

Encuesta de población activa

Matrices de doble codificación CNAE 93 – CNAE 2009

Antecedentes

La Encuesta de Población Activa (EPA) recoge trimestralmente, entre otras muchas variables, información sobre la actividad del establecimiento donde trabajan las personas ocupadas (tanto en el empleo principal como, en su caso, en el empleo secundario) y, para aquellas personas que no están ocupadas pero tienen experiencia profesional, la actividad del establecimiento donde trabajaron por última vez si hace un año o menos que dejó su empleo¹.

Estas variables se codifican siguiendo las clasificaciones estadísticas oficiales correspondientes.

A lo largo de 2008 la EPA inició la codificación de sus variables de actividad con la nueva Clasificación Nacional de Actividades 2009 (CNAE-2009). Una breve referencia sobre el cambio a esta clasificación puede consultarse en:

<http://www.ine.es/daco/daco42/clasificaciones/cnae09/cnae2009.pdf>.

Para construir el enlace de las series en las que intervienen las variables de actividad económica, a lo largo de los cuatro trimestres de 2008 las variables de actividad se codificaron utilizando tanto la clasificación nueva CNAE-2009 como la anterior CNAE1993. Ambas codificaciones han sido incluidas en los ficheros anonimizados de microdatos que se ponen a disposición de los usuarios interesados, para que puedan analizar la correspondencia entre ambas clasificaciones a lo largo de 2008.

Con el fin de facilitar la transición entre ambas clasificaciones a otros periodos distintos de 2008, se ha incluido, en la página dedicada a la EPA en la WEB del INE, una matriz de conversión, calculada a partir de la **doble codificación de la actividad del establecimiento del empleo principal para las personas ocupadas en alguno de los cuatro trimestres de 2008**. Se considera que este colectivo es el más idóneo para proporcionar una estimación robusta de la correspondencia empírica entre ambas clasificaciones.

En todo caso, hay que tener en cuenta que **la correspondencia entre ambas clasificaciones**, para aquellas agrupaciones en las que no existe identidad conceptual en su contenido (es decir, cuando la correspondencia no es biunívoca), **tiende a ser menos robusta cuanto más se separa del periodo en el que la doble codificación ha sido realizada** (año 2008 en este caso), ya que la propia estructura de actividad de la economía española es dinámica.

A continuación se describe la matriz y cómo se han calculado las series retrospectivas del periodo 2000-2007 con CNAE-2009.

¹ También se obtiene información sobre la actividad del establecimiento para las personas con experiencia profesional que han dejado su empleo hace ocho años o menos pero sólo en el caso de aquellas personas que están en la última entrevista. Esta información se explota en el fichero anual de submuestra de la EPA y no ofrece resultados trimestrales. También en el fichero anual se ofrece la información de la actividad del establecimiento hace un año.

Descripción de la matriz

La matriz de conversión se ha determinado para una **desagregación de dos dígitos en ambas clasificaciones** considerando esta elección el punto óptimo entre el nivel de detalle necesario y la fiabilidad de las estimaciones obtenidas de la doble codificación en media anual 2008. **Las estimaciones para agrupaciones de actividades se han realizado por agregación de las correspondientes actividades a dos dígitos que las constituyen.**

La matriz se ha calculado como **media anual de los cuatro trimestres de 2008, tabulando la actividad principal del total de ocupados al nivel de divisiones (códigos de dos dígitos) de ambas clasificaciones.** Hay 60 divisiones en CNAE1993 y 88 en CNAE2009. La información obtenida en 2008 se puede utilizar tanto para estimaciones retrospectivas con CNAE 2009 como estimaciones futuras con CNAE 1993, simplemente aplicando la distribución porcentual adecuada. Así pues, se proporcionan, para el mencionado nivel de desglose de división, una matriz de paso de CNAE2009 a CNAE1993 y otra de CNAE1993 a CNAE2009.

Para facilitar el manejo de la matriz se ha desarrollado la misma de forma 'secuencial' en virtud de la distribución de cada una de las categorías de una clasificación en las de la otra.

Utilización

El uso de las matrices es, conceptualmente, muy sencillo. Se trata de una simple transformación de vectores, multiplicando un vector (el de categorías de actividad según una CNAE) por la matriz resultando otro vector (el de categorías de actividad según la otra CNAE).

Ilustraremos el manejo de la matriz con un ejemplo. Utilizaremos la matriz de paso CNAE1993 a CNAE2009 correspondiente a la actividad CNAE2009=68, *Actividades inmobiliarias*, para estimar la proporción de ocupados en el segundo trimestre de 2007 que trabajaban en esta actividad a partir de la CNAE1993=70, *Actividades inmobiliarias*:

El número de ocupados en esta rama en T2-2007 (resultados INEBASE) es:

Encuesta de Población Activa
Ocupados

Ocupados por sexo y rama de actividad. Valores absolutos y porcentajes respecto del total de cada sexo

Unidades: Porcentaje, Miles Personas




Tabla		Gráfico	
		Ambos sexos	
		Valor absoluto	
		2007TII	
70 Actividades inmobiliarias		181,7	

Utilizaremos el desarrollo secuencial de la matriz para la CNAE93=70:

CNAE-93 Rev.1	CNAE-2009	Porcentaje	Subtotales.
70	41	23,5414	
70	68	60,3939	
70	81	16,0647	
			100,000

La matriz anterior indica que el total de la rama actividades inmobiliarias según CNAE 93 se distribuye un 23,5414% en CNAE2009=41 ('Construcción de edificios'), un 60,3939% en CNAE2009=68 ('Actividades inmobiliarias' según CNAE2009) y un 16,0647 en CNAE2009=81 (Servicios a edificios y actividades de jardinería).

Así pues, el total de ocupados en actividades inmobiliarias en T2-2007 según la CNAE2009 sería:

Resultado: $181,7 \times 60,3939\% = 109,73572$

Aparte del uso propio que cada usuario pueda realizar de estas matrices de paso entre clasificaciones, siempre es posible solicitar a los servicios de información del INE la ejecución de tablas ad-hoc siguiendo los trámites habituales.

Los resultados obtenidos serán en cualquier caso los resultantes de la aplicación de esta matriz de doble codificación a dos dígitos, salvo por eventuales errores de redondeo fruto de la precisión en miles de personas de los datos de INEBASE.

Madrid, 16 de junio de 2014