

L'ús de microdades en l'anàlisi economètrica aplicada

Jaume García
Instituto Nacional de Estadística

*Facultat de Ciències, Universitat Autònoma de Barcelona
Bellaterra, 27 de novembre de 2009*

IN

e

❖ **Motivación**

❖ **Implicaciones econométricas**

❖ **Un ejemplo del uso de microdatos:**

✓ *Diferencias salariales por razón de género*

IN

—

e

Motivación

❖ Ejemplos de microdatos (datos individuales):

- *EPA, EPF, EET, ERRHHCyT*
- *EIE, EIP, EAI+D, EAS*

❖ Microdatos *versus* series temporales

❖ Característica básica de los microdatos

❖ Accesibilidad a los microdatos

- ✓ www.ine.es/prodyser/microdatos.htm

Instituto Nacional de Estadística. (National Statistics Institute) - Windows Internet Explorer

http://www.ine.es/

Archivo Edición Ver Favoritos Herramientas Ayuda

Instituto Nacional de Estadística. (National Statistics I... Página Herramientas

INE Instituto Nacional de Estadística

EL INE INEbase IPC Productos y Servicios Área de Prensa Ayuda English

Guía de navegación
Preguntas frecuentes
Contactar con el INE
Calendario de estadísticas
 Toda la información estadística
 Entorno físico / medio ambiente
 Demografía y población
 Sociedad
 Economía
 Ciencia y tecnología
 Agricultura
 Industria, energía, construcción
 Servicios
 Clasificaciones
 Internacional
 Síntesis estadística
 Lista completa de operaciones
 TEMPUS Banco de series
 IPC Índice Precios Consumo
 Oficina del Censo Electoral
 Colaboradores en encuestas económicas

Dato	Fecha	Valor	%
Puntos de atención al público			
IPC. Índice de precios de consumo	05/09	106,8	-0,9
RSS. Servicio de noticias			
IPRI. Índice de Producción Industrial	05/09	111,8	-4,4
Publicación	05/09	81,9	-22,8*
PIB. Producto Interior Bruto	1ºT/09	267.317	-3,0
Biblioteca			
ETCL. Encuesta de Turismo en Alojamiento	1ºT/09	2.434,32	3,9
Información europea			
Ocupación. Encuesta de Ocupación en Alojamiento de Turismo Rural	05/09	22.836,8	-8,87
Carta de servicios	05/09	47,40	-10,48
Precios productos difusión			
EPA. Encuesta de Población y Vivienda	1ºT/09	60,15	0,80
Oficina virtual			
EPA. Tasa de paro (%)	1ºT/09	17,36	7,73
Hipotecas constituídas: Importe medio (€)	04/09	142.817	-12,0
Padrón municipal: Población total (miles) **	1/1/09	46.662,0	1,1
Padrón municipal: Población extranjera (miles) **	1/1/09	5.598,7	6,3

* Media de lo que va de año
** Datos provisionales

Última hora **RSS** Servicio de noticias

03 Jul 09. Índice de Producción Industrial (IPI). Mayo 2009: -22,3%
El Índice de Producción Industrial corregido de calendario baja un 20,5% en mayo respecto al mismo mes de 2008. [Información detallada.](#)

02 Jul 09. Encuesta de Ocupación en Alojamiento de Turismo Rural 2008
Se publican los datos de la Encuesta de ocupación en alojamientos de turismo rural del año 2008.

02 Jul 09. Profesionales sanitarios colegiados 2008
Se publican los datos de Profesionales sanitarios colegiados. Año 2008.

IPC EPA PIB POB

Indice de precios de consumo. IPC

Var. mensual
Var. anual

Indice general nacional (%)

Cuántos habitantes tiene ...
Escriba una población

Apellidos y nombres más frecuentes

Historia Anuarios y censos siglos XIX y XX

Novedades Catálogo

http://www.ine.es/prodyser/microdatos.htm

Internet 100%

Inicio Ine 1... Zi... U... ac... pr... In... 13:10

Productos y servicios / Información / Ficheros de microdatos - Windows Internet Explorer

http://www.ine.es/prodyser/microdatos.htm

Archivo Edición Ver Favoritos Herramientas Ayuda

Productos y servicios / Información / Ficheros de micr...

Sociedad	Periodicidad	Serie disponible
Educación		
Encuesta de Transición Educativo-Formativa e Inserción Laboral	Única	2005
Encuesta sobre la Participación de la Población Adulta en las Actividades de Aprendizaje (EADA)	Quinquenal	2007
Salud		
Encuesta de salud y hábitos sexuales	Única	2003
Encuesta sobre Discapacidades, Deficiencias y Estado de Salud	Únicas	1986 y 1999
Encuesta de Morbilidad Hospitalaria	Anual	Desde 1978
Condiciones de vida		
Encuesta de Presupuestos Familiares (Base 2006)	Anual	Desde 2006
Encuesta básica de Presupuestos Familiares	Únicas	1980-81 y 1990-91
Encuesta continua de Presupuestos Familiares	Trimestral	1tr/1985 - 4tr/ 2005
Encuesta continua de Presupuestos Familiares. Fichero longitudinal	Anual	1998 - 2004
Panel de hogares de la UE. Datos de España	Anual	Desde 1994
Encuesta sobre Equipamiento y Uso de Tecnologías de la Información y Comunicación en los hogares	Anual	Desde 2002
Encuesta de Empleo del Tiempo	Única	2002-2003
Mercado laboral		
Encuesta de Población Activa. EPA. Resultados trimestrales de la encuesta	Trimestral	Desde 1tr/1999
Encuesta de Estructura Salarial	Única	2006
Análisis sociales		
Encuesta de hogares y medio ambiente 2008	Única	2008
Ciencia y Tecnología		
Investigación y Desarrollo Tecnológico		
Encuesta sobre Innovación Tecnológica en las Empresas		
Agricultura		
Periodicidad		
Serie disponible		
Censo agrario	Únicos	1982, 1989 y 1999
Encuesta sobre la Estructura de las Explotaciones Agrícolas	Bienal	Desde 1987

Esperando a http://www.ine.es/prodyser/microdatos.htm...

Internet 100%

Inicio | Internet Explorer | A... | INE | 1... | Zi... | U... | ac... | pr... | Pr... | 13:09

INE

Implicaciones econométricas

Punto de partida: Modelo de regresión

$$Y_i = \beta_1 + \beta_2 X_{2,i} + \dots + \beta_K X_{K,i} + u_i$$

Variables cualitativas explicativas

Ejemplos: género, sector CNAE, región, educación, grupos de edad, etc.

Variables ficticias (*dummies*)

$$D_i = \begin{cases} 1 & \text{si se tiene un determinado atributo} \\ 0 & \text{en caso contrario} \end{cases}$$

IN

—

e

Variables dependientes cualitativas

Variable dependiente: decisión o situación

Ejemplos: decisión de consumir, decisión de exportar, situación en el empleo, decisión acerca del tipo de transporte

Valor ajustado: probabilidad de escoger una determinada alternativa

Modelos de elección discreta con dos alternativas:
Logit, Probit

Modelos de elección discreta con tres alternativas:
Logit multinomial, modelos anidados
(D. McFadden, Premio Nobel 2000)

IN

VARIABLES DEPENDIENTES LIMITADAS

Variable dependiente: no negativa y con una proporción importante de ceros.

Ejemplos: gasto en determinados bienes, número de horas trabajadas, asistencia a un espectáculo.

Valor ajustado: valor esperado de la variable dependiente, probabilidad de que la variable dependiente sea positiva.

Modelo más sencillo: modelo Tobit (1958)

Interpretación de los efectos de las variables explicativas: dos componentes

Importancia de la modelización de los ceros

e

IN

—

e

Importancia de la selección de la muestra

Muestra no aleatoria: seleccionada según algún criterio

Ejemplos: estimación de ecuaciones de salarios con la muestra de trabajadores, estimación de ecuaciones de ingresos

Valor ajustado de la variable dependiente condicionado o no a pertenecer a la muestra

Modelo de regresión con término de corrección
(J.J. Heckman, Premio Nobel 2000)

Importancia del término de corrección

IN

—

e

Modelos de duración

Variable dependiente: tiempo que un individuo lleva en un determinado estado

Ejemplos: duración de los episodios de desempleo, duración de los procesos de huelga, demografía empresarial

Valor ajustado: Probabilidad de abandonar un estado condicionado a llevar un determinado tiempo en el mismo

Modelo exponencial, Weibull y otros

Importancia de la heterogeneidad no observada

IN

e

Modelos para datos de panel

Ejemplo: Estimación de la tasa de rentabilidad de la educación (β)

$$\ln w_i = \alpha + \beta S_i + G_i + E_i + u_i$$

Importancia del efecto de los factores inobservables



Estimaciones sesgadas de β

Muestra de parejas de hermanos gemelos

$$G_{j,1} = G_{j,2}$$

$$E_{j,1} = E_{j,2}$$

IN E

$$\ln w_{j,1} = \alpha + \beta S_{j,1} + G_{j,1} + E_{j,1} + u_{j,1}$$

$$\ln w_{j,2} = \alpha + \beta S_{j,2} + G_{j,2} + E_{j,2} + u_{j,2}$$

$$\ln w_{j,1} - \ln w_{j,2} = \beta (S_{j,1} - S_{j,2}) + (u_{j,1} - u_{j,2})$$

Fuentes de variabilidad

$$(X_{i,t} - \bar{X})^2 \quad (X_{i,t} - \bar{X}_i)^2 \quad (\bar{X}_i - \bar{X})^2$$

Efectos fijos *versus* efectos aleatorios

Estimadores tradicionales:

Estimador intragrupos (*within-group*)

Estimador entregrupos (*between-group*)

Estimador de componentes de error

IN

Ejemplo:

**Análisis de las diferencias salariales por
razón de género**

e

IN
e

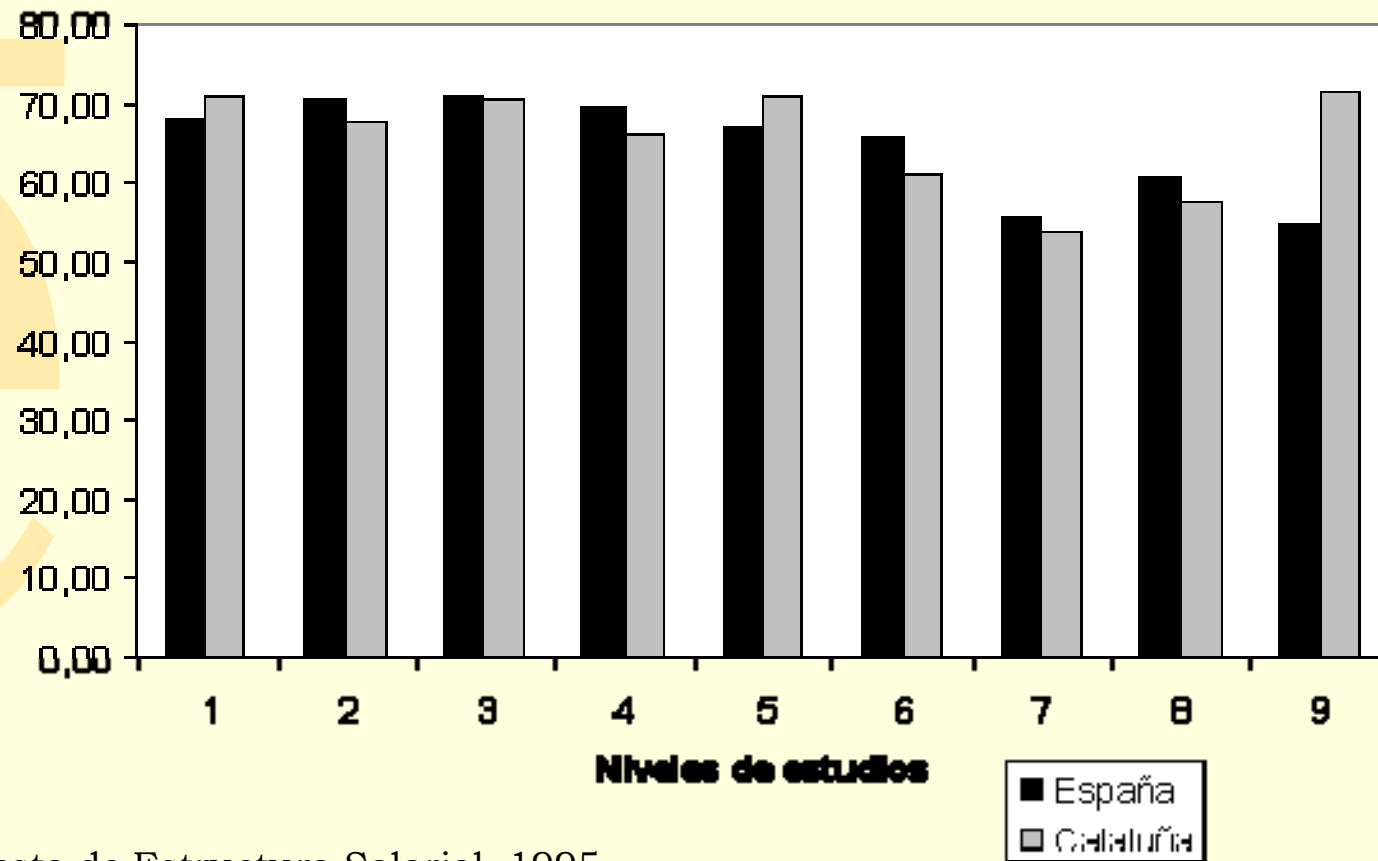
Salario medio de las mujeres / salario medio de los hombres (%)

Alemania	80,6
Austria	78,9
Bélgica	92,7
Dinamarca	89,6
España	85,7
Francia	89,2
Gran Bretaña	75,7
Grecia	86,8
Holanda	78,9
Irlanda	75,8
Italia	91,4
Portugal	94,1
Unión Europea	83,8

Fuente: Comisión Europea (2002), PHOGUE-1998.

IN

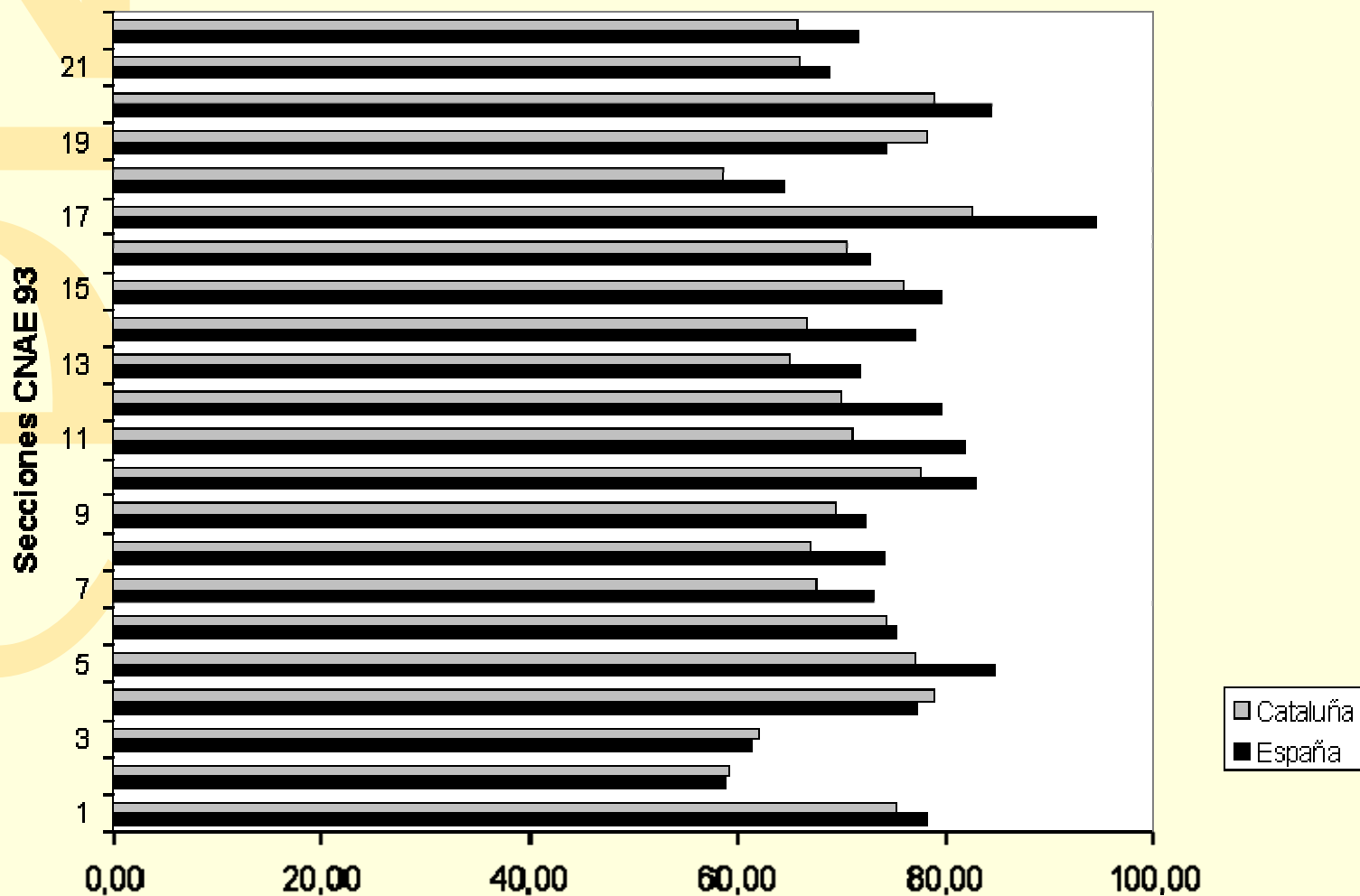
Salario medio de las mujeres / salario medio de los hombres (%) por niveles de estudios



Fuente: Encuesta de Estructura Salarial, 1995
Instituto Nacional de Estadística

INI

Salario medio de las mujeres / salario medio de los hombres (%) por ramas de actividad (CNAE)



Fuente: Encuesta de Estructura Salarial, 1995
Instituto Nacional de Estadística

Divisiones CNAE 1993**Media 1996-2001**

	% Hombres	Salario/hora
18 Industria de la confección y de la piel	77.70	5.94
52 Comercio minorista	59.58	5.99
70 Actividades inmobiliarias	59.17	8.05
74 Otras actividades empresariales	56.19	7.52
67 Actividades auxiliares en la intermediación financiera	48.91	10.36
55 Hostelería	46.94	5.74
19 Industria del cuero y otros	42.59	5.53
65 Intermediación financiera	29.19	17.50
11+12 Petróleo, gas natural	24.40	19.22
23 Tratamiento del petróleo y combustibles nucleares	12.10	18.53
14 Extracción de minerales no metálicos ni energéticos	7.57	8.53
27 Metalurgia	7.24	11.72
45 Construcción	4.53	6.92
13 Extracción de minerales metálicos	2.98	11.36
10 Extracción de carbón	1.47	13.89

Fuente: Encuesta de Salarios en la Industria y los Servicios
Instituto Nacional de Estadística

IN

e

Salario de las mujeres / salario de los hombres (%) en diferentes puntos de la distribución

Percentil	Cataluña	España
10	67.6	63.2
25	69.6	75.1
50 (mediana)	66.0	69.5
Mitjana	65.3	69.0
75	68.0	70.3
90	64.6	70.4

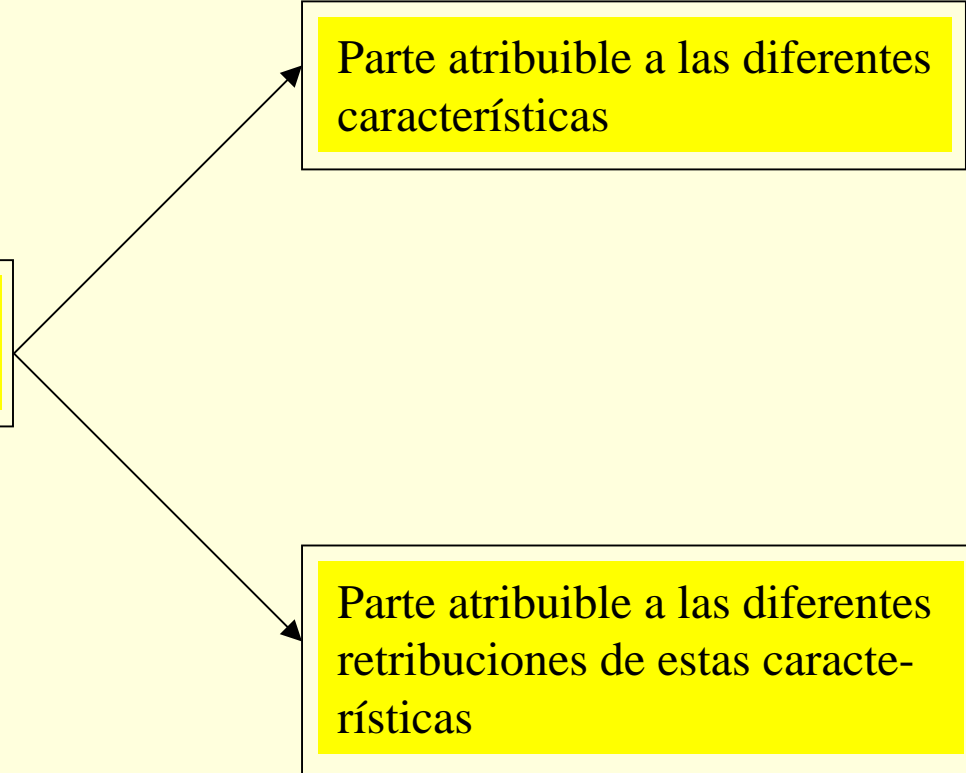
Fuente: Encuesta de Estructura Salarial, 1995
Instituto Nacional de Estadística

Descomposición de las diferencias salariales entre hombres y mujeres

Diferencia entre el salario medio de los hombres y de las mujeres

Parte atribuible a las diferentes características

Parte atribuible a las diferentes retribuciones de estas características



IN

Análisis econométrico de la discriminación salarial por razón de género

$$\bar{w}_h - \bar{w}_m$$

$$w_i = f(\text{edad}, \text{educación}, \text{experiencia}, \text{sexo})$$

$$w_i = \alpha + \beta X_i + [GÉNERO] + \varepsilon_i$$

$$w_i = \alpha + \beta X_i + \delta H_i + \varepsilon_i$$

$$H_i = 1 \quad \text{hombre}$$

$$= 0 \quad \text{mujer}$$

e

IN

e

Descomposición de Oaxaca

$$w_h = X_h' \beta_h + u_h \Rightarrow \bar{w}_h = \bar{X}_h' \hat{\beta}_h$$

$$w_m = X_m' \beta_m + u_m \Rightarrow \bar{w}_m = \bar{X}_m' \hat{\beta}_m$$

$$\bar{w}_h - \bar{w}_m = (\bar{X}_h' - \bar{X}_m') \hat{\beta}_h + \bar{X}_m' (\hat{\beta}_h - \hat{\beta}_m)$$

$$\bar{w}_h - \bar{w}_m = (\bar{X}_h' - \bar{X}_m') \hat{\beta}_m + \bar{X}_h' (\hat{\beta}_h - \hat{\beta}_m)$$

Parte atribuible a diferentes características

Parte atribuible a diferentes retribuciones

IN

Análisis de la discriminación en Cataluña y España (EES, 1995)

Salario medio de las mujeres / salario medio de los hombres

Cataluña	65.2%
España	69.0%

Diferencia salarial atribuible a la discriminación

Cataluña	72.1%
España	76.8%

Características: Edad, antigüedad, educación, tipo de contrato, tipo de jornada, tamaño de la empresa, rama de actividad

IN

e

Importancia de la discriminación a lo largo de la distribución de los salarios

Centil	%
10è	71.9
25è	64.6
50è (mediana)	69.1
75è	71.0
90è	90.5

Fuente: García, Hernández y López (2001)

IN

—

e

Una frase para cerrar:

*“La Estadística sirve para aliviar
el sufrimiento humano”*

José Aranda

L'ús de microdades en l'anàlisi economètrica aplicada

Jaume García
Instituto Nacional de Estadística

*Facultat de Ciències, Universitat Autònoma de Barcelona
Bellaterra, 27 de novembre de 2009*