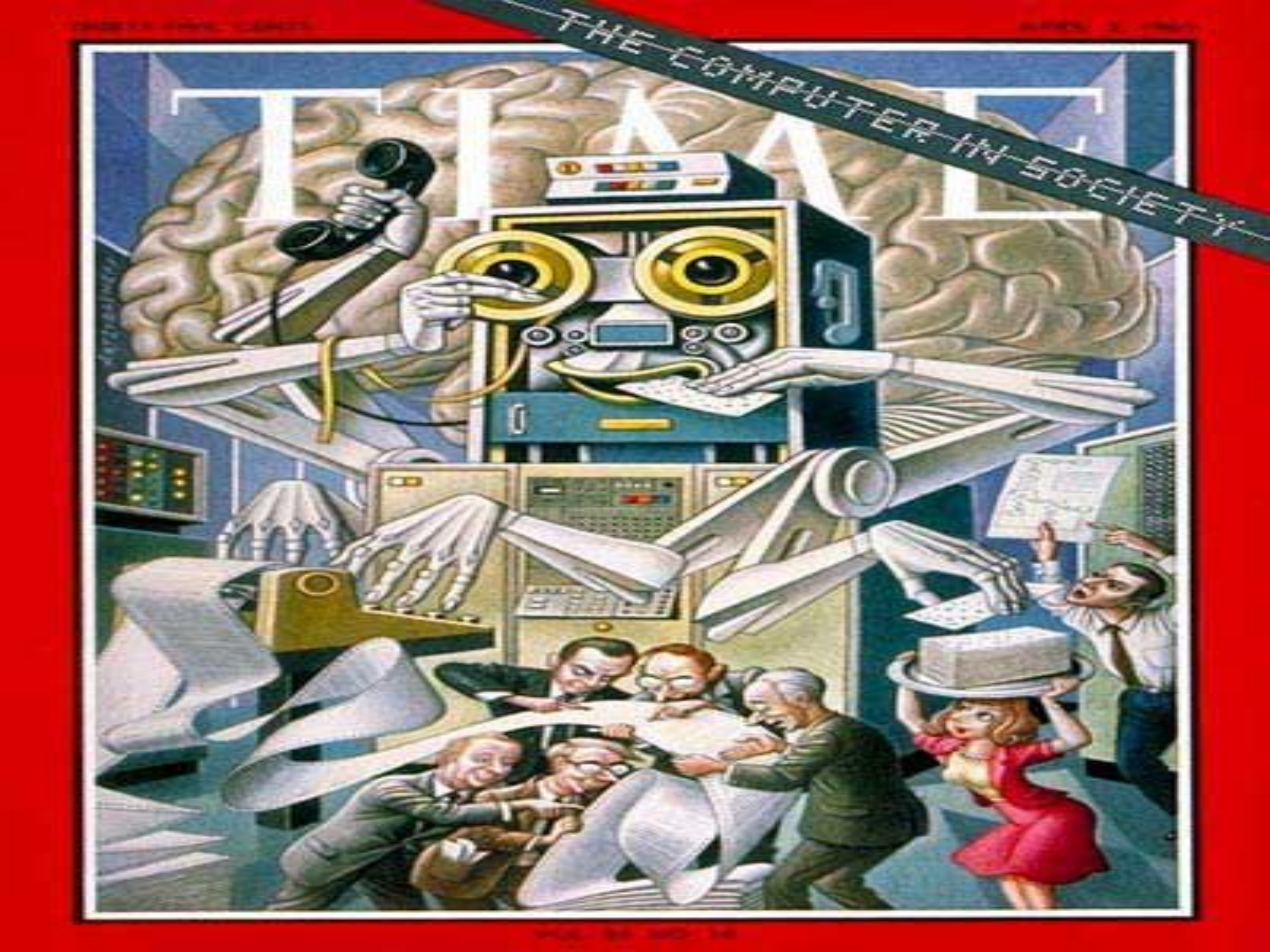
TIME
THE COMPUTER IN SOCIETY

Tema 7

Los modelos Log-lineales

Los modelos Log-lineales

- Introducción a los modelos LOG-LINEALES
- Fases para la elaboración de los modelos LOG-LINEALES
- Principales modelos LOG-LINEALES
  - El modelo saturado
  - El modelo de independencia
  - El modelo jerárquico
- Criterios de selección del modelo:
  - Principio de parsimonia
  - Significación estadística
  - Interpretación sustantiva

Los modelos Log-lineales

- Los modelos LOG-LINEALES tienen como objetivo el análisis de las relaciones entre variables cualitativas representadas en tablas de contingencia multidimensionales.
- Los modelos LOG-LINEALES resumen el proceso por muy complejo que sea en una serie de componentes llamados parámetros Lambda. De este modo una tabla de contingencia de dos variables presenta cuatro efectos:
  - El efecto de las filas
  - El efecto de las columnas
  - El efecto de la interacción entre las variables
  - El efecto debido al promedio de la casilla

Los modelos Log-lineales

- Para elaborar un modelo LOG-LINEAL, las frecuencias observadas en cada casilla se transforman en sus logaritmos naturales, con lo que el modelo multiplicativo se convierte en un modelo aditivo ya que:

- log (a * b) = log a + log b

(recordemos que el logaritmo de un número es la potencia a la que hay que elevar la base para obtener ese número)

Los modelos Log-lineales

FASES PARA LA ELABORACIÓN DE LOS MODELOS LOG-LINEALES

1 Selección y especificación de modelos (obtener las frecuencias teóricas a partir de los diferentes modelos posibles o partir de uno dado desde el marco teórico)

2 Prueba de ajuste y evaluación del modelo (comparación de frecuencias observadas y teóricas a partir de Chi-cuadrado o Razón de Verosimilitud)

3 Calculo y estimación de los parámetros (para observar su importancia relativa)

4 Interpretación del modelo y establecimiento de las relaciones

Los modelos Log-lineales

- Ln F ij = μ+ λAi + λBj + λABij

En el caso de tres variables tendríamos:

- Ln F ij = μ + λ Ai + λ Bj + λ Ck + λABij + λACik + λBCjk + λABCijk

Tam Hab	RENTA	DECLARA	NO DECLARA	TOTAL
-100000	R Alta	76	3	79
	R. Baja	394	236	630
+100000	R. Alta	48	2	50
	R. Baja	289	149	438
TOTAL		807	390	1147

Los modelos Log-lineales

RENTA	DECLARA	NO DECLARA	TOTAL
R Alta	124	50	129
R. Baja	683	385	1068
	807	390	1147

RENTA	DECLARA	NO DECLARA	TOTAL
R Alta	4,82	1,61	3,21
R. Baja	6,53	5,95	6,24
	5,67	3,78	4,727

Los modelos Log-lineales

μ= 4,72
λ R alta =4,72 -3,215 = 1,5
λ B baja = 4,72 – 6,24 = -1,5

λ D si = 4,72 – 5,675 = -0,948
λ D no = 4,72 – 3,78 = 0,948

λRD alta si = 4,82 – (4,72 + 1,52 – 0,948) = - 0,471
λRD alta no = 1,61 – (4,72 + 1,51+ 0,948) = -5,577

λRD baja si = 6,53 – (4,72 + 1,52 – 0,948) = 4,236
λRD baja no = 5,95 – (4,72 + 1,51+ 0,948) = 1,787

Ln Fijk = μ + λR alta + λD si + λRD alta si =
4,727 + 1,512 – 0,948 – 0,471 = 4,82

Los modelos Log-lineales

MODELO SATURADO

$Ln\ F\ ij = \mu + \lambda Ri + \lambda Dj + \lambda RDij$

MODELO DE INDEPENDNECIA

$Ln\ F\ ij = \mu + \lambda Ri + \lambda Dj + ... + \lambda Km$

MODELO JERÁRQUICO

En un modelo de más de tres variables, se incluye el término λABC si están también incluidos los términos λA, λB, λC, λAB, λAC y λBC

$Ln\ F\ ij = \mu + \lambda Hi + \lambda Rj + \lambda Dk + \lambda RDjk$

Los modelos Log-lineales

MODELO JERÁRQUICO

Los modelos jerárquicos son los que cumplen la siguiente condición: si hay un término de interacción de un grupo de variables, entonces tiene que haber términos de orden inferior para todas las combinaciones posibles de estas variables. Un modelo de más de tres variables, que incluye el término λABC incluye también λA, λB, λC, λAB, λAC y λBC. Es decir, los modelos jerárquicos se rigen por la regla siguiente: si el parámetro relacionado con un conjunto de variables V, se incluye en el modelo, entonces el modelo debe de incluir todos los parámetros relacionados con cualquier subconjunto de V.

Los modelos Log-lineales

CONCEPTO DE CLASE GENERADORA

La clase generadora expresa de manera sintética los efectos incluidos en el modelo, así, si la clase generadora en un análisis con tres variables es:

(AB) (AC) sabemos que el modelo incluye estos efectos y sus derivados es decir, A, B, C, pero no (CB). Se trata pues de un modelo jerárquico que no incluye interacción entre las tres variables y tampoco todos las posibles interacciones de dos variables

Los modelos Log-lineales

Las frecuencias esperadas en nuestro caso son

	-100000		+100000	
	Si declara	**No declara**	**Si declara**	**No declara**
Renta alta	**58,6**	**2,4**	**65,4**	**2,6**
Renta baja	**421,3**	**238,7**	**261,7**	**148,9**

μ = 4,72

λ Ra =1,5

λ Bb = -1,5

λ D si = -0,94

λ D no = 0,94

λ H -100000 = -0,09

λ H +100000 = 0,09

Los modelos Log-lineales

λ H = 0,277* λ RD = 0,695

λR = -1,536* λ RDH = - 0,181

λ D = 2,020*

λ HR = -0,028

λ HD = -0,205

RENTA

HABITAT

¿ DECLARACIÓN ?

Modelo	Valor de Chi-cuadrado	G.L.	Sig.
$\mu + \lambda H + \lambda R + \lambda D + \lambda RH + \lambda RD$	2,549	2	0,28
$\mu + \lambda H + \lambda R + \lambda D + \lambda HD + \lambda RD$	10,49	2	0,005
$\mu + \lambda H + \lambda R + \lambda D + \lambda RD$	12,59	3	0,006
$\mu + \lambda R + \lambda D + \lambda RD$	61,59	4	0,000

Los modelos Log-lineales

SELECCIÓN DEL MODELO

1 Ningún modelo queda ajustado
2 Sólo un modelo se ajusta
3 Varios modelos se ajustan (criterios de selección)
- Significación estadística
- Principio de parsimonia
- Interpretación sustantiva

Los modelos Log-lineales

DECLARACIÓN

¿

RENTA

?

HABITAT

El modelo obtenido indica que hay relación entre Hábitat y Renta y entre Renta y Declaración y no interacción conjunta entre las tres variables. Es decir, la relación entre Renta y Declaración es igual en los pequeños municipios que en los grandes, y la relación entre Hábitat y Renta es la misma entre los que declaran, que entre los que no declaran

Los modelos Log-lineales

Factor	Level	Label
A	2	
B	2	
C	2	

Factor	Code	Obs	Exp	Resid.	Std. Resi
A	1				
B	1				
C	1	624.5	624.5	.00	.00
C	2	143.5	143.5	.00	.00
B	2				
C	1	7501.5	7501.5	.00	.00
C	2	1746.5	1746.5	.00	.00
A	2				
B	1				
C	1	1413.5	14313.5	.00	.00
C	2	344.5	344.5	.00	.00
B	2				
C	1	7060.5	7060.5	.00	.00
C	2	1817.5	1817.5	.00	.00

	GOODNESS- of - FIT TEST STATISTICS		
Likelihood Ratio	.0000	DF=0	P=1.000
Pearson Chi square	.0000	DF =0	P=1.000

PRUEBA DE LOS EFECTOS DE ORDEN K O SUPERIOR
Tests that K-way and Higer order effects are zero

K	DF	L.R. CHI	Prob	P.CHI	Prob
3	1	.117	.7323	.117	.7318
2	4	395.579	.0000	385.096	.0000
1	7	21884.7	.0000	23901.9	.0000

PRUEBA DE LOS EFECTOS DE ORDEN K
Tests that K-way effects are zero

K	DF	L.R. CHI	Prob	P.CHI	Prob
1	3	21489.2	.0000	23516.8	.0000
2	3	395.462	.0000	384.979	.0000
3	1	.117	.7323	.117	.7318

Los modelos Log-lineales

Test of PARTIAL ASSOCIATIONS

Effect Name	DF	Partial Chi	Prob
A*B	1	388.317	.0000
A*C	1	7.403	.0065
B*C	1	.632	.4268
A	1	18.621	.0000
B	1	13288.387	.0000
C	1	8182.213	.0000

PARAMETROS λ
ESTIMATES FOR PARAMETERS

Paramet	Coeff.	Std.err	Z-value	Low95 CI	Upp95 CI
A*B*C	-.00519	.0145	-.356	-.0337	.0233
A*B	-.2141	.0145	-14.70	-.242	-.185
A*C	.019	.014	1.377	-.008	.048
B*C	.008	.014	.581	-.020	.037
A	-.208	.014	-14.348	-.237	-.180
B	-1.032	.014	-70.85	-1.060	-1.003
C	.712	.014	48.890	.683	.740

Los modelos Log-lineales

BACKWARD ELIMINATIO FOR DESING 1 GENERATING CLASS

A*B*C

Likelihood Ratio chi Square=	.0000	DF=0	P=1.000

If delete Simple effect is	DF	L.R. Chis Change	Prob
A*B*C	1	.117	.7323

STEP 1

The best model has generating class
A*B
A*C
B*C

Likelihood Ratio chi Square=	.117	DF=1	P=.732

If delete Simple effect is	DF	L.R. Chis Change	Prob
A*B	1	388.317	.0000
A*C	1	7.403	.0006
B*C	1	.632	.426

STEP 2

The best model has generating class
A*B
A*C

Likelihood Ratio chi Square=	.748	DF=2	P=.688

If delete Simple effect is	DF	L.R. Chis Change	Prob
A*B	1	387.872	.0000
A*C	1	6.958	.0083

STEP 3

The best model has generating class
A*B
A*C

Likelihood Ratio chi Square=	.748	DF=2	P=.688

The final model has generating class
A*B
A*C

Los modelos Log-lineales

FRECUENCIAS ESPERADAS PARA EL MODELO DE LA CLASE GENERADORA

A*B
A*C

Factor	Code	Obs	Exp	Resid.	Std. Resi
A	1				
B	1				
C	1	624.5	622.3	1.68	.07
C	2	143.5	144.7	-1.68	-1.14
B	2				
C	1	7501.5	7502.7	-1.68	-.02
C	2	1746.5	1744.3	1.68	.04
A	2				
B	1				
C	1	1413.5	1399.9	13.05	.35
C	2	344.5	357.1	-13.05	-.69
B	2				
C	1	7060.5	7073.1	-13.05	-.16
C	2	1817.5	1803.9	13.05	.31