Análisis de datos longitudinales

Grado en Estadística

Juan R González
Departamento de Matemáticas, UAB
Insituto de Salud Global de Barcelona, ISGlobal

juanr.gonzalez@isglobal.org
https://github.com/isglobal-brge/TeachingMaterials
http://www.creal.cat/brge.thml

Análisis de datos longitudinales

Grado en Estadística

Tema 1

Regresión *joinpoint*

Juan R González

Departamento de Matemáticas, UAB

Insituto de Salud Global de Barcelona, ISGlobal

juanr.gonzalez@isglobal.org

https://github.com/isglobal-brge/TeachingMaterials

http://www.creal.cat/brge.thml

Tasas de Mortalidad

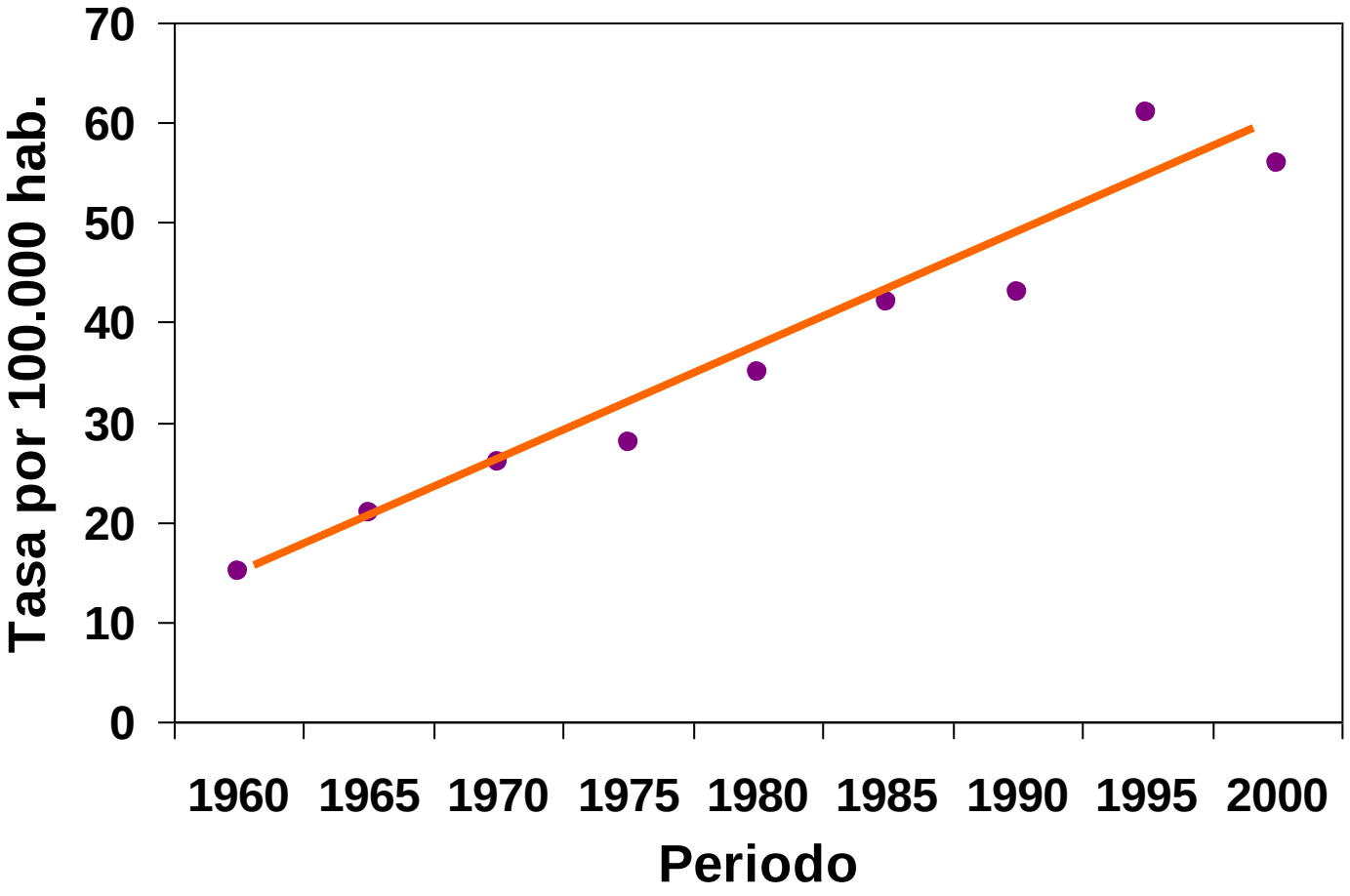
Nota: misma definición para incidencia

Defunciones en un año
(registro civil, MNP)

$$\text{Mortalidad} = \frac{\text{defunciones}}{\text{población a riesgo}} \times 10^{n} \text{ y año}$$

Población a 1 de julio del año

Tendencia temporal



Cambios en las tendencias

- **Efectos cohorte**
  - **Diferentes generaciones expuestas a diferentes riesgos**
  - **Cambios en las tasas sucesivos grupos de edad en sucesivos periodos**
  - **Largas exposiciones o hábitos (tabaco, dieta,...)**
- **Efectos periodo**
  - **Cambios en las tasas en un momento del tiempo**
  - **Influyen a todos los grupos de edad a la vez**
  - **Exposición puntual, cambio procesos dx (Chernobil, PSA, ...)**

Cambios en las tendencias

- **Interés evaluar efectos periodo**
  - **Cáncer:**
    - **Monitorizar las tasas**
    - **Evaluar cambios en el diagnóstico**
    - **Evaluar el efecto de un cribado**
    - **Evaluar efecto un tratamiento**
  - **SIDA**
    - **Monitorizar las tasas**
    - **Evaluar la influencia de un nuevo tratamiento**

Modelo lineal generalizado de Poisson

$\ln(E[t_{ij}]) = \mu + \beta \text{periodo}$

como

$\ln(E[t_{ij}]) = \ln(\theta_{ij} / N_{ij}) = \ln(\theta_{ij}) - \ln(N_{ij})$

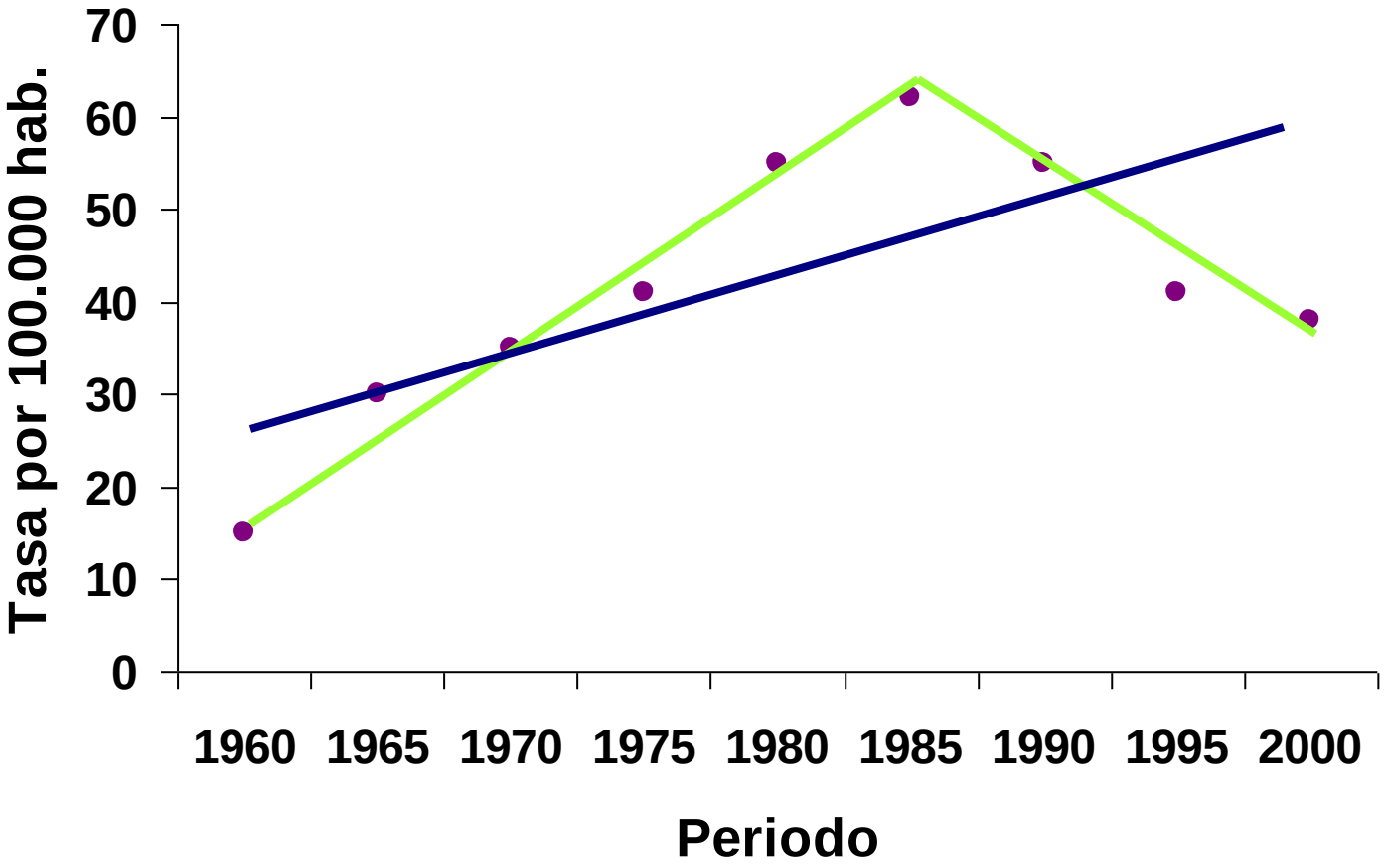
entonces

$\ln(\theta_{ij}) = \mu + \beta \text{periodo} + \ln(N_{ij})$

PCA: porcentaje de cambio anual $(1-\exp(\beta))\%$

Problema: Sobredispersión (binomial negativa)

Cambios en las tendencias. Motivación



Cambios en las tendencias

Pergamon

European Journal of Cancer, Vol. 35, No. 10, pp. 1477–1516, 1999

Printed in Great Britain
0959-8049/99/$ - see front matter

PII: S0959-8049(99)00154-9

Original Paper

Cancer Mortality in Europe, 1990–1994, and an Overview of Trends from 1955 to 1994

F. Levi,[1,2] F. Lucchini,[2] E. Negri,[3] P. Boyle[4] and C. La Vecchia[3,5]

[1]Registre Vaudois des Tumeurs, Institut Universitaire de Médecine Sociale et Préventive, Centre Hospitalier Universitaire Vaudois, Falaises 1, 1011 Lausanne; [2]Unité d'Épidémiologie du Cancer, Institut Universitaire de Médecine Sociale et Préventive, Bugnon, Lausanne, Switzerland; [3]Istituto di Ricerche Farmacologiche 'Mario Negri'; [4]European Institute of Oncology; and [5]Istituto di Statistica Medica e Biometria, Università degli Studi di Milano, Milan, Italy

Mortality data, abstracted from the WHO database, are presented in tabular form for 26 cancer sites or groups of sites, plus total cancer mortality, in 35 European countries during the period 1990–1994. Trends in mortality are also given in graphical form for 24 major countries over the period 1955–1994. In most western European countries total cancer mortality was—for the first time—moderately downwards in the early 1990s. Such favourable trends included some decline in lung cancer mortality for males, the persistent decline in stomach cancer for both sexes, and of cervical cancer for women, as well as some decline in breast and colorectal cancers, plus other neoplasms (testis, lymphoid neoplasms), whose treatment has further improved over the last few years. However, cancer mortality was still upwards in a few southern and eastern European countries, including Hungary and Poland, where total cancer mortality rates in middle-aged males are now the highest ever registered in Europe. The favourable trends in western Europe over the recent years are similar to those observed in the U.S.A.

Key words: epidemiology, mortality, neoplasms, time trends, Europe
Eur J Cancer, Vol. 35, No. 10, pp. 1477–1516, 1999

Cambios en las tendencias

Int. J. Cancer: **81,** 62–66 (1999)

Publication of the International Union Against Cancer
Publication de l'Union Internationale Contre le Cancer

RECENT DECLINES IN WORLDWIDE MORTALITY FROM CUTANEOUS MELANOMA IN YOUTH AND MIDDLE AGE

Carlo LA VECCHIA[1,2], Franca LUCCHINI[3], Eva NEGRI[1] and Fabio LEVI[3]*

[1]*Istituto di Ricerche Farmacologiche "Mario Negri", Milan, Italy*
[2]*Istituto di Statistica Medica e Biometria, Università di Milano, Milan, Italy*
[3]*Registre Vaudois des Tumeurs, Institut Universitaire de Médecine Sociale et Préventive, Lausanne, Switzerland*

TABLE I – TRENDS IN MORTALITY[1] FROM SKIN CANCER AT AGE 20 TO 44 IN SELECTED COUNTRIES, 1955–1984 AND 1985–1995

	Death rate/100,000 males					Death rate/100,000 females				
	1955–1959	1980–1984	1990–1995	Percent change/year[2] 1955–1984	Percent change/year[2] 1985–1995	1955–1959	1980–1984	1990–1995	Percent change/year[2] 1955–1984	Percent change/year[2] 1985–1995
Austria	0.83	1.46	0.99	2.2[3]	−4.1	0.52	1.01	0.89	2.5[3]	−0.6
Belgium (90–92)	0.33	0.62	1.02	1.8[3]	2.2	0.45	0.54	0.63	2.0[3]	−1.9
Denmark (90–93)	0.88	1.64	1.39	2.6[3]	−2.5	0.83	1.30	1.16	0.7	1.6
Finland	0.74	1.04	0.83	1.5	−5.8	0.91	0.71	0.58	0.6	0.7
France (90–94)	0.33	0.79	0.81	3.7[3]	−1.0	0.26	0.78	0.67	4.1[3]	0.7
Germany[4]	0.78	0.95	0.84	0.4	−3.7	0.60	0.81	0.68	1.0[3]	−2.1[3]
Hungary	0.44	1.50	1.32	4.7[3]	0.8	0.46	1.05	0.96	3.7[3]	0.7
Ireland (90–94)	0.70	0.97	0.83	0.9	−5.1	0.66	0.99	0.93	1.3	−2.3
Italy (90–94)	0.43	0.91	0.84	2.8[3]	−1.4	0.32	0.74	0.71	2.8[3]	0.3
Netherlands	0.74	1.40	1.41	3.0[3]	−1.6	0.56	1.03	1.22	2.4[3]	1.8
Norway (90–94)	0.96	2.12	1.87	3.1[3]	−1.6	0.88	1.42	1.38	2.1[3]	−2.2
Poland	0.26	0.87	1.04	3.9[3]	−0.5	0.19	0.18	1.06	5.1[3]	2.2
Portugal	0.24	0.39	0.54	2.6[3]	−2.9	0.25	0.21	0.55	−0.5	12.0[3]
Spain (90–94)	0.17	0.54	0.83	4.8[3]	2.7	0.14	0.37	0.60	5.0[3]	5.8[3]
Sweden	0.99	1.24	1.35	0.9	0.1	0.52	1.04	0.81	2.5[3]	−6.0[3]
Switzerland (90–94)	1.14	1.54	1.27	1.1	0.4	0.73	1.05	0.97	1.2	0.8
UK, England and Wales	0.62	1.10	1.19	2.1[3]	0.4	0.74	1.01	0.93	1.4[3]	−0.4
UK, Scotland	0.58	0.84	0.97	1.6[3]	1.3	0.63	0.98	0.91	2.1[3]	0.4
United States	1.10	1.77	1.40	1.6[3]	−5.0[3]	0.90	0.97	0.78	0.3[3]	−2.2[3]
Canada	0.70	1.20	0.99	2.1[3]	−3.2[3]	0.62	0.97	0.67	1.4[3]	−4.4[3]
Australia	2.83	2.81	3.01	0.4	0.0	1.89	1.62	1.43	−0.6[3]	−3.3
New Zealand (90–94)	1.74	3.02	2.73	2.8[3]	−5.4	1.12	2.76	1.81	2.7[3]	−1.1

[1]Age-standardized rates on the world population. [2]From a log-linear model based on single calendar year rates. [3]$p < 0.05$. [4]Before 1973, rates refer to Federal Republic of Germany.

Percent change/year

1955-84 1985-95

Regresión Joinpoint

- $(x_1,y_1), \ldots, (x_n,y_n)$ con $x_1 \leq \ldots \leq x_n$
- $E[y|x]=\beta_0+\beta_1 x+\delta_1(x-\tau_1)^+ +\ldots+\delta_k(x-\tau_k)^+$

  Donde τ_k es un joinpoint desconocido

  $a^+ = a$ para $a>0$
- **Solución** (regresión no lineal con restricciones):

  regresión "piecewise"

  regresión segmentada

  regresión "broken line"

  regresión multi-fase

Inferencia sobre los Joinpoints

PERMUTATION TESTS FOR JOINPOINT REGRESSION WITH APPLICATIONS TO CANCER RATES

Hyune-Ju Kim[1], Michael P. Fay[2], Eric J. Feuer[2] and Douglas N. Midthune[3]

Test de permutaciones

A	B	B	A	B	A	B	B	A	A
2,3	2,8	3,6	2,5	5,6	4,2	3,6	5,8	1,6	2,9

media A=2,70
media B=4,28
Δ=-1.58

H_0: A=B H_1: A≠B

Si H_0 cierta cualquier combinación de A y B debería ser igual

B	A	B	A	B	A	B	B	A	A
2,3	2,8	3,6	2,5	5,6	4,2	3,6	5,8	1,6	2,9

media A=2,80
media B=4,18
Δ=-1.38

.....

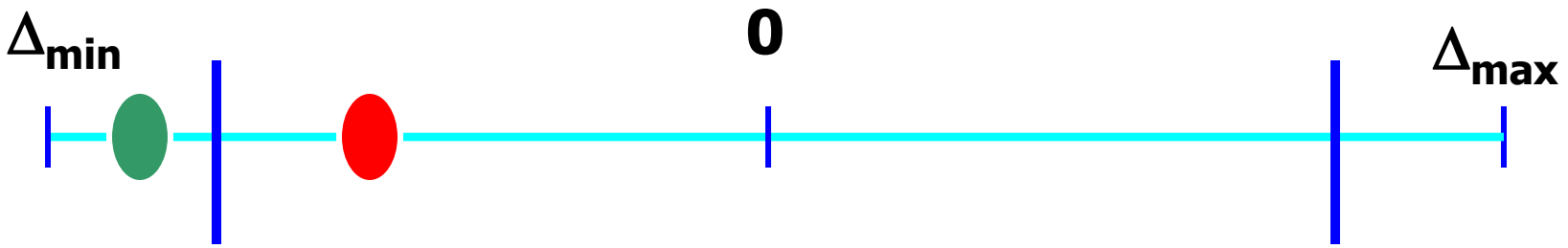
A	B	B	A	B	A	B	A	B	A
2,3	2,8	3,6	2,5	5,6	4,2	3,6	5,8	1,6	2,9

media A=3,54
media B=3,44
Δ=0.10

Test de permutaciones

H_0: A=B H_1: A≠B

Si H_0 cierta cualquier combinación de A y B debería ser igual



Para α=0,025 y 1000 muestras $\rightarrow$ 25

Otra forma es calcular p-valor directamente (#Δ>obs)/n

Test de permutaciones

Otra forma es calcular p-valor directamente (#Δ>obs)/n

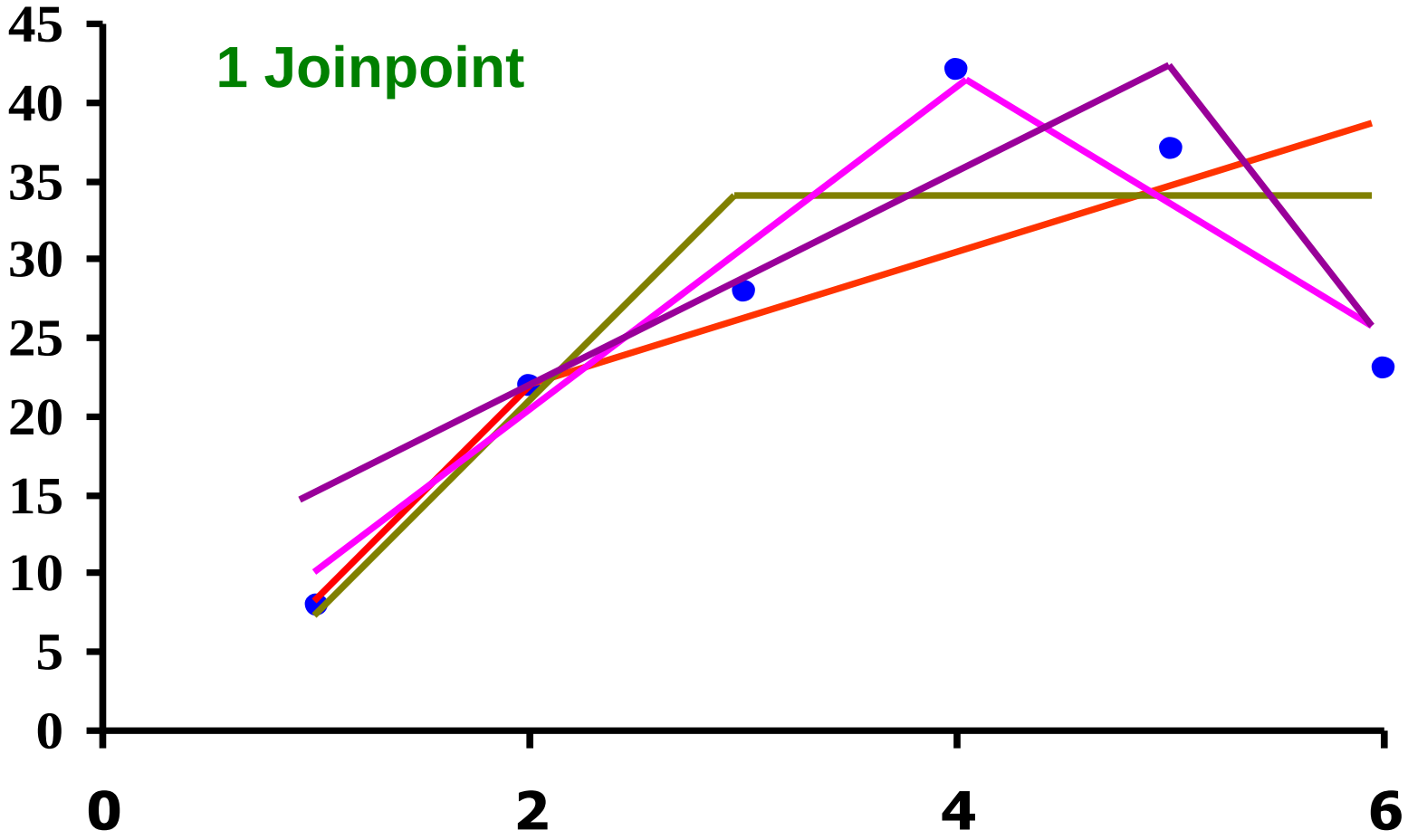
-1.96, -1.74, -1.54, -2.30, -0.3, 0.1, 0.3, -1.94, … , -1. 92

contamos (#Δ>-1.58) y dividimos por n con n el total de permutaciones

PROBLEMA: n/2 A's y n/2 B's n es muy grande n!

SOLUCION: Montecarlo tomar p.e. 1000 de esas permutaciones al azar y calcular el p-valor igual

Test de permutaciones (Joinpoint)



Regresión Joinpoint. Algoritmo

Testamos k_0 joinpoints vs. k_1 joinpoints

Rech. H_0 con α_1

Testamos k_1+1 joinpoints vs. k_1 joinpoints

Testamos k_0 joinpoints vs. k_1-1 joinpoints

- **K+1 joinpoints si rechazamos el último test**
- **K si no**
- **k_1 - k_0 tests Ajustamos α por Bonferroni**
- **ICA entre τ_j y τ_{j+1} es: $100*\exp(\beta_1+\delta_1+...+\delta_j)-1$**

Regresión Joinpoint.

1. Estimar el modelo bajo H_0 (modelo nulo)
2. Permutar los residuales del modelo nulo. Obtenermos N_p data sets permutados
3. Para cada uno de estos set de datos, estimamos el modelo alternativo (bajo H_1) y calculamos un estadístico de bondad de ajuste
4. El p-valor se determina de la distribución permutacional de dicho estadístico.

Regresión Joinpoint. Ejemplo

H_0: k=0 vs H_1: k=1

1º Estimar el modelo bajo la Hipótesis nula

$$E[y|x]=\beta_0+\beta_1 x+\delta_1(x-\tau_1)^+$$

Se lleva a cabo mediante regresión segmentada (paquete 'segemented' en R)

Regresión Joinpoint. Ejemplo

2º Permutar los residuales y añadirlos al modelo nulo

$$y'_{(a)} = \hat{\mu}^{(k_0)'} + [\hat{\epsilon}^{(k_0)}_{\pi_{a1}}, \ldots, \hat{\epsilon}^{(k_0)}_{\pi_{an}}]$$

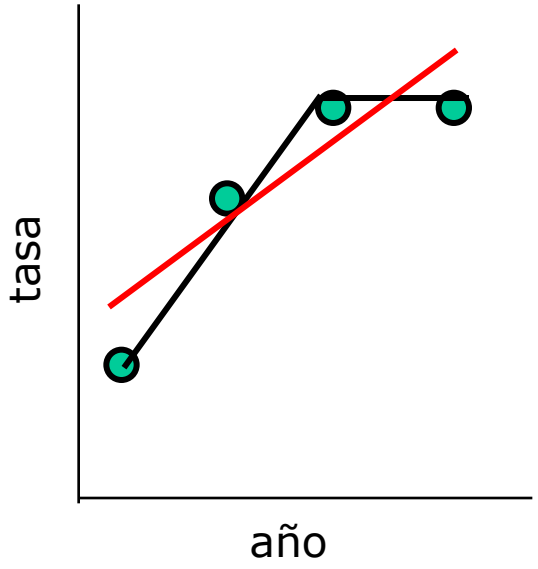
Ej.		
Año	**tasa**	**error**
1985	**12**	**2**
1986	**16**	**3**
1987	**19**	**1**
1988	**20**	**4**

Ej.	
Año	tasa’
1985	13
1986	17
1987	22
1988	22

Permutación: 1 3 4 2

Regresión Joinpoint. Ejemplo

3º Estimar el modelo bajo H_1 para esos datos permutados



Bajo H_0

Bajo H_1

$T(y_{(a)})$ = error bajo H0 / error bajo H1

Regresión Joinpoint. Ejemplo

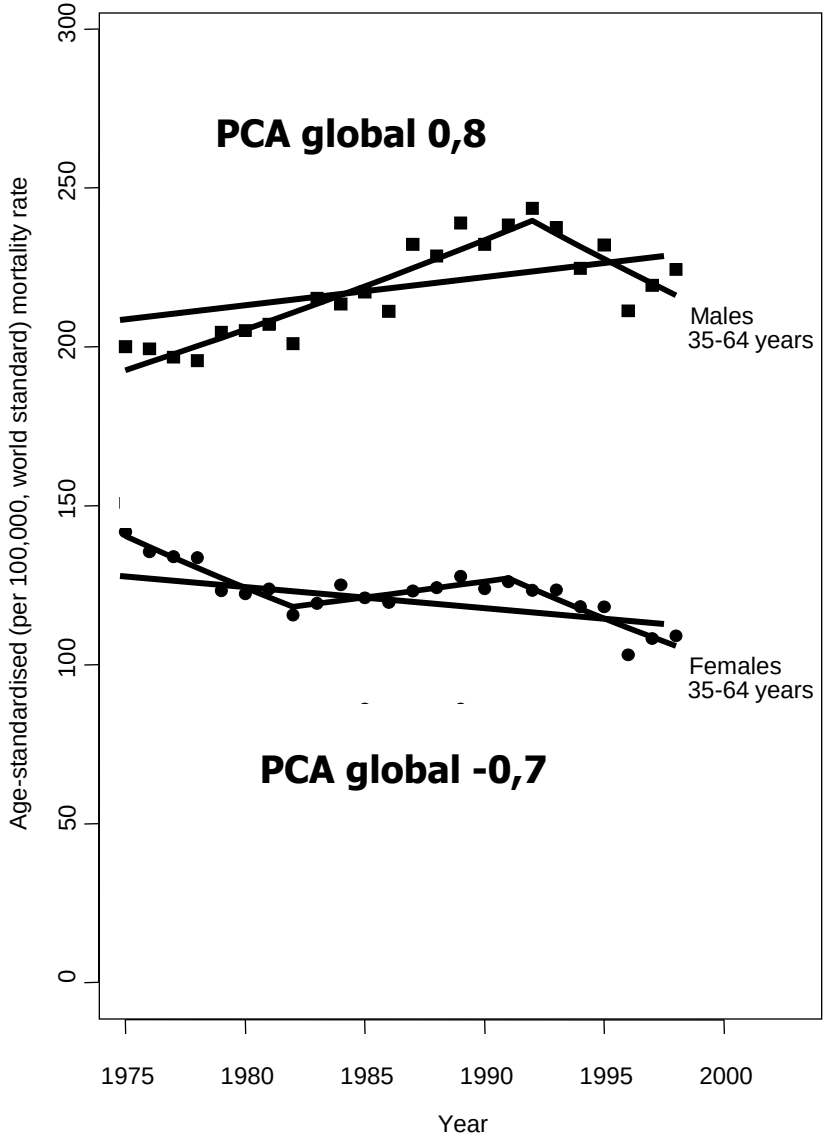
4º Realizamos el cálculo mediante Monte-Carlo

N_p - 1 permutaciones de $T(y_{(a)})$ a=1, … , N_p - 1
Tomamos el estadístico $T(y) \equiv T(y_{(0)})$ de los datos originales. El p-valor viene dado por:

$$p = \frac{\#\text{veces que } [T(y_{(a)}) \geq T(y)] \text{ para } a \in \{0,1,\ldots,N_p - 1\}}{N_p}$$

Mortalidad Cataluña

E Fernández, JR González, JM Borràs et al. Eur J Cancer, 2001



Hombres 35-64

Año	1975-92	1992-97
PCA	1,3	-2,2

Mujeres 35-64

Año	1975-82	1982-92	1992-97
PCA	-2,4	0,7	-3,5

Mortalidad Cataluña

Cáncer de Pulmón

	Global	Regresión Joinpoint					
	1975-1998	Tendencia 1		Tendencia 2		Tendencia 3	
Hombres	1,9	**1975-1982**	2,3	**1982-1989**	4,1	**1989-1997**	-0,9
Mujeres	-0,2	**1975-1988**	-1,8	**1988-1997**	2,2		

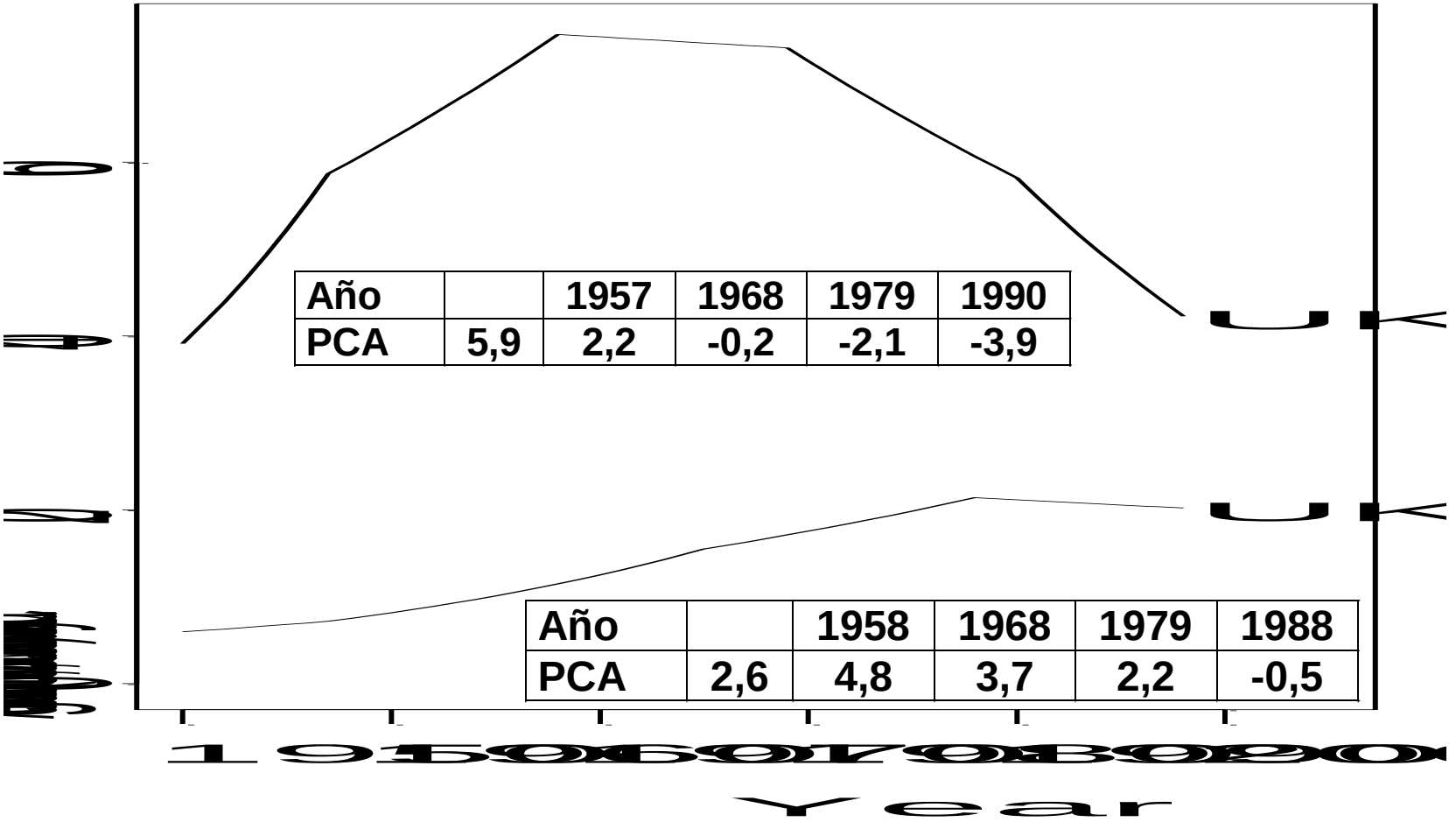
E Fernández, JR González, JM Borràs et al. Eur J Cancer, 2001

Cáncer Cataluña 1975-98. Hombres

	Age-standardised mortality rates[a]		Annual percent of change 1975-98	Joinpoint analysis (1975-1998)							
				Trend 1		Trend 2		Trend 3		Trend 4	
	1975-76	1997-98		Years	APC	Years	APC	Years	APC	Years	APC
All sites	150.92	171.81	0.7[b]	1975-1981	0.5	1981-1992	1.6[b]	1992-1996	-2.5[b]	1996-1998	1.2
Smoke related	61.86	83.98	1.5[b]	1975-1990	2.8[b]	1990-1998	-1.1[b]				
Oral Cavity and pharynx (140-149)	3.64	6.93	3.4[b]	1975-1998	3.4[b]						
Oesophagus (150)	5.35	5.05	0.3	1975-1998	0.3						
Stomach (151)	20.50	10.42	-2.9[b]	1975-1981	-4.3[b]	1981-1998	-2.6[b]				
Colon (153)	8.08	13.45	3.3[b]	1975-1978	0.6	1978-1981	-8.1	1981-1988	8.8[b]	1988-1998	2.0[b]
Recto (154)	5.24	4.75	-0.6[b]	1975-1988	0.1	1988-1998	-3.2[b]				
Colorectal (153, 154, 159.0)	13.31	18.47	1.9[b]	1975-1982	0.2	1982-1991	3.5[b]	1991-1998	-0.1		
Liver (155)	9.78	8.90	0.1	1975-1983	0.3	1983-1987	-6.0	1987-1990	9.9	1990-1998	-1.5
Gallbladder (156)	1.17	1.37	1.4[b]	1975-1977	-16.7	1977-1987	7.9[b]	1987-1991	-7.4	1991-1998	1.9
Pancreas (157)	3.87	6.50	2.3[b]	1975-1987	3.9[b]	1987-1998	0.8				
Larynx (161)	7.87	5.25	-1.7[b]	1975-1991	-1.3[b]	1991-1998	-3.4[b]				
Lung (162)	32.77	48.77	1.8[b]	1975-1982	2.3[b]	1982-1989	4.0[b]	1989-1998	-0.8[b]		
Melanoma (172)	0.42	1.14	5.0[b]	1975-1986	10.3[b]	1986-1998	2.1				
Prostate (185)	12.90	14.24	0.4[b]	1975-1998	0.4[b]						
Urinary bladder (188)	6.68	8.57	1.1[b]	1975-1992	2.3[b]	1992-1998	-3.3				
Kidney (189)	2.00	3.08	2.4[b]	1975-1990	3.9[b]	1992-1998	-0.4				
Brain (191.192)	5.57	4.27	-1.6[b]	1975-1998	-1.6[b]						
Thyroid (193)	0.15	0.27	1.8	1975-1998	1.8						
Hodgkin disease (201)	1.46	0.87	-2.9[b]	1975-1998	-2.9[b]						
Non-Hodgkin lymphoma (200.202)	1.76	3.76	3.4[b]	1975-1991	5.6[b]	1991-1998	-1.9				
Myeloma (203)	1.02	1.93	2.8[b]	1975-1990	4.9[b]	1991-1998	-0.9				
Leukemias (204-208)	5.35	4.91	-0.2	1975-1991	0.9[b]	1991-1998	-3.3				

Cáncer Pulmón Europa. 1955-1998

JM Borràs, E Fernández, JR González, Eva Negri,
Franca Lucchini, Carlo La Vecchia, Fabio Levi
Annals of Oncology, 2003

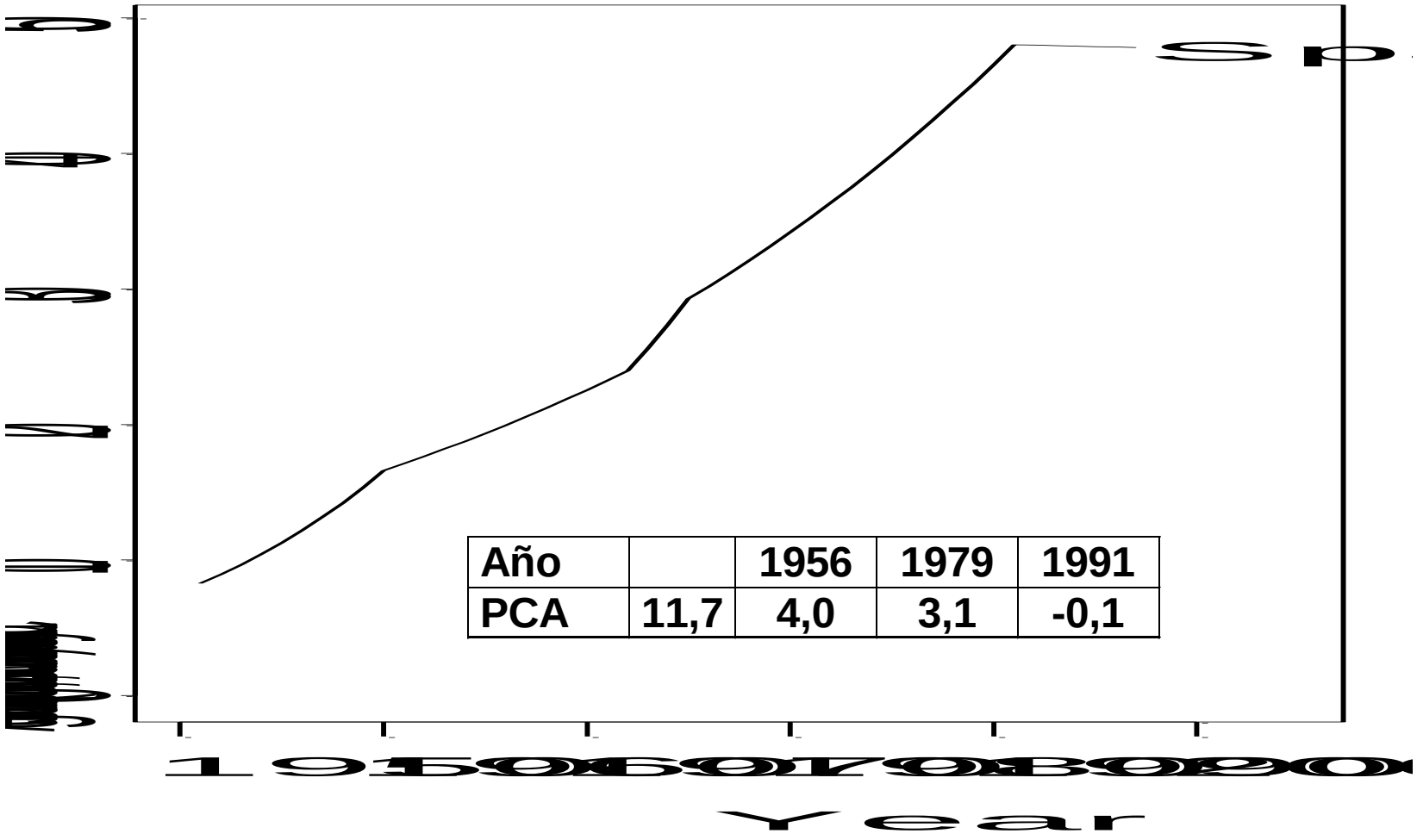


Año		1957	1968	1979	1990
PCA	5,9	2,2	-0,2	-2,1	-3,9

Año		1958	1968	1979	1988
PCA	2,6	4,8	3,7	2,2	-0,5

Cáncer Pulmón Europa. 1955-1998

JM Borràs, E Fernández, JR González, Eva Negri,
Franca Lucchini, Carlo La Vecchia, Fabio Levi
Annals of Oncology, 2003



Año		1956	1979	1991
PCA	11,7	4,0	3,1	-0,1

Cáncer Pulmón Europa. 1955-1998

JM Borràs, E Fernández, JR González, Eva Negri,
Franca Lucchini, Carlo La Vecchia, Fabio Levi
Annals of Oncology, 2003

Año		1957	1973	1989
PCA	7,8	1,2	-1,1	2,1