

ASPECTOS TÉCNICOS DE LAS ESTADÍSTICAS OFICIALES

José Ignacio Alonso Cimadevilla¹

Instituto Nacional de Estadística de León

Resumen

Los múltiples cambios sociales que han tenido lugar en las últimas décadas, han repercutido de forma especial en la elaboración de las estadísticas en nuestro país. En este trabajo se describen las estadísticas que mayor repercusión tienen a nivel económico y que son elaboradas por el organismo oficial de las estadísticas oficiales, el Instituto Nacional de Estadística.

Concretamente, se desarrolla la metodología utilizada en la elaboración de la Encuesta de Población Activa (EPA), cuyos resultados permiten conocer la actividad económica en relación al componente humano. Por otro lado, y respecto al Índice de Precios al Consumo (IPC), se describen los aspectos técnicos del índice base 2006 y base 2011, entendiéndose que éste último constituye una actualización del anterior. El trabajo se completa con los Censos Demográficos 2011, cuya metodología ha experimentado una gran variación. Aunque se describe el proceso de elaboración del Censo de Edificios y del Censo de Viviendas, se concreta de forma más detallada el Censo de Población, considerado como el mayor proyecto estadístico de un país.

Palabras clave: Encuesta de población activa; Índice de precios al consumo; Censo de población.

INTRODUCCIÓN

El Instituto Nacional de Estadística (INE) es un organismo autónomo de carácter administrativo, con personalidad jurídica y patrimonio propio, adscrito al Ministerio de Economía y Competitividad a través de

la Secretaría de Estado de Economía y Apoyo a la Empresa. Se rige, básicamente, por la Ley 12/1989, de 9 de mayo, de la Función Estadística Pública, que regula la actividad estadística para fines estatales, la cual es competencia exclusiva del Estado, y por el Estatuto aprobado por Real Decreto

¹ Delegado Provincial del Instituto Nacional de Estadística. León.

508/2001 de 11 de mayo, y modificado por el Real Decreto 947/2003, de 18 de julio, por el Real Decreto 759/2005, de 24 de junio y por el Real Decreto 950/2009, de 5 de junio.

La Ley asigna al INE un papel destacado en la actividad estadística pública encomendándole expresamente la realización de las operaciones estadísticas de gran envergadura (censos demográficos y económicos, cuentas nacionales, estadísticas demográficas y sociales, indicadores económicos y sociales, coordinación y mantenimiento de los directorios de empresas, formación del Censo Electoral...). También se ocupa de las siguientes funciones: la formulación del Proyecto del Plan Estadístico Nacional con la colaboración de los Departamentos Ministeriales y del Banco de España; la propuesta de normas comunes sobre conceptos, unidades estadísticas, clasificaciones y códigos; y las relaciones en materia estadística con los Organismos Internacionales especializados y, en particular, con la Oficina de Estadística de la Unión Europea (EUROSTAT).

Su creación se debe a la Ley de 31 de diciembre de 1945 (BOE del 3 de enero de 1946), asignándole como misiones principales la elaboración y perfeccionamiento de las estadísticas demográficas, económicas y sociales ya existentes, la creación de otras nuevas y la coordinación con los servicios estadísticos de las áreas provinciales y municipales. Además de regular la coordinación entre otros servicios estadísticos como el Servicio Sindical de Estadística, la Ley crea el Consejo Superior de Estadística y se organiza en Servicios Centrales, Delegaciones provinciales y Delegaciones en los Ministerios.

El comienzo de la estadística oficial en España viene marcado por el Decreto firmado el 3 de noviembre de 1856 por el general Narváez, presidente del Consejo de Ministros de Isabel II, por el que se creaba una Comisión, compuesta por personas de reconocida capacidad, para la formación de la Estadística General del Reino. Unos meses más tarde, el 21 de abril de 1857, la Comisión pasa a denominarse Junta de Estadística. Su primer trabajo es el Censo de Población, con fecha de referencia del 21 de mayo del mismo año. Es la Ley de Instrucción Pública, de 9 de septiembre de 1857, la que establece que la Estadística será una disciplina académica.

Ya en el siglo XX, el Decreto de 1 de octubre de 1901 establece la formación de las estadísticas oficiales y la publicación de las mismas. Se crea la Dirección General y se crean departamentos en los Ministerios para completar su labor. En 1924, el Consejo del Servicio Estadístico, creado en 1921, es reformado, cuatro años antes de que pase a depender del Ministerio de Trabajo y Previsión. En 1931, la adscripción se hace al Ministerio de la Presidencia. Durante la Guerra Civil (1936-1939) comienza a funcionar el Servicio Sindical de Estadística en coordinación con los Servicios de Estadística del Estado, dentro de la zona controlada por los militares sublevados.

Dentro del marco de la Comunidad Económica Europea, el Sistema Estadístico Europeo (SEE) está formado por Eurostat (la oficina de estadística de la UE), las oficinas de estadística de todos los estados miembros y otros organismos que elaboran estadísticas europeas. El SEE garantiza que las estadísticas europeas elaboradas en todos los Estados miembros de la Unión Europea sean fiables, siguiendo unos criterios y definiciones comunes y

tratando los datos de la manera adecuada para que sean siempre comparables entre los distintos países de la UE.

En la actualidad el SEE se regula fundamentalmente por la Ley Estadística Europea, aprobada en 2009 mediante el Reglamento (CE) 223/2009 del Parlamento Europeo y del Consejo. Las estadísticas de la Unión Europea se preparan, elaboran y difunden tanto por el Sistema Estadístico Europeo (SEE) como por el Sistema Europeo de Bancos Centrales (SEBC). El SEE cuenta con un Programa Estadístico Europeo que recoge la planificación estadística para un periodo de cinco años. Este Programa es aprobado por el Parlamento Europeo y por el Consejo y para asegurar que se tengan en cuenta las necesidades de los usuarios en la elaboración del Programa Estadístico Europeo, se creó el Comité Consultivo Europeo de Estadística. En él están representados los usuarios, informantes, instituciones académicas y sociales y la administración comunitaria.

No obstante, la planificación de la actividad se hace conjuntamente entre los INE y Eurostat, la producción de estadísticas nacionales armonizadas corresponde a las autoridades de los Estados miembros, mientras que Eurostat recopila los datos que aportan los Estados, los analiza y en base a ellos ofrece cifras comparables y armonizadas, de forma que se puedan definir, acometer y analizar las políticas comunitarias. Además, Eurostat se encarga de asegurar la coordinación necesaria para garantizar el funcionamiento de este complejo sistema (lenguas diferentes, formas de organización administrativa muy diversa, nomenclaturas específicas...) y para asegurar la coherencia y calidad de los datos.

En este trabajo se van a desarrollar los principales aspectos técnicos de tres de las investigaciones con mayor repercusión de las muchas que se realizan en el INE. En primer lugar, la Encuesta de Población Activa (EPA) como ejemplo de muestreo probabilístico; a continuación, el Índice de Precios de Consumo (IPC) como ejemplo de muestreo no probabilístico y, finalmente, el Censo de Población, operación que, aunque tradicionalmente ha sido dirigida a toda la población, en esta última ocasión, se ha efectuado por muestreo.

1. ENCUESTA DE POBLACIÓN ACTIVA (EPA)

La EPA es una investigación por muestreo de periodicidad trimestral, dirigida a la población que reside en viviendas familiares del territorio nacional y cuya finalidad es averiguar las características de dicha población en relación con el mercado de trabajo.

Se publicó por primera vez en 1964 y hasta finales de 1968 se obtuvieron resultados con referencia trimestral; de 1969 a 1974 la referencia temporal pasó a ser semestral y a partir de 1975 volvió a ser trimestral. En 1987 se modificó el cuestionario de la encuesta para adaptarse a las últimas recomendaciones internacionales de entonces (Conferencia Internacional de Estadísticos del Trabajo) y la exigencia de adaptar la EPA a la Encuesta de Fuerza de Trabajo de la Comunidad Económica Europea con motivo de la incorporación de España en las Comunidades Europeas en 1986. En 1999 se convierte en una 'encuesta continua' dado que las entrevistas se realizan a lo largo de las 13 semanas de cada trimestre y no de 12 de las 13 como se venía haciendo hasta ese momento.

En 2002 se introduce una nueva definición operativa de paro produciéndose una ruptura en las series de parados y activos, cuyo impacto se valoró elaborando doble estimación de ambas definiciones a lo largo del año 2001; en 2005 se produjo el último cambio metodológico sustancial hasta el momento, introduciendo un nuevo cuestionario y el control centralizado del sistema de recogida mediante encuesta telefónica asistida por ordenador. También en este año se calcularon series retrospectivas para el periodo 1996-2004 con la nueva base de población instaurada ese año, con el fin de mantener la homogeneidad de las estimaciones. El periodo anterior quedó sin variaciones. Las cifras actuales de la encuesta se encuadran en la metodología instaurada en 2005.

Su **finalidad** principal es conocer la actividad económica en lo relativo a su componente humano. Está orientada a dar datos de las principales categorías poblacionales en relación con el mercado de trabajo (ocupados, parados, activos, inactivos) y a obtener clasificaciones de estas categorías según diversas características. También posibilita confeccionar series temporales homogéneas de resultados. Por último, al ser las definiciones y criterios utilizados coherentes con los establecidos por los organismos internacionales que se ocupan de temas laborales, permite la comparación con datos de otros países.

Las **ventajas** que presenta son:

- se puede realizar de forma continua con la periodicidad que se desee.
- permite profundizar en los aspectos que interesen en relación con la fuerza

laboral, al ser una investigación enfocada directamente a estos temas.

- entrevistadores especializados se encargan de la cumplimentación de los cuestionarios.
- los resultados se obtienen con rapidez, al ser una encuesta por muestreo.
- las definiciones y tratamiento de la información son uniformes a lo largo de las sucesivas realizaciones de la encuesta, lo que origina series homogéneas de resultados.
- se pueden obtener resultados para el total nacional y para subconjuntos territoriales (fundamentalmente para las comunidades autónomas y las provincias).

El principal inconveniente se deriva de su propia condición de encuesta por muestreo y es el no poder dar información de algunas características con la mayor desagregación posible. Así, por ejemplo, el número de activos de cada una de las sesenta divisiones de la Clasificación Nacional de Actividades Económicas en cada provincia resulta poco fiable, ya que a una mayor desagregación en la información corresponde un mayor error de muestreo.

Las **unidades** son de dos tipos: por una parte, de muestreo, donde las primarias son las secciones censales (áreas geográficas perfectamente delimitadas) y las secundarias son las viviendas; y, por otra, como unidades de análisis se toman las viviendas y las personas. Cubre todo el territorio nacional desde la inclusión de Ceuta y Melilla en el segundo trimestre de 1988 y va dirigida a la población que reside en viviendas familiares principales, es decir, las utilizadas todo el año o la

mayor parte de él como vivienda habitual o permanente.

Se distinguen dos **periodos de referencia**: el de los resultados, que es el trimestre natural y, el periodo de referencia de la información, que es, en general, la semana natural inmediatamente anterior a la de la entrevista según el calendario, aunque hay preguntas con periodos de referencia especiales.

1.1. Principales definiciones en la EPA

Todas las definiciones están basadas en las recomendaciones aprobadas por la Organización Internacional del Trabajo (OIT) en la Decimotercera y Decimosexta Conferencia Internacional de Estadísticos del Trabajo (Ginebra, 1982 y 1998, respectivamente). Por otra parte, todas las características definidas están referidas al concepto nacional y no al interior de acuerdo con las definiciones del Sistema Europeo de Cuentas Nacionales y Regionales (SEC-95). Esto se debe a que no es posible recoger información de la población que trabaja en España y reside en el extranjero, ya que la Encuesta va dirigida a la población que habita en las viviendas familiares del territorio nacional.

- **Población económicamente activa**: Es el conjunto de personas de unas edades determinadas que, en un periodo de referencia dado, suministran mano de obra para la producción de bienes y servicios económicos o que están disponibles y hacen gestiones para incorporarse a dicha producción. Por tanto, la población económicamente activa comprende todas las personas de 16 o más años que durante la semana de referencia (la ante-

rior a aquélla en que corresponde realizar la entrevista según el calendario) satisfacen las condiciones necesarias para su inclusión entre las personas ocupadas o paradas, según se define más adelante.

- **Población ocupada**: Es la formada por todas aquellas personas de 16 o más años que durante la semana de referencia han tenido un trabajo por cuenta ajena o han ejercido una actividad por cuenta propia.
- **Población parada o desempleada**: Se consideran paradas a todas las personas de 16 o más años que reúnen simultáneamente las condiciones de estar sin trabajo, buscarlo y estar disponible para trabajar.
- **Población económicamente inactiva**: Abarca a todas las personas de 16 o más años, no clasificadas como ocupadas ni paradas ni población contada aparte durante la semana de referencia.
- **Población contada aparte**: Los varones que cumplían el servicio militar obligatorio (o servicio social sustitutorio) se consideraban población contada aparte, esto es, no se les incluían entre los activos ni entre los inactivos, independientemente de que en la semana de referencia hubieran trabajado o no. El servicio militar obligatorio desapareció en diciembre de 2001.

Tabla 1. Principales definiciones en la EPA

Menores de 16 años				
Personas de 16 y más años	Activos	Ocupados	Asalariados	del sector público del sector privado
			Trabajadores por cuenta propia	Empleadores Empresarios sin asalariados y trabajadores independientes Miembros de cooperativas Ayudas familiares
			Otros	
	Parados	que buscan primer empleo que han trabajado antes		
	Inactivos	Estudiantes Jubilados o pensionistas Labores del hogar Incapacitados para trabajar Otra situación (rentistas, ...) No sabe		
Población Contada Aparte (PCA)*	que trabaja que no trabaja			
* El año 2002 desapareció el servicio militar obligatorio, por lo que a partir de esa fecha no existe esta categoría poblacional.				

Las principales tasas que se consideran en la EPA son:

- **Tasa global de actividad**, cociente entre el número total de activos y la población total. Se calcula para ambos sexos y para cada uno de ellos por separado.
- La **tasa específica de actividad** para un intervalo de edad determinado, es el cociente entre el número de activos de esas edades y la población correspondiente al intervalo.
- La **tasa global de empleo** es el cociente entre el número total de ocupados y la población total. Se calcula para ambos sexos y para cada uno de ellos por separado.
- La **tasa específica de empleo** para un intervalo de edad determinado es el cociente entre el número de ocupados

de esas edades y la población correspondiente al intervalo.

- La **tasa de paro** es el cociente entre el número de parados y el de activos. Se calcula para ambos sexos y para cada uno de ellos por separado.
- La **tasa específica de paro** para un intervalo de edad determinado, que es el cociente entre los parados de edades comprendidas entre los extremos del intervalo y los activos de dicho intervalo.
- **Tasa de asalarización**, cociente entre el número de asalariados y el número total de ocupados.
- **Tasa de temporalidad**, cociente entre el número de asalariados con contrato temporal y el número total de asalariados.
- **Tasa de trabajo a tiempo parcial**, cociente entre el número de ocupados a

tiempo parcial y el número total de ocupados.

1.2. El diseño de la muestra

Tomando como referencia el marco de la encuesta, se utiliza un muestreo bietápico con estratificación de las unidades de primera etapa. Las unidades de primera etapa están constituidas por las secciones censales, cuya muestra permanece fija indefinidamente salvo pocas excepciones. Siempre que una sección sale de la muestra es sustituida por otra seleccionada aleatoriamente.

Las unidades de segunda etapa están constituidas por las viviendas familiares principales (ocupadas permanentemente) y los alojamientos fijos (chabolas, cuevas, etc.). No se consideran encuestables las viviendas secundarias (ocupadas sólo una parte del año), ni las disponibles para alquiler o venta, ya que no forman parte del ámbito poblacional definido anteriormente. Dentro de las unidades de segunda etapa no se realiza submuestreo alguno, recogiendo información de todas las personas que tengan su residencia habitual en las mismas.

Las unidades de primera etapa se estratifican atendiendo a un doble criterio:

- Criterio geográfico (de estratificación). Las secciones se agrupan en estratos dentro de cada provincia, de acuerdo con la importancia demográfica del municipio al que pertenecen.
- Criterio socioeconómico (de subestratificación). Las secciones censales se agrupan en subestratos dentro de cada uno de los estratos, según las características socioeconómicas de las mismas.

Para llegar a la formación de los estratos se consideran los siguientes tipos de municipios:

- **Municipios autorrepresentados:** Son aquellos que dada su categoría dentro de la provincia deben tener siempre secciones en la muestra. Se consideran así la capital de la provincia y aquéllos que tienen un número de habitantes tal, que en la afijación proporcional dentro de la provincia le corresponden al menos 12 secciones en la muestra, así como los que teniendo una situación demográfica destacada dentro de la provincia no hay otros similares con que agruparlos, aunque proporcionalmente le correspondan menos de 12 secciones en la muestra.
- **Municipios correpresentados:** Son aquellos que dentro de la misma provincia forman parte de un grupo de municipios demográficamente similares y que son representados en común.

De acuerdo con esta clasificación se consideran nueve estratos, de tal manera que en los tres primeros se encuentran las tres categorías de municipios autorrepresentados y en el resto los correpresentados, abarcando desde el estrato 4 los municipios entre cincuenta y cien mil habitantes hasta el estrato 9 los menores de dos mil habitantes. Cada diez años, con la información procedente de los Censos de Población, se actualiza la definición de los estratos en cada provincia.

Tamaño de la muestra: Para determinar el tamaño de la muestra, se aplica un procedimiento de mínima varianza para coste fijo. Se parte de un presupuesto (Q) y a partir de él se procede a determinar el número de secciones (n) y el número de viviendas (m) que minimizan la varianza de las estimaciones. Para ello se utiliza una

función de coste de tipo lineal y la expresión del coeficiente de variación para una proporción en el muestreo de conglomerados con submuestreo.

- La función de coste considerada es:
 $Q = n Q_S + n m Q_V$, con
 $Q_S = Q_F + d Q_D$, donde:

Q = Presupuesto total; Q_S = Coste por unidad primaria (sección); Q_V = Coste por unidad última (vivienda)

n = Número de secciones; m = Número de viviendas por sección

Q_F = Coste fijo por sección; Q_D = Coste diario del trabajo de campo

d = Número de días necesarios para el trabajo de campo

- El coeficiente de variación para una proporción P viene dado por la siguiente expresión:

$$CV^2(\hat{P}) = \frac{V(\hat{P})}{\hat{P}^2} = \frac{1-\hat{P}}{\hat{P}} \frac{1+\delta(m-1)}{nm} = \frac{1-\hat{P}}{\hat{P}} F(\delta, m, n)$$

siendo $F(\delta, m, n) = \frac{1+\delta(m-1)}{nm}$

siendo δ el coeficiente de correlación intraclásica, que para el caso de la población activa se ha calculado y vale 0,05. El mínimo de la expresión $CV^2(\hat{P})$ respecto de las variables m y n se obtiene calculando el mínimo de la expresión $F(\delta, m, n)$ que es independiente de \hat{P} . Para distintos valores de m compatibles con el trabajo de campo, $m = 4, 6, 8, 10, 11, 14, 17, 18, 19, \dots, 91, 100$ y los correspondientes valores de n , dados por

$$n = \frac{Q}{Q_S + m Q_V}$$

se obtienen distintos valores para $F(\delta, m, n)$.

El valor mínimo de $F(\delta, m, n)$ respecto de m y n corresponde a $m=20$ y $n=3000$ y en base a este resultado, la muestra se fija en un total de 3.000 secciones, investigán-

dose una media de 20 viviendas por sección. Posteriormente, la muestra ha tenido diferentes ampliaciones con el objeto de dar cumplimiento a las exigencias de la Unión Europea y mejorar la representación de áreas más desagregadas. A partir del primer trimestre de 2005 se establece un tamaño muestral de 3.588 secciones y 18 viviendas por sección, excepto en las provincias de Madrid, Barcelona, Sevilla, Valencia y Zaragoza, en las cuales el número de entrevistas por sección es de 22. En el tercer trimestre de 2009 se firma un convenio con la Comunidad Autónoma de Galicia por el que se incrementa la muestra en esta comunidad hasta un total de 468 secciones y se asignan estratos separados a los municipios de Santiago de Compostela y Ferrol. El tamaño de muestra final para el total nacional se establece en 3.822 secciones.

Para distribuir las secciones de la muestra entre las provincias, dentro de la provincia entre estratos y dentro de éstos entre substratos (afijación de la muestra), se tienen en cuenta los siguientes aspectos:

- Los resultados nacionales deben tener la mayor fiabilidad posible. A este respecto hay que recordar que, en general, cuanto más lejos se esté de la afijación óptima por provincias y estratos, mayor es la pérdida de precisión en la estimación nacional para un tamaño fijo de la muestra.
- Disponer en cada provincia de un tamaño mínimo de muestra que permita dar estimaciones de la misma.
- En cada provincia el número de secciones debe ser múltiplo de trece. Con ello se facilita la distribución de la muestra entre las semanas del trimestre.

Para compatibilizar las tres condiciones antes expuestas se ha adoptado una

afijación de compromiso entre la uniforme y la proporcional, a base de agrupar provincias de importancia demográfica similar y asignarles de 3 a 12 entrevistadores, es decir, de 39 a 156 secciones muestrales (con las excepciones de Ceuta y Melilla, que debido a su reducido tamaño poblacional tienen un solo agente y por tanto 13 secciones en la muestra cada una de ellas).

Dentro de cada provincia la afijación entre estratos es proporcional al tamaño de cada uno de ellos, si bien se han potenciado los estratos donde se encuentran los municipios de mayor tamaño, ya que se espera que la mayor parte de las características que se estudian estén correlacionadas con los niveles económico-sociales y culturales de los habitantes y es precisamente en estos estratos donde, en general, la dispersión debe ser mayor y donde el costo por entrevista es menor. Dentro de los estratos, la afijación entre subestratos es estrictamente proporcional al tamaño (medido en número de viviendas familiares).

Selección de la muestra: se realiza de forma que dentro de cada estrato cualquier vivienda familiar tenga la misma probabilidad de ser seleccionada, es decir, se tengan muestras autoponderadas dentro de cada estrato. Este tipo de muestras proporciona pesos de diseño iguales por estrato en los estimadores. Para ello, las unidades de primera etapa (secciones censales) se seleccionan con probabilidad proporcional al número de viviendas familiares principales, según los datos del último Censo o del Padrón Continuo. Dentro de cada sección seleccionada en primera etapa, se selecciona un número fijo de viviendas familiares con igual probabilidad, mediante la aplicación de un muestreo sistemático con arranque aleatorio. Para esta encuesta se ha

determinado seleccionar 18 viviendas por sección.

- La probabilidad de selección de la vivienda i , perteneciente a la sección j del estrato h , donde se han afijado K_h secciones, será:

$$P(V_{ijh}) = P(S_{jh}) \cdot P\left(\frac{V_{ijh}}{S_{jh}}\right) = K_h \cdot \frac{V_{jh}}{V_h} \cdot \frac{18}{V_{jh}} = K_h \cdot \frac{18}{V_h}$$

$P(S_{jh})$ = Probabilidad de selección de la sección j del estrato h

$P\left(\frac{V_{ijh}}{S_{jh}}\right)$ = Probabilidad de selección de la vivienda i condicionada a la selección de la sección j .

V_{jh} = Total de viviendas de la sección j ;
 V_h = Total de viviendas del estrato h .

Esta probabilidad no depende ni de la vivienda (i) ni de la sección (j), por lo tanto, la muestra es autoponderada.

La muestra se distribuye uniformemente en el tiempo, lo que equivale a que en cada provincia el número de secciones por semana es constante. Además, se ha procurado que la distribución de secciones muestrales por provincia, estrato y semana sea homogéneo, al igual que por provincia, turno de rotación y semana. Cada periodo de la encuesta es de un trimestre siendo cada una de las secciones de la muestra visitada en una de las 13 semanas del mismo. La totalidad de la muestra está dividida en tres submuestras independientes representativas, cada una de ellas, de toda la población.

Como se ha dicho en el párrafo anterior, cada periodo de la encuesta es de un trimestre, repitiéndose ésta sucesivamente. Las secciones censales permanecen fijas en la muestra indefinidamente (salvo las excepciones), sin embargo las viviendas

familiares son renovadas parcialmente cada trimestre de encuesta, a fin de evitar el cansancio de las familias. Esta renovación se efectúa en una sexta parte de las secciones.

A estos efectos, la muestra total se halla dividida en seis submuestras que denominamos *Turnos de rotación*. Cada trimestre se renuevan las viviendas que pertenecen a las secciones de un determinado turno de rotación. Por tanto cada vivienda permanece en la muestra durante seis trimestres consecutivos, al cabo de los cuales sale de la misma para ser reemplazada por otra de la misma sección.

La distribución del número de secciones por estrato y semana es similar en cada turno de rotación. De esta forma se trata de evitar posibles sesgos de medida en las estimaciones, debidos al diferente comportamiento de las familias colaboradoras en función del tiempo que lleven en la encuesta.

Cada turno de rotación puede ser considerado, por tanto, como una submuestra representativa. Este hecho facilita la obtención de estimaciones de variables estructurales mediante la unión de dichas submuestras.

Hasta el año 2001, se utilizaron estimadores de razón tomando como variable auxiliar las Proyecciones Demográficas de población elaboradas por el INE, siendo la expresión del estimador de una determinada característica Y en un trimestre de encuesta la siguiente:

$$\hat{Y} = \sum_h \frac{P_h}{p_h} \sum_{i=1}^{n_h} y_{hi},$$

extendiéndose el sumatorio h a los estratos de una provincia, una

comunidad autónoma o al total nacional, y donde:

P_h : es la proyección de la población, que reside en viviendas familiares, en el estrato h , referida a la mitad del trimestre.

p_h : es el número de personas que habitan en las viviendas de la muestra, en el estrato h , en el momento de la entrevista.

n_h : es el número de viviendas en las secciones de la muestra en el estrato h .

y_{hi} : es el valor de la característica investigada en la vivienda i -ésima, del estrato h .

Las continuas variaciones de población, bien en sus características, bien en su distribución espacial exigen realizar actualizaciones en el marco que necesariamente repercuten en la estructura muestral. En el marco de la EPA se consideran cuatro tipos de actualizaciones: las relativas a las secciones muestrales, consecuencia de las modificaciones producidas por diversas incidencias como particiones, fusiones o variaciones de límites en las secciones seleccionadas; las relativas a las viviendas; las correspondientes a las probabilidades de selección y las relativas a todas las secciones y viviendas de la población.

La forma más directa de actualizar las probabilidades de selección del seccionado es la selección de una nueva muestra a partir del marco disponible más actualizado. Pero un cambio tan radical en una encuesta continua, como es la EPA, genera tres tipos de problemas:

- Pérdida de información imprescindible para la selección y visita de las viviendas que resulten elegidas en la segunda etapa. Esta información, que es necesario rehacer, tiene aspectos tangibles como los directorios de viviendas o la planimetría de la zona, y otros intangibles pero no

menos importantes, como el conocimiento por parte de la población de la sección de la figura del entrevistador, hecho éste que facilita el acceso a las familias y disminuye notablemente la falta de respuesta.

- Pérdida de precisión en las estimaciones de variaciones trimestrales interanuales, al disminuir considerablemente la muestra común entre ambos periodos.
- Posible presencia de discontinuidades en la serie temporal de la encuesta, debidas a la causa citada en el apartado anterior.

A causa de ello se decidió arbitrar un procedimiento que, sin distorsionar las probabilidades de selección que realmente corresponden a cada sección, mantenga la muestra de secciones con las mínimas variaciones, considerando dos tipos de actualizaciones de las probabilidades de selección en función de la información disponible para las mismas.

Errores: los errores que afectan a toda encuesta pueden agruparse en dos grandes grupos:

- **Errores de muestreo**, que se originan por la obtención de resultados sobre las características de una población, a partir de la información recogida en una muestra de la misma.
- **Errores ajenos al muestreo**, que son comunes a toda investigación estadística, tanto si la información es recogida por muestreo como si se realiza un Censo. Estos errores se presentan en cualquier fase del proceso estadístico, tanto antes como durante o tras la recogida de datos: deficiencias del marco e insuficiencias en las definiciones y cuestionarios, por defectos en la labor de los entrevistadores e incorrecta declaración por parte de los informantes o en la depuración, codificación, grabación o tabulación de los resultados.

- Los errores de muestreo se calculan trimestralmente de las estimaciones de algunas de las principales características investigadas. Para la obtención de los errores de muestreo se utiliza el método de las semimuestras reiteradas, consistente en obtener sucesivas semimuestras de la muestra inicial. A partir de cada semimuestra se calcula la estimación de la característica de la que queremos obtener el error de muestreo.

Una vez calculadas todas las estimaciones con cada una de las semimuestras, así como la estimación con la muestra completa, el estimador de la varianza viene dado por:

$$\hat{V}(\hat{Y}) = \frac{1}{r} \sum_{i=1}^r (\hat{y}_i - \hat{Y})^2, \text{ donde:}$$

r : es el número de semimuestras obtenidas, esto es, el número de reiteraciones.

\hat{y}_i : es la estimación obtenida con la i -ésima reiteración. Para cada reiteración se repite el proceso de estimación general, es decir, se aplica la técnica de reponderación (software CALMAR).

\hat{Y} : es la estimación basada en la muestra completa.

En el caso de la EPA el número de reiteraciones que se utiliza es de 40. Para formarlas se agrupan todas las secciones de cada estrato por pares, procurando que las dos secciones de cada par pertenezcan al mismo turno de rotación de la EPA. A continuación se asigna aleatoriamente la primera sección de cada par a 20 reiteraciones y la otra sección a las otras 20. De esta forma cada reiteración queda constituida por un número de secciones equivalente al 50 por ciento de la muestra (semimuestra) y cada sección

aparece en la mitad de las reiteraciones. En las tablas se publica el error de muestreo relativo en porcentaje (coeficiente de variación).

- El estudio de los errores ajenos al muestreo presenta numerosas dificultades debido a la gran variedad de causas que los originan, así como a las hipótesis en que se basan los modelos teóricos que, en general, no se cumplen en la realidad, lo que lleva a obtener resultados aproximados.

En la EPA el análisis de los errores ajenos al muestreo se basa en el modelo matemático elaborado por la Oficina de Censos de los Estados Unidos, debido a Hansen, Hurwitz y Bershad (1961), y que, operativamente, consiste en repetir las entrevistas de la encuesta en una submuestra de la muestra de viviendas originalmente seleccionada. Posteriormente se cotejan los datos obtenidos en ambas ocasiones, con objeto de investigar las inconsistencias y cuantificar los errores mediante la aplicación de diversos índices de calidad.

Esta encuesta de evaluación persigue un doble objetivo: por una parte, controlar el trabajo de recogida de la información en todas las comunidades autónomas y, por otra, evaluar la calidad de los resultados.

La comparación de los resultados obtenidos en la encuesta de evaluación (entrevista repetida, ER) con los obtenidos en la entrevista original (EO) permite evaluar dos grandes tipos de errores ajenos al muestreo:

- **Errores de cobertura**, producidos por la omisión o por la inclusión errónea de unidades (viviendas y personas) en la encuesta original.

- **Errores de contenido**, que afectan a las características investigadas en las personas encuestables.

El trabajo de campo se lleva a cabo por agentes especializados, los cuales realizan la entrevista repetida a lo sumo tres semanas después de la original, refiriéndose los datos de ambas entrevistas al mismo periodo de tiempo.

El hecho de que más del 70 por ciento de las negativas por primera vez se producen en la primera entrevista a las familias, unido a la existencia de dificultades técnicas para la realización de la encuesta de evaluación (ER) mediante entrevistas telefónicas, han determinado que se investiguen en ER únicamente secciones que en EO se encuentran en primera entrevista. El método de recogida utilizado en estas secciones, tanto en EO como en ER, es de visita domiciliaria. Como consecuencia se dispone de menos muestra en la encuesta de evaluación, respecto a años anteriores, por lo que las cuatro muestras trimestrales se van a agrupar para ofrecer los resultados en cómputo anual, a fin de que éstos sean más representativos.

Para la selección trimestral de la muestra de la encuesta de evaluación se han creado cuatro zonas, agrupando en cada una varias comunidades autónomas, de forma que cada una de éstas esté incluida en una y sólo una de las zonas. Cada semana se investigan las secciones (de primera entrevista) de la muestra en una de las zonas, siendo aleatoria la asignación de las semanas a las zonas, de modo que cada una de éstas se investigue al menos en tres de las semanas del trimestre. De este modo se investigan aproximadamente entre 130 y 150 secciones cada trimestre.

En las secciones seleccionadas se repite la entrevista en la mitad de las viviendas, utilizándose en ER un cuestionario ligeramente reducido respecto al de EO, es decir, con algunas preguntas menos. Con este procedimiento se investiga un número de viviendas de entre 1.300 y 1.500, lo que representa aproximadamente un 2 por ciento de la muestra de la EPA.

Además de la encuesta de evaluación, y con objeto de detectar errores cometidos en el proceso de actualización de las secciones de la muestra, cada trimestre se selecciona una muestra de cincuenta secciones (una de cada provincia, salvo Ceuta y Melilla) para evaluar la calidad de las actualizaciones.

Aparte de la entrevista repetida se realiza un estudio específico de aquellas unidades seleccionadas que son encuestables pero que se negaron a facilitar los datos solicitados. Para estas unidades que se niegan a colaborar en la encuesta se cumplimenta un cuestionario de negativas, en el que se recogen una serie de características básicas, como son el sexo, la edad y la relación con la persona principal de la persona que rehúsa ser entrevistada, así como la edad, el sexo, la nacionalidad, los estudios terminados, la relación con la actividad, la rama de actividad y la ocupación de la persona principal.

2. EL ÍNDICE DE PRECIOS DE CONSUMO (BASE 2006)

La operación del cambio de Sistema del Índice de Precios de Consumo (IPC) consiste, fundamentalmente, en revisar y actualizar cada uno de sus componentes y determinar las mejores opciones para conseguir un indicador representativo y

preciso que se adapte a las tendencias de la economía.

Hasta la entrada en vigor de la base 2001, el IPC basaba su cálculo en lo que se denomina sistema de base fija, cuya principal característica es que tanto la composición de la cesta de la compra como sus ponderaciones se mantienen inalterables a lo largo del tiempo que dura la base. Los cambios de base se llevaban a cabo cada ocho o nueve años, debido a que ésa era la periodicidad de la Encuesta Básica de Presupuestos Familiares (EBPF), la fuente utilizada para la elaboración de las ponderaciones y de la cesta de la compra.

A partir de 1997, las dos encuestas de presupuestos familiares que convivían (una continua, con periodicidad trimestral, y una básica, que se realizaba cada ocho o nueve años) fueron sustituidas por una sola, con periodicidad trimestral, que proporcionaba una información más cercana a la encuesta básica, en cuanto al nivel de desagregación. Esta encuesta, denominada Encuesta Continua de Presupuestos Familiares (ECPF), proporcionó la información necesaria para la actualización de las ponderaciones así como la renovación de la composición de la cesta de la compra en el cambio de base del IPC 2001. Y, además, posibilitó la actualización permanente de dichas ponderaciones y la revisión de la cesta de la compra, lo que supuso una mejora en los cambios de Sistema del IPC.

El IPC base 2001 se actualizó con la revisión permanente de su sistema metodológico, mejorándolo contactando con los distintos foros académicos y organismos productores nacionales e internacionales. Pero, también, es más dinámico que sus predecesores en la medida en que anualmente revisa las

ponderaciones para ciertos niveles de desagregación funcional e incluye en el plazo más breve posible cualquier cambio detectado en los componentes del mercado, ya sea la aparición de nuevos productos, cambios en la estructura de consumo o en la muestra de municipios o establecimientos. Además, establece los cambios de base cada cinco años, realizando una revisión completa de la metodología y la muestra y la actualización de ponderaciones a todos los niveles de desagregación.

Como consecuencia de este nuevo esquema de funcionamiento, en enero de 2007, entra en vigor el Sistema de Índices de Precios de Consumo, con base de referencia en el año 2006. Este Sistema sustituye al IPC que, con base 2001, estuvo vigente hasta diciembre de 2006. El IPC, base 2006, mantiene las principales características del IPC, base 2001, y, al igual que éste, revisa anualmente las ponderaciones para cierto nivel de desagregación funcional. Para realizar esta actualización utiliza la información proporcionada por la nueva Encuesta de Presupuesto Familiares (EPF) que, desde el año 2006, sustituye a la ECPF-97 y cuya principal característica es su periodicidad anual. Asimismo la información proporcionada por esta nueva encuesta también se utilizará en los cambios de base posteriores al año 2008.

El Índice de Precios de Consumo, que se publica mensualmente, tiene como objetivo medir la evolución del nivel de precios de los bienes y servicios de consumo adquiridos por los hogares residentes en España. En el Sistema Base 2006 se utiliza la siguiente definición de gasto de consumo de la EPF: "el gasto de consumo es el flujo monetario que destina el hogar y cada uno de sus miembros al pago de determinados bienes y servicios,

con destino al propio hogar o para ser transferidos gratuitamente a otros hogares o instituciones".

Las aplicaciones del IPC son numerosas y de gran importancia en los ámbitos económico, jurídico y social. Entre ellas cabe destacar su utilización como medida de la inflación. También se aplica en la revisión de los contratos de arrendamiento de inmuebles, como referencia en la negociación salarial, en la fijación de las pensiones, en la actualización de las primas de seguros y otros tipos de contrato, y como deflactor en la Contabilidad Nacional.

El **campo de consumo** es el conjunto de los bienes y servicios que los hogares del estrato de referencia (población que reside en viviendas familiares en España) destinan al consumo; por lo tanto no se consideran los gastos en bienes de inversión, los autoconsumos y autosuministros, ni los alquileres imputados, ni los gastos subvencionados por las administraciones públicas. Tampoco forman parte del campo de consumo algunos impuestos no considerados consumo desde el punto de vista de la EPF ni otros gastos, como los destinados a loterías y juegos de azar.

La **cesta de la compra** es el conjunto de los bienes y servicios seleccionados en el IPC cuya evolución de precios representa la de todos aquellos que componen la parcela COICOP a la que pertenecen. La selección de los artículos que componen la cesta de la compra se realiza a partir del IPC, base 2001, y los datos de la ECPF 2004-2005. El criterio para determinar qué parcelas deben estar incluidas sigue siendo el mismo que para la base 2001: se han tenido en cuenta todas aquellas parcelas que superan el 0,3 por mil del gasto total. Una vez determinadas las parcelas de gasto que están representadas en el

índice, se revisan los artículos que componen la cesta de la compra de la base 2001, aumentando, disminuyendo o manteniendo los artículos de cada parcela, en función de la ponderación de ésta y de la variabilidad de los precios de dichos artículos. Así, el número total de artículos que componen la cesta de la compra del IPC base 2006 es 491. Para cada uno de los artículos se elabora su descripción o especificación con el fin de facilitar su identificación por parte del encuestador y permitir la correcta recogida de los precios. Estas especificaciones tienen en cuenta las particularidades propias de cada región.

El IPC base 2006 se adapta completamente a la clasificación internacional de consumo COICOP. Los artículos de la cesta de la compra se agregan en subclases, éstas en clases, posteriormente en subgrupos, y por último, los subgrupos en grupos. La estructura funcional del IPC consta de 12 grupos, 37 subgrupos, 79 clases y 126 subclases. Además, se mantienen las 57 rúbricas y los 28 grupos especiales existentes en el IPC base 2001.

Los artículos están distribuidos en los doce grandes grupos siguientes:

1	Alimentos y bebidas no alcohólicas	176
2	Bebidas alcohólicas y Tabaco	12
3	Vestido y calzado	67
4	Vivienda	18
5	Menaje	60
6	Medicina	13
7	Transporte	31
8	Comunicaciones	3
9	Ocio y cultura	43
10	Enseñanza	7
11	Hoteles, cafés y restaurants	23
12	Otros bienes y servicios	38

Total	491
-------	-----

Como en la mayoría de los países de la Unión Europea (UE), el diseño de la muestra de los precios que intervienen en el cálculo del IPC es intencional, y por tanto se trata de un diseño no probabilístico, dadas las características de la población objeto de estudio.

2.1. Selección de la muestra

Para obtener indicadores significativos en todos los niveles de desagregación funcional y geográfica para los que se publica el IPC, se estructura el proceso de selección

de la muestra en tres grandes apartados, cada uno de los cuales tiene como objetivo la selección de los diferentes componentes de la misma, a saber: municipios, zonas comerciales y establecimientos y artículos.

La selección de los **municipios** que forman parte del nuevo Sistema de IPC se realiza atendiendo a criterios demográficos y a la representatividad geográfica. Los datos oficiales de población que se han utilizado para realizar la selección de municipios son los obtenidos de la revisión del Padrón

Municipal de Habitantes a 1 de enero de 2003. Se parte de los criterios demográficos que se utilizaron en el IPC base 2001, y se han introducido algunos adicionales, con el fin de obtener indicadores representativos para cada nivel de desagregación funcional y geográfica.

En la base 2001, el criterio de cobertura geográfica se basaba, principalmente, en la población del conjunto de municipios seleccionados. De esta forma, los municipios seleccionados debían cubrir el 30% de la población de la provincia y el 50% de la población de la comunidad autónoma. Con este criterio se iban seleccionando los municipios por tamaño hasta cumplir el requisito, sin tener en cuenta la distribución geográfica de los mismos dentro de la provincia. Además, la muestra de municipios donde se realizaba la recogida de precios de artículos de alimentación era siempre mayor que para el resto de artículos.

Para la base 2006 se ha mantenido este criterio de partida, pero se ha completado de la siguiente forma:

- representatividad geográfica: es importante que los municipios de la muestra estén repartidos por toda la provincia, evitando la concentración en determinados focos de población.
- representatividad poblacional: se hace especial hincapié en la representatividad de los municipios pequeños; hasta ahora, dado que el criterio de selección era exclusivamente poblacional, quedaba excluida de la muestra parte de la población residente en municipios de menor tamaño.
- representatividad de la cesta: todos los municipios deben contener artículos de todos los grupos; para ello se ha elaborado una cesta reducida, a partir de la cesta total, en la que se han

incluido artículos de consumo básicos. Con ello, se ha aumentado considerablemente la representatividad del IPC.

Así, la muestra de municipios obtenida con los criterios antes citados consta de 177 (las 52 capitales de provincia y 125 municipios no capitales), frente a los 141 municipios de la Base 2001.

En 97 de estos 177 municipios se recogen precios de toda la cesta de artículos, en 44 se recogen precios de toda la cesta de alimentación y parte del resto de la cesta, y en los 36 restantes se recogen precios de una parte reducida de la cesta (compuesta por el 48% de los artículos). Estos últimos son los nuevos municipios que han entrado a formar parte del IPC base 2006.

Por lo tanto, en la base 2006 se ha mejorado la representatividad y diversificación poblacional, ya que de los 36 municipios nuevos, 31 tienen menos de 50.000 habitantes; se ha mejorado la cobertura de bienes y servicios, ya que en todos los municipios se recogen precios de todos los grupos de consumo.

Es importante destacar que en la práctica se cubren porcentajes de población superiores a los indicados, ya que algunos establecimientos incluidos en la muestra, como hipermercados, centros comerciales, talleres de reparaciones o tiendas de muebles, se encuentran en las afueras de los municipios o en municipios limítrofes, con lo que la población real representada en el índice es mayor que la teórica.

Para la selección del **número de establecimientos**, se utiliza como punto de partida la muestra del IPC base 2001, y se estudia la red existente de establecimientos disponibles en cada provincia, prestando especial atención a los distintos tipos y características de dichos establecimientos.

Como criterio general, el número de establecimientos que mensualmente informan de los precios de un artículo se calcula en función de la ponderación del artículo en el índice y de la variabilidad de sus precios: cuanta más ponderación y/o variabilidad de precios, mayor número de establecimientos se deben seleccionar. Además, para el cálculo del número de establecimientos, se establece un número mínimo para cada artículo en cada provincia, dependiendo del tipo de artículo y del tipo de recogida del mismo.

En relación a la selección de los tipos de establecimiento, se considera la distribución de los porcentajes de ventas por tipo de establecimiento (hipermercados, supermercados, mercados y tiendas especializadas), dependiendo de cada artículo. Para ello se cuenta con información de diversas fuentes, entre ellas la Encuesta de Comercio (INE) y el Ministerio de Agricultura, Pesca y Alimentación.

Se dedica especial atención a los centros comerciales, hipermercados y supermercados, dada su importancia en cuanto al volumen de ventas, si bien este aspecto ya se recogía en el IPC base 2001. En muchos casos la situación de dichos centros, así como la presencia de mercados, condiciona la creación de las "zonas comerciales" que de forma explícita se definen, en cada municipio de la muestra, para los artículos perecederos de alimentación (carnes, pescados, frutas y hortalizas frescas), e implícitamente para el resto de artículos.

Para los artículos perecederos de alimentación se definen tres tipos de áreas comerciales, atendiendo al tamaño del municipio y al número de establecimientos susceptibles de ser seleccionados para los distintos tipos de artículos considerados. Estos artículos se clasifican en dos grandes grupos atendiendo a la variabilidad de los

precios que presentan y al peso que tienen en la cesta de la compra. Esta clasificación determina el número de establecimientos en los que se recogen precios según el tipo de zona comercial y el tipo de artículo de que se trate.

Para el resto de artículos, aunque no se ha efectuado la delimitación estricta de las zonas comerciales, la selección de establecimientos se hace cumpliendo el objetivo de representatividad: la muestra de establecimientos debe representar, con la evolución de los precios de los artículos que en ellos se venden, a todos los establecimientos de la localidad.

Partiendo de las premisas anteriores, la selección de los establecimientos informantes se realiza tomando en consideración las siguientes normas básicas:

- En la muestra deben estar representadas todas las zonas comerciales y los distintos tipos de establecimientos que existen.
- Los establecimientos deben ser los más frecuentes y los de mayor afluencia de público en la localidad, y/o los de mayor volumen de venta.
- Los establecimientos deben ser representativos del tipo de artículo del cual se recoge información.
- En cada establecimiento no se pueden recoger, en el mismo día, más de un precio del mismo artículo.
- Un establecimiento no debe concentrar un número importante de observaciones de precios de diferentes artículos. Se intenta evitar con esto que la política de precios de un solo establecimiento pueda condicionar la evolución del índice.
- No forman parte de la muestra los establecimientos de acceso restringido a un sector de la población como cooperativas, economatos o establecimientos

similares. Tampoco se considera la venta ambulante ni la venta a domicilio.

- Los establecimientos seleccionados han de ofrecer suficientes garantías de continuidad en la venta de artículos de los cuales se recogen precios, ya que esta muestra permanecerá fija a lo largo del tiempo, salvo que se produzca el cierre, el cambio de actividad, la pérdida de representatividad en lo que al consumo se refiere o se deje de comercializar el artículo del cual se recogían precios. En estos casos, el establecimiento será sustituido por otro que cumpla los requisitos necesarios para pertenecer a la muestra.

En el IPC base 2006 se han seleccionado aproximadamente 33.000 establecimientos con estas características, repartidos por todo el territorio nacional.

Para seleccionar los **artículos** representativos de las parcelas de gasto de la ECPF se consultan diferentes organismos, asociaciones de empresarios, fabricantes, comerciantes y establecimientos, los cuales facilitan información de aquellos artículos que mejor representan las distintas parcelas, de acuerdo con los siguientes criterios de selección:

- La evolución de los precios de los artículos seleccionados debe ser similar a la del resto de los artículos de la parcela a la que representan.
- Los artículos deben ser los consumidos habitualmente por la población.
- Deben tener precios que sean fácilmente observables.
- Deben ofrecer garantías razonables de permanencia en el mercado.

Así, en el IPC base 2006, la cesta de la compra está compuesta por 491 artículos, frente a los 484 con los que se elaboraba el IPC base 2001. Se han eliminado artículos cuyo consumo o representa-

tividad había disminuido, como el tejido para confección o la reparación de ciertos electrodomésticos, y se han incorporado nuevos artículos como los relacionados con los productos dietéticos e infantiles y nuevos servicios relacionados con la salud, entre otros.

2.2. Cálculo

La fórmula empleada para calcular los índices del IPC, base 2006, es la fórmula de Laspeyres encadenado, que se empezó a utilizar en el IPC, base 2001. El índice general correspondiente al mes m del año t se expresa matemáticamente del siguiente modo, y, al ser encadenado establece comparaciones entre el periodo corriente (t) y el periodo base (0), pero considerando las situaciones intermedias (k).

$$I_0^t = \prod_{k=1}^t \frac{\sum_i p_i^k q_i^{k-1}}{\sum_i p_i^{k-1} q_i^{k-1}}$$

- El principal inconveniente de los índices encadenados es la falta de aditividad. Esto hace que no sea posible obtener el índice de cualquier agregado como media ponderada de los índices de los agregados que lo componen. Así, por ejemplo, el índice general no se puede calcular como media ponderada de los índices de los doce grupos.
- En el IPC español se calcula un índice elemental para cada artículo de la cesta de la compra en cada una de las provincias, por lo que el agregado elemental es el artículo-provincia.
 - El índice del agregado elemental i se obtiene como cociente del precio medio de dicho agregado elemental en el periodo actual y el

precio medio en el periodo de referencia de los precios, es decir, diciembre del año anterior.

$$I_{dic(t-1),i}^{mt} = \frac{\bar{p}_i^{mt}}{\bar{p}_i^{dic(t-1)}} \cdot 100, \text{ donde:}$$

$I_{dic(t-1),i}^{mt}$ es el índice, referido a diciembre del año $(t-1)$, del agregado elemental i , en el mes m del año t .

\bar{p}_i^{mt} es el precio medio del agregado elemental i , en el mes m del año t .

$\bar{p}_i^{dic(t-1)}$ es el precio medio del agregado elemental i , en diciembre del año $(t-1)$.

- El precio medio del agregado i , en el periodo (m,t) es la media geométrica simple de los precios recogidos en dicho periodo y otorga la misma importancia a las variaciones de todos los precios, independientemente del nivel de los mismos.
- Las ponderaciones que intervienen en el cálculo de los índices agregados provienen de la EPF.

En el cálculo de las mismas, para los artículos que componen la cesta de la compra del IPC, base 2006, ha sido necesaria la desagregación de estas parcelas para obtener información más detallada. Los datos empleados en el cálculo de las ponderaciones, utilizadas durante el año 2007, son los correspondientes a los años 2004 y 2005. Para obtener el gasto total se calculó una media aritmética ponderada de las estructuras anuales, donde el segundo año tiene asignado un peso mayor que

el primero por estar más próximo al momento de actualización.

Además, para corregir el desfase que se producía entre este periodo de referencia de las ponderaciones y el periodo de referencia de los precios (diciembre del año 2006) se actualizaron las ponderaciones mediante la utilización de información sobre evolución de precios y de cantidades, procedente del IPC y de otras fuentes. De esta forma, el periodo de referencia de las ponderaciones, utilizadas durante el año 2007, es diciembre de 2006.

Las ponderaciones de cada artículo representan la relación entre el gasto realizado en las parcelas representadas por dicho artículo y el gasto total realizado en todas las parcelas cubiertas por el índice:

W_i = gasto realizado en las parcelas representadas por el artículo i / gasto total.

Estas ponderaciones son diferentes en cada una de las agregaciones geográficas (provincias, comunidades autónomas y conjunto nacional) y a partir de ellas se obtienen las ponderaciones de las distintas agregaciones funcionales. Así, la ponderación de un agregado funcional se obtiene como suma de las ponderaciones de los artículos que componen dicha agregación. Las actualizaciones anuales de ponderaciones, que se llevan a cabo en el IPC base 2006, se realizan con la última información anual disponible de la nueva ECP.

En líneas generales, los índices agregados se pueden obtener como:

- **Agregaciones funcionales dentro de una provincia:** el índice, referido a diciembre del año anterior, de cualquier agregación funcional en una provincia, se obtiene como agregación de los índices elemen-

tales de los artículos pertenecientes a dicha agregación con las ponderaciones vigentes en el año del índice. Una vez calculados los índices agregados como se ha detallado anteriormente, es preciso encadenarlos. Estos índices son los que finalmente se difunden y dan continuidad a las series publicadas en base 2006.

- **Agregaciones geográficas de una agregación funcional:** De la misma forma que en el caso anterior, se calculan índices de una agregación geográfica superior a la provincia para una agrupación funcional.

Se calculan las siguientes tasas de variación:

- La **tasa de variación mensual** de un índice en un periodo: es el cociente entre el índice del mes corriente y el índice del mes anterior.
- La **tasa de variación acumulada** (o en lo que va de año): es el cociente entre el índice del mes corriente y el índice de diciembre del año anterior.

$$R_i^{mt/(m-1)t} = \frac{I_{dic(t-1),i}^{mt} - I_{dic(t-1),i}^{(m-1)t}}{I_{dic(t-1),G}^{(m-1)t}} \times W_{dic(t-1),i} \times 100, \text{ donde:}$$

$I_{dic(t-1),i}^{mt}$ es el índice, referido a diciembre del año $(t-1)$, del artículo i , en el mes m del año t .

$I_{dic(t-1),G}^{(m-1)t}$ es el índice general, referido a diciembre del año $(t-1)$, en el mes $(m-1)$ del año t .

La repercusión de la variación en lo que va de año (o variación acumulada) de un artículo o conjunto de artículos en el índice general, representa la variación acumulada que experimentarían el índice general si el resto de artículos no hubieran sufrido variación de precios alguna en lo que va de año. O lo que es lo mismo, es la

- La **tasa de variación anual:** cociente entre los índices publicados del mes corriente y del mismo mes del año anterior, ambos en base 2006.

La repercusión de la variación mensual de un artículo o conjunto de artículos en el índice general se define como la parte de la variación mensual del índice general que corresponde a dicho artículo o conjunto de artículos. Por tanto, la suma de las repercusiones mensuales de todos los artículos de la cesta de la compra es igual a la variación mensual del índice general. Esto significa que la repercusión que la variación mensual de precios de un artículo o conjunto de artículos tiene en la variación mensual del índice general, es la variación que éste habría experimentado si todos los precios del resto de artículos hubieran permanecido estables ese mes. La fórmula de la repercusión mensual de un artículo (o agregado funcional) determinado i , en el mes m del año t , es la siguiente:

parte de la variación acumulada debida a dicho artículo o conjunto de artículos. La fórmula de la repercusión acumulada de un artículo (o agregado funcional) determinado i en el mes m del año t es la siguiente:

$$R_i^{mt/dic(t-1)} = \frac{I_{dic(t-1),i}^{mt} - I_{dic(t-1),i}^{dic(t-1)}}{I_{dic(t-1),G}^{dic(t-1)}} \times W_{dic(t-1),i} \times 100, \text{ donde:}$$

$I_{dic(t-1),i}^{mt}$ es el índice, referido a diciembre del año $(t-1)$, del artículo i , en el mes m del año t .

$W_{dic(t-1),i}$ es la ponderación, referida a diciembre del año $(t-1)$, del artículo i , en tanto por uno.

2.3. Cesta de la compra

Las operaciones incluidas en el proceso de cálculo del IPC, desde la recogida de precios hasta el cálculo de los índices, son diferentes en función de las particularidades propias de cada artículo de la cesta de la compra. Así, la periodicidad de la recogida de los precios varía según la frecuencia con la que se modifican los

precios de los artículos. La forma de recogerlos también es diferente dependiendo de la homogeneidad geográfica de los precios y de la disposición de los mismos. Por último, según las características de cada artículo, el método de cálculo de los índices es diferente.

En la siguiente tabla se muestran los diferentes tipos de artículo según los criterios utilizados en su clasificación:

Tabla 2. Criterios de clasificación de los artículos de la cesta de la compra

Criterios	Tipos de artículo			
Periodicidad en la recogida de precios	grupos 1 y 2	mensual	perecedero	estacional
			no perecedero	
	grupos del 3 al 12	mensual		
			trimestral	trimestral
Lugar de recogida y grabación de precios	provincias			
	servicios centrales			
Método de cálculo de índices elementales	estacionales			
	de recogida centralizada			
	de recogida provincial con precio elaborado			
	alquiler de vivienda			
	sin tratamiento especial			

➤ Según la periodicidad y frecuencia en la recogida de precios se establece una primera forma de clasificar los artículos. Así, se consideran dos tipos de artículos, los mensuales y los trimestrales.

- **Artículos de recogida mensual:** Los precios de los artículos mensuales

se observan todos los meses en todos los establecimientos de la muestra, mediante visita personal, en la mayoría de los casos. En general, cada establecimiento se visita una vez al mes, excepto aquellos en donde se recogen precios de artículos perecederos que, debido a la mayor frecuencia

con que varían sus precios, el entrevistador visita dos o tres veces al mes, dependiendo del municipio. En esta categoría se encuentran los alimentos frescos no elaborados, que sufren fluctuaciones periódicas en sus precios y presentan constantes cambios de calidad. Dentro de éstos, se encuentran los artículos estacionales (frutas y hortalizas frescas), algunos de los cuales sólo se comercializan unos determinados meses del año; para estos artículos sólo se lleva a cabo la recogida de precios en los meses en que están disponibles. Dentro de los artículos de recogida mensual destacan los artículos de temporada, que son aquellos cuyo consumo tiene lugar únicamente algunos meses del año, ya que el resto del mismo no se comercializan. En el IPC se consideran de temporada los artículos de vestido y calzado, que tienen 2 temporadas definidas (primavera-verano y otoño-invierno). La recogida de precios de estos artículos se realiza una vez al mes durante la temporada en que se comercializan y el tratamiento de los precios los meses que el artículo desaparece es la repetición del último precio recogido.

- **Artículos de recogida trimestral:** Los artículos trimestrales son aquellos cuyos precios tienen un comportamiento bastante estable, es decir, no suelen experimentar muchas variaciones de precios a lo largo del tiempo (electrodomésticos, muebles, servicios de reparaciones, etc.). La recogida trimestral permite ampliar el número de precios recogidos con el mismo coste. El tratamiento de estos precios consiste en dividir la muestra de establecimientos seleccionados en

tres submuestras, de modo que cada mes sólo se visitan los establecimientos de una de ellas y se repite el último precio recogido en los establecimientos de las otras dos submuestras. Con ello se consigue que todos los meses haya establecimientos que informen sobre los precios de estos artículos. Además, en caso de que varíen más de la mitad de los precios recogidos en el mes, el mes siguiente se solicitará información en todos los establecimientos. Con la inclusión de los precios rebajados, a partir de enero de 2002, se estableció una nueva categoría de artículos dentro de los trimestrales: los trimestrales de rebajas. Son aquellos que, a pesar de cumplir con el requisito de la estabilidad de precios propia de los trimestrales, muestran variaciones significativas en periodos típicos de rebajas (electrodomésticos, muebles, ropa de cama, etc.). Por ello, la recogida de precios en los meses de rebajas se realiza visitando todos los establecimientos de la muestra y no sólo los de la submuestra correspondiente a ese mes.

- Según el lugar donde se recogen y graban los precios, se puede distinguir entre artículos de recogida provincial y artículos de recogida centralizada. Los precios de los primeros se recogen en cada provincia, por medio de visita personal, a través del teléfono o fax, o de los boletines oficiales de las comunidades autónomas o las provincias, y se graban en las delegaciones provinciales.

Por otro lado, el seguimiento de los precios de los artículos de recogida centralizada se realiza desde los Servicios Centrales del INE. Forman parte de este

tipo de artículos aquellos bienes y servicios que tienen una o varias de las siguientes características:

- sus precios son los mismos en una amplia zona geográfica,
- sus precios están sujetos a tarifas publicadas en el BOE,
- existen pocas empresas que comercializan el artículo,
- se dispone de un directorio perfectamente definido de informantes,
- son artículos con cambios de calidad habituales (como los artículos tecnológicos), lo que conlleva dificultad para realizar ajustes de calidad; al hacerse su recogida de forma centralizada se homogeneiza el tratamiento de estos ajustes.

- En lo que respecta al método de cálculo, existen ciertos grupos de artículos cuyos índices elementales se obtienen de forma diferente a la fórmula general descrita anteriormente. Como ya se ha especificado, el índice elemental de cualquier artículo de la cesta de la compra se obtiene como media simple de los precios recogidos, sin tener en cuenta ningún tipo de ponderación. La excepción a esta norma general la constituye la fórmula de cálculo de los artículos de recogida centralizada y los artículos con precio elaborado; en ambos casos, el índice elemental se calcula teniendo en cuenta un conjunto de variedades o modalidades representativas del artículo, ponderadas adecuadamente por el gasto realizado en cada una de ellas.

En el IPC base 2006, la ponderación de cada modalidad se obtiene a partir del gasto realizado por los consumidores. De esta forma, se mantiene la coherencia con la estructura general de ponderaciones de

la cesta de la compra. Otro conjunto de artículos que, por sus características, recibe un tratamiento especial son los estacionales. Debido a las oscilaciones periódicas de sus precios y cantidades, los índices de las frutas, verduras y hortalizas frescas se calculan con un método diferente, basado en medias móviles, que tiene en cuenta los calendarios de producción y comercialización de las mismas. Por último, y debido a las características especiales del mercado, el alquiler de vivienda también recibe un tratamiento diferenciado de la fórmula general del IPC.

- La recogida de precios de los artículos se realiza tanto en provincias como en Servicios Centrales, mediante visita personal de los agentes del INE a los establecimientos en las fechas correspondientes, con la excepción de algunos de ellos para los que, por sus características especiales, la recogida de la información se realiza por teléfono, fax, correo electrónico, catálogo o Internet. La recogida se realiza mediante un cuestionario generado automáticamente para cada establecimiento, en el que el entrevistador anota los precios e incidencias relativos a los artículos que aparecen en el mismo. Cada establecimiento es visitado por un solo entrevistador, excepto los hipermercados y las grandes superficies. Los precios recogidos son precios efectivos de venta al público con pago al contado.

La recogida de precios de un mismo artículo en los distintos establecimientos informantes, se ha distribuido a lo largo de ese periodo para recoger el mayor número posible de fluctuaciones de precios. Todos los meses se visitan los establecimientos seleccionados aproxima-

damente el mismo día; con ello se pretende que la variación reflejada por el índice corresponda a una variación mensual. Como los artículos perecederos están sujetos a fluctuaciones importantes de precios, sus precios se recogen tres veces a lo largo del mes en cada uno de los establecimientos seleccionados en todas las capitales de provincia, manteniendo una distancia de al menos siete días entre las tres visitas al establecimiento. En el resto de municipios, se recogen los precios de estos artículos dos veces en cada uno de los establecimientos que pertenecen a la muestra. Para el resto de artículos cada establecimiento se visita una sola vez al mes, a excepción de los artículos trimestrales, para los cuales la recogida de precios en cada establecimiento se realiza una vez cada tres meses.

Se recogen los precios que han sufrido reducciones por motivos tales como ofertas y promociones, así como aquellos cuyos descuentos son debidos a los periodos oficiales de rebajas. Esto afecta a la mayoría de las parcelas que componen el IPC, aunque los descuentos por rebajas se producen de forma más acentuada en las parcelas de Vestido y calzado y Menaje, donde éstas son más habituales.

Los criterios, seguidos en la base 2006, para la recogida de precios con descuento son los siguientes:

- que el descuento se realice sobre artículos que se espera estén disponibles de nuevo a sus precios habituales, es decir, no se trata de descuentos por liquidaciones o saldos;
- que el descuento se realice sobre artículos que puedan ser adquiridos por todos los consumidores, no sólo por una parte de ellos (por ejemplo, no se tendrán en cuenta descuentos realizados por tener tarjetas de fidelidad del

establecimiento o por cumplir determinadas condiciones);

- y que estos descuentos sean efectivos en el momento de la compra (por ejemplo, no se consideran los reembolsos posteriores a la compra).

Se recogen, por tanto, descuentos debidos a:

- Las rebajas de temporada (periodos de rebajas oficiales regidos por la Ley de Ordenación de Comercio Minorista).
- Ofertas de cualquier tipo (siempre que no se trate de liquidaciones o saldos).

➤ Un aspecto muy relevante en cualquier IPC es el ajuste que se debe realizar sobre los precios cuando hay un cambio en la variedad del artículo o en el establecimiento, ya que el IPC tiene como objetivo recoger la evolución de los precios de los mismos productos a lo largo del tiempo, sin que ésta se vea influida por dichos factores. Estos ajustes se conocen como ajustes por cambio de calidad. Los cambios de calidad son un problema al que se enfrentan todos los países, y que en los últimos años se ha visto acentuado por el rápido progreso técnico que han experimentado algunos artículos. Por ello, es uno de los temas de los que, con mayor prioridad, se ocupa EUROSTAT, en el ámbito de la armonización de los IPC de los países de la UE.

En la elaboración del IPC español, en las distintas bases, han sido varios los procedimientos que se han utilizado para la estimación de los cambios de calidad. La elección de estos métodos ha venido determinada por la disponibilidad de información en cada momento y por el tipo de artículo de que se trate.

Un ajuste por cambio de calidad es necesario cuando un artículo (producto, variedad o modalidad), cuyo precio forma parte del cálculo del IPC, se sustituye por otro, y en ese momento es necesario determinar qué parte de la diferencia de precios entre el artículo sustituto y el sustituido se debe a una calidad diferente entre los mismos.

Las sustituciones de los artículos pueden deberse a varios motivos: cuando deja de ser representativo y surge otro más representativo en el mercado; cuando desaparece del mercado y cuando el establecimiento donde se recoge el precio del artículo deja de ser representativo, cierra o cambia de actividad económica.

En definitiva, los métodos de ajuste de calidad utilizados de forma más habitual en el IPC, base 2006, son los siguientes:

- Ajuste total de calidad. Parte del supuesto de que la diferencia entre el precio del artículo sustituido y del artículo sustituto está totalmente motivada por la diferencia de calidad entre ambos, o que los artículos son tan diferentes que no se pueden comparar. Se considera, entonces, que la diferencia de precios entre ambos artículos es debida únicamente a la distinta calidad de los mismos, con lo que el índice no reflejará variación de precios. Con este ajuste se supone que de haber seguido a la venta el artículo sustituido, su precio no habría variado.
- Ajuste por calidad idéntica. Se parte de la idea de que el artículo sustituto tiene la misma calidad que el artículo sustituido, es decir, que la diferencia de precios existente entre ambos se debe a una variación real de precios. Con este ajuste se supone que de haber seguido a la venta el artículo sustituido, su precio

habría sido el mismo que el del artículo sustituto.

- Otros ajustes. Se incluyen en este apartado todos los ajustes para los cuales se estima el valor de la diferencia de calidad entre un artículo y su sustituto.

Hasta el IPC base 2006, en los cambios de Sistema de Índices de Precios de Consumo se producía una ruptura en la continuidad de las series. La actualización de ponderaciones, la composición de la nueva cesta de la compra y especialmente, los cambios metodológicos, hacían que la serie nueva difiriera de la antigua. Estas diferencias, desde el punto de vista teórico eran insalvables. No obstante, la necesidad de disponer de series continuadas por parte de los usuarios hizo necesario el cálculo de unos coeficientes de enlace que unían las series publicadas en base antigua con las series en base nueva. Sin embargo, para el IPC, base 2006, por tratarse de un índice encadenado, no ha sido necesario calcular ningún coeficiente de enlace, ya que el método de cálculo del encadenamiento permite realizar cambios en ponderaciones, muestra y metodología cada mes de diciembre y encadenar los índices obtenidos con los nuevos cálculos, con la serie que se venía publicando calculada con muestra, ponderaciones y metodología antigua.

Así, en el IPC, base 2006, sólo se ha cambiado el periodo de referencia de los índices o periodo base, que ha pasado de ser el año 2001 a ser el año 2006. Para ello se ha calculado un coeficiente de re-escala, que ha convertido los índices publicados en base 2001, desde enero de 2002 hasta diciembre de 2006, en índices en base 2006.

Este coeficiente es aquel que hace que la media aritmética simple de índices publicados del año 2006, en base 2001, sea igual a 100:

$$\frac{1}{12} \sum_{m=1}^{12} I_{01}^{m06} \times C_{re-escala} = 100 \quad \Rightarrow \quad C_{re-escala} = \frac{100}{\frac{1}{12} \sum_{m=1}^{12} I_{01}^{m06}}$$

Multiplicando la serie publicada en base 2001 por este coeficiente de re-escala, se obtiene una serie de índices en base 2006, que conserva las tasas de variación publicadas, y con la que se han encadenado los nuevos índices en base 2006, calculados a partir de enero de 2007.

2.4. El Índice de Precios de Consumo (Base 2011)

El IPC base 2011 se publica en febrero del año 2011 con el objetivo de mejorar la representatividad de este indicador mediante cambios en la composición de la cesta de la compra y la actualización de la estructura de ponderaciones. Introduce un nuevo tratamiento para los artículos estacionales (frutas frescas y verduras y hortalizas frescas), permitiendo realizar una medición más precisa de la evolución de precios en el corto plazo para este tipo de productos.

2.4.1. Cambios en la cesta de la compra

Los cambios más relevantes en la cesta de la compra están relacionados con los bienes y servicios relativos a los soportes para el registro de imagen y sonido y con el material para el tratamiento de la información. Así, se incorporan los discos duros portátiles y se excluye el CD grabable y el alquiler de película. En lo referente a materiales para el tratamiento de la información la nueva base incluye los *notebooks* y las *tablets*.

También cabe destacar la incorporación en la cesta de la compra de nuevos

servicios de estética, como la fotodepilación y la depilación láser, y paramédicos, como el logopeda.

Como consecuencia de estos ajustes, la cesta de la compra del IPC base 2011 pasa a tener **489 artículos**, frente a los 491 de la base anterior.

2.4.2. Actualización de las ponderaciones

La continua adaptación del IPC a los cambios en el comportamiento de los consumidores incluye también la revisión permanente de su estructura de ponderaciones. Cada año se actualiza el peso o importancia de los grandes agregados que componen este indicador, lo que mantiene la actualidad del mismo.

Además de la revisión anual de las ponderaciones para los grandes agregados, cada cinco años se actualiza la estructura completa para todos los niveles de desagregación. Así pues, el IPC base 2011 incluye una nueva estructura de ponderaciones que representa de forma más precisa las pautas de consumo de los hogares. Para su elaboración se ha considerado la EPF como fuente principal de información, y otras fuentes como la evolución del consumo privado de la Contabilidad Nacional, la evolución de precios del IPC y otras fuentes de diferentes sectores.

En la siguiente tabla se incluye el peso de cada uno de los 12 grandes grupos y su comparación con los pesos vigentes hasta el año 2011.

Tabla 3. Ponderación de los grupos de artículos de la cesta de la compra

Grupo	Ponderaciones de grupos (%)		
	2011	2012	%
01. Alimentos y bebidas no alcohólicas	18,16	18,26	0,6
02. Bebidas alcohólicas y tabaco	2,87	2,89	0,7
03. Vestido y calzado	8,59	8,34	-
04. Vivienda	11,70	00	2,6
05. Menaje	6,84	6,67	-
06. Medicina	3,21	3,14	-
07. Transporte	14,74	15,16	2,9
08. Comunicaciones	3,98	3,85	-
09. Ocio y Cultura	7,64	7,54	-
10. Enseñanza	1,38	1,42	2,8
11. Hoteles, cafés y restaurantes	11,52	11,46	-
12. Otros bienes y servicios	9,37	9,26	-
TOTAL	100	100	1,2

3. LOS CENSOS DEMOGRÁFICOS 2011

Los Censos Demográficos son el proyecto estadístico de mayor envergadura que periódicamente debe acometer la oficina de estadística de cualquier país. Bajo la denominación *Censos Demográficos* se engloban en realidad tres censos diferentes: el Censo de Población, el Censo de Viviendas y el Censo de Edificios. De los tres, el Censo de Población es, sin duda, el de mayor repercusión y el de más amplia tradición.

El primer censo moderno de población, entendiendo como tal el que utiliza a la

persona como unidad de análisis, se realizó en España en 1768 por el Conde de Aranda bajo el reinado de Carlos III. También son de destacar por su interés el Censo efectuado en 1787 por Floridablanca y el realizado diez años más tarde por Godoy en tiempos de Carlos IV. No obstante, la serie de censos de la organización estadística oficial se inicia en 1857 con el primero de la Comisión General de Estadísticas del Reino, al que siguió, en un lapso inusualmente corto, el de 1860. Después vinieron los de 1877, 1887 y 1897. A partir de 1900 ha habido Censo de Población cada diez años sin excepción alguna.

El Censo de Población de 2011 ha sido el decimoséptimo de los Censos oficiales realizados en España. Su realización se enmarca dentro del Programa Mundial 2010 que abarca el periodo 2005-2014 promovido por Naciones Unidas y en cuyo marco se han finalizado, en el momento de redactar este proyecto, los censos de 121 países. Continuando con las actuaciones internacionales que impulsan la realización de los censos, hay que destacar que por primera vez se ha desarrollado una reglamentación comunitaria. El reglamento 763/2008 del Parlamento Europeo y del Consejo (junto con otros que lo desarrollan), además de implantar la obligatoriedad de realizar el Censo durante el año 2011, asegura la comparabilidad de los resultados a nivel de la Unión Europea.

El marco metodológico general en el que se desarrolla el proyecto de censo para España está fijado por las recomendaciones de la Conferencia de Estadísticos Europeos para la ronda censal de 2010 y con un mayor nivel de concreción por el reglamento antes citado y los tres reglamentos de la Comisión que desarrollan el anterior (sobre definiciones de variables y clasificaciones, sobre hipercubos de datos y sobre calidad de la operación).

La Reglamentación desarrollada por la Unión Europea contempla un amplio rango de opciones posibles para recopilar la información de las variables censales. Ese rango va desde los censos clásicos basados en una recogida exhaustiva de los datos, hasta un censo basado en información tomada exclusivamente de registros administrativos. Entre ambos extremos cita un número de situaciones intermedias generadas por el mayor o menor peso de la recogida de datos en campo y de los registros administrativos. Entre ellas figura expresamente el modelo de un censo

basado en registros administrativos completado con una encuesta por muestreo. Este ha sido el modelo censal para España en 2011. De hecho, España, con el Padrón Municipal como registro de población consolidado, se sitúa entre los países con mejores condiciones para realizar un censo de estas características.

La introducción de elementos como la georreferenciación de los edificios, el aprovechamiento de la abundante información administrativa y la recogida de datos multicanal (Internet, entre ellos) son algunos de los ejes sobre los que se construye el primer *censo basado en registros y encuesta por muestreo* en España.

Para su elaboración, se llevó a cabo una recopilación de información procedente de diversas fuentes estadísticas y administrativas que permitieron desarrollar esta estrategia formando un directorio territorial inicial y acumulando datos relativos a personas para su posterior uso. Asimismo, se realizó un recorrido del territorio para completar y contrastar la información territorial disponible, enumerando las viviendas y recogiendo las variables de edificios, disponiendo al final de este recorrido de un directorio completo de la operación.

El Censo de 2011 se planteó como una operación basada en la combinación de los siguientes elementos:

- Un “**fichero precensal**” realizado a partir de un aprovechamiento máximo de los registros administrativos disponibles, tomando al Padrón como elemento básico de su estructura.
- Un trabajo de campo que incluye dos grandes operaciones:
 - Un **Censo de Edificios** exhaustivo que permite la georreferenciación de todos

los edificios y conocer sus características.

- Una **gran encuesta por muestreo**, dirigida a un porcentaje relativamente alto de la población para conocer el resto de características de las personas y las viviendas.

Algunos de los aspectos sustanciales en la estrategia del Censo 2011 han sido:

- La combinación del fichero precensal con la información obtenida de la encuesta proporciona toda la información censal. En particular, la cifra de población se obtiene mediante el recuento de los registros que contenga el fichero precensal, ponderados —cuando sea necesario— con unos factores de recuento obtenidos a partir de la encuesta.
- En tanto que no hay operación de campo exhaustiva, no se prevé que el Censo se utilice para introducir rectificaciones en las inscripciones padronales. El Censo tiene, por tanto, fines exclusivamente estadísticos. Sin embargo sí que se basa en el Padrón, como ya se ha apuntado, y sus resultados se utilizan para contrastar los datos padronales.
- La fracción muestral global está en torno al 12,3% de la población (11,9% de viviendas). Por otro lado, la distribución territorial no es uniforme, al pretender proporcionar un conjunto de tablas e indicadores para los municipios de menor tamaño y, al mismo tiempo, descender del nivel municipal en la difusión de los datos para los municipios que superen un determinado tamaño.
- Se plantea un censo completo de edificios, con especial hincapié en la enumeración de todos los inmuebles de los edificios destinados a vivienda. Análogamente a lo planeado para la población, se realiza una fase de “fichero precensal” de territorio, consistente en un cruce previo entre los datos del Censo

2001, Padrón y Catastro fundamentalmente, completado con datos de otras fuentes como los procedentes de Oficinas de Estadística de las Comunidades Autónomas. Esta fase se complementa con un recorrido para completar el cruce anterior.

- Los datos muestrales se elevan al Fichero Precensal ponderado, calibrados de forma que se reprodujeran las distribuciones marginales del mismo a nivel municipal.
- Las nuevas tecnologías juegan un papel relevante en la operación censal, utilizándose dispositivos portátiles (*tablets*) para el Censo de Edificios y para las entrevistas a la población. Además se ofrecen diversos canales de cumplimentación de la información a los ciudadanos: cuestionarios en papel para ser devueltos por correo e Internet.

3.1. Fichero precensal (FPC)

El fichero precensal se basa en un aprovechamiento máximo de los registros administrativos disponibles, tomando al Padrón como elemento de partida de su estructura al que se va asociando información de otros registros administrativos y de operaciones estadísticas. Esta operación tiene objetivos múltiples:

- Disponer de información adicional a la padronal para decidir si un registro concreto debe ser incluido en el recuento censal o no.
- Aportar información directa de variables censales.
- Servir de marco inicial para realizar una primera selección de la muestra de personas y viviendas que formarán parte de la encuesta.
- Ser el directorio de partida del recorrido del Censo de Edificios.
- Aportar información adicional para las fases de tratamiento de los datos.

El FPC abarca en sentido amplio dos aspectos: territorio y personas. El Fichero territorial intenta reflejar la situación del territorio en el momento más cercano posible a la fecha de referencia censal. Por otra parte, el contenido del fichero de personas se fundamenta sobre la base padronal.

El **Censo de Edificios** es una operación estadística coincidente en el tiempo con la fase postal de la encuesta de población. Dicha operación está diseñada para recopilar información exhaustiva de todos los edificios del territorio nacional en los que hay situada alguna vivienda, con enumeración de todos los inmuebles situados en ellos. Sus objetivos son:

- Enumerar y georreferenciar todos los edificios que tengan algún inmueble que sea una vivienda.
- Determinar las características de los edificios mediante un cuestionario de edificio.
- Enumerar todos los inmuebles contenidos en cada edificio.
- Seleccionar los inmuebles, a medida que se daban altas en el recorrido, que formarán parte de la muestra de la encuesta de población y viviendas.

Al efectuarse con carácter exhaustivo permite disponer de un directorio georreferenciado completo de edificios con alguna vivienda y de todos sus inmuebles. La georreferenciación permite la identificación del edificio ante cambios en la dirección postal y, en consecuencia, su identificación inequívoca.

Como ya se ha apuntado, el **Censo de Población y Viviendas 2011** se basa en tres pilares: el fichero precensal, un Censo de Edificios exhaustivo y una encuesta por muestreo, para conocer las características

de las personas y las viviendas, con un tamaño de muestra adecuado para cumplir con la normativa de cobertura establecida por Eurostat. Sus objetivos son:

- Estimar el total poblacional correspondiente a determinados colectivos, de forma que se corrija la cifra padronal correspondiente a los mismos.
- Estimar las características de la población y de las viviendas a distintos niveles de desagregación geográfica.

De acuerdo con lo anterior, y considerando que el censo de población es el único medio que se dispone para obtener información desagregada a nivel de sección censal, unidad primaria de muestreo utilizada en las encuestas de hogares, se selecciona una muestra en todas las secciones censales.

3.2. Selección de la muestra

El marco para la selección de la muestra ha sido el FPC obtenido del cruce de los ficheros de Padrón, Catastro y otros ficheros de tipo administrativo. Después de los distintos cruces realizados, las viviendas figuran clasificadas en el FPC como localizables o no localizables. Las primeras son aquellas que a través de la dirección postal es posible que se puedan localizar en el trabajo de campo. Las segundas son aquellas que no tienen una dirección completa y consecuentemente no pueden ser localizadas en el trabajo de campo. Durante la realización del trabajo de campo se dieron altas y bajas en las viviendas del FPC.

A efectos de selección de la muestra, el total de viviendas de cada municipio se agrupa en dos marcos: Marco A formado por el conjunto de viviendas localizables del FPC tal y como se ha definido anteriormente, y Marco B formado por el

conjunto de inmuebles que son dados de alta durante el recorrido exhaustivo que se realiza en campo.

Para alcanzar los objetivos del censo se selecciona una muestra de inmuebles diferente según se trate del marco A o B. La muestra se selecciona en todas las secciones censales.

- MARCO A: La muestra procedente de este marco selecciona de entre los inmuebles cuyo uso fuera el de vivienda, antes de comenzar los trabajos de campo. De este grupo se contacta por correo ordinario con las que según el FPC sean principales, dando la opción de cumplimentar el cuestionario censal bien por Internet o bien devolverlo cumplimentado por correo ordinario. Pasado un periodo de tiempo, del conjunto residual de viviendas que no habían respondido se seleccionó una submuestra para recoger el cuestionario censal mediante entrevista personal asistida por ordenador (CAPI). Las viviendas no principales se investigaron durante la realización del recorrido exhaustivo y aquellas que no pudieron ser resueltas durante el mismo pasaron a la muestra de recogida por CAPI. En consecuencia, se consideran dos grupos diferentes de viviendas, uno conteniendo a las principales y el otro a las no principales.
- MARCO B: La muestra procedente de este marco se selecciona durante el recorrido exhaustivo en campo, para lo cual se utiliza el procedimiento de Bernouilli de selección aleatoria. Durante dicho recorrido, a esta parte de la muestra se le deja un cuestionario en papel y el hogar pudo colaborar de la misma forma a la indicada para las viviendas del marco A.

Teniendo en cuenta los objetivos de obtener una cierta información a nivel municipal y el presupuesto disponible, el tamaño de la muestra es de, aproximadamente, tres millones de viviendas, representando una fracción de muestreo global del 11,9% por ciento. En porcentaje de población representa una fracción de muestreo del 12,3%. La distribución de esta muestra depende del tamaño del municipio, fijando unos criterios de precisión para las estimaciones. De esta forma las fracciones de muestreo varían desde las aplicadas a los municipios de menor tamaño que eran investigados exhaustivamente hasta las de los municipios de mayor tamaño a los que corresponden las menores fracciones de muestreo.

La muestra se selecciona en cada municipio, dentro de cada grupo de viviendas, con igual probabilidad. La procedente del marco A, se selecciona mediante muestreo sistemático con arranque aleatorio. La muestra del marco B se selecciona del conjunto de inmuebles que son alta en el recorrido exhaustivo utilizando el procedimiento de Bernouilli que asigna igual probabilidad a las unidades de muestreo.

Los estimadores de las características de viviendas y personas, en un determinado municipio, son estimadores de expansión con corrección de falta de respuesta a los que se aplican técnicas de calibrado, según los casos, y el factor de expansión inicial se obtiene como la inversa de la probabilidad de selección.

BIBLIOGRAFÍA

Carrasco Carpio, C. y García Serrano, C. (2012). Inmigración y mercado de trabajo. *Informe 2011*. Ministerio de Trabajo y Seguridad Social.

Censos de Población y Viviendas 2011. *Estadística y Sociedad* (2011), nº 48.

Delgado, M. (2011). El próximo censo de población de 2011. *Estadística y Sociedad*, 48.

Escuder Bueno, J. y Escuder Vallés, R. (2009). Síntesis histórica y metodológica de los Índices de Precios al Consumo españoles. *Estadística Española*, 51(171), 257-280.

Fuentes Castro, D. (2007). Análisis del poder adquisitivo de los asalariados desde la entrada en circulación del euro. *Boletín Económico del ICE* (Información Comercial Española), nº 2926, 39-64.

García, M.A. (2005). Cambios en la encuesta de población activa en 2005. *Estadística y Sociedad*, 11.

García Villar, J. y Gómez del Moral, M. (2010). La estadística oficial como bien público en una sociedad democrática. *Estadística y Sociedad*, 43, 5-7.

Hansen, M.H., Hurwitz, W.N. y Bershad, M.A. (1961). Measurement errors in censuses and surveys. *Bulletin of the International Statistical Institute*, 38, 359-374.

www.ine.es/censos2011/censos2011.htm

www.ine.es/daco/daco43/notaepa.htm

www.ine.es/daco/daco43/meto_res_ipc.htm