

Tema 6: Proyecciones de población. Procedimientos matemáticos de estimación. El método de los componentes. Proyección de componentes. Estimadores intercensales de población.

6.1. PROYECCIONES DE POBLACIÓN

Tanto los agentes públicos como los privados necesitan estimaciones y proyecciones de población para su actividad planificadora: desde la elaboración de políticas educativas, sociales, sanitarias, de pensiones, de urbanismo, etc., hasta la planificación de comportamientos de consumo. También son útiles, por ejemplo, para evaluar los efectos que tendría una determinada política de control o fomento de la natalidad.

El tamaño y la composición de la población varían en función de los cambios que se produzcan en la mortalidad, la natalidad y los movimientos migratorios, lo que obliga a establecer hipótesis sobre estos tres elementos de forma combinada. Esto hace que la tarea de estimar la población de un área en un determinado momento sea realmente complicada.

Esta tarea es razonablemente más sencilla cuanto mayor sea la población a estimar (por ejemplo, es más sencillo estimar la población del total nacional que de las provincias). No obstante, dado que el menos previsible de los componentes es el migratorio, incluso a nivel nacional se pueden producir importantes desviaciones en las estimaciones, pues es fácil, además, que su magnitud sea muy superior a la del crecimiento vegetativo.

Es importante distinguir entre dos conceptos que, a primera vista, podrían parecer equivalentes, pues utilizan el mismo método de cálculo, y sin embargo tienen distinta filosofía. Se trata de las previsiones y las proyecciones demográficas.

El objetivo de las **previsiones** es “acertar” la evolución futura de la población, incorporando tendencias probables basadas en el pasado para proporcionar escenarios futuros altamente probables. Pero esto es muy complejo y subjetivo, pues depende de un conjunto mucho más amplio de parámetros (económicos, sociales, etc...), a menudo de difícil cuantificación.

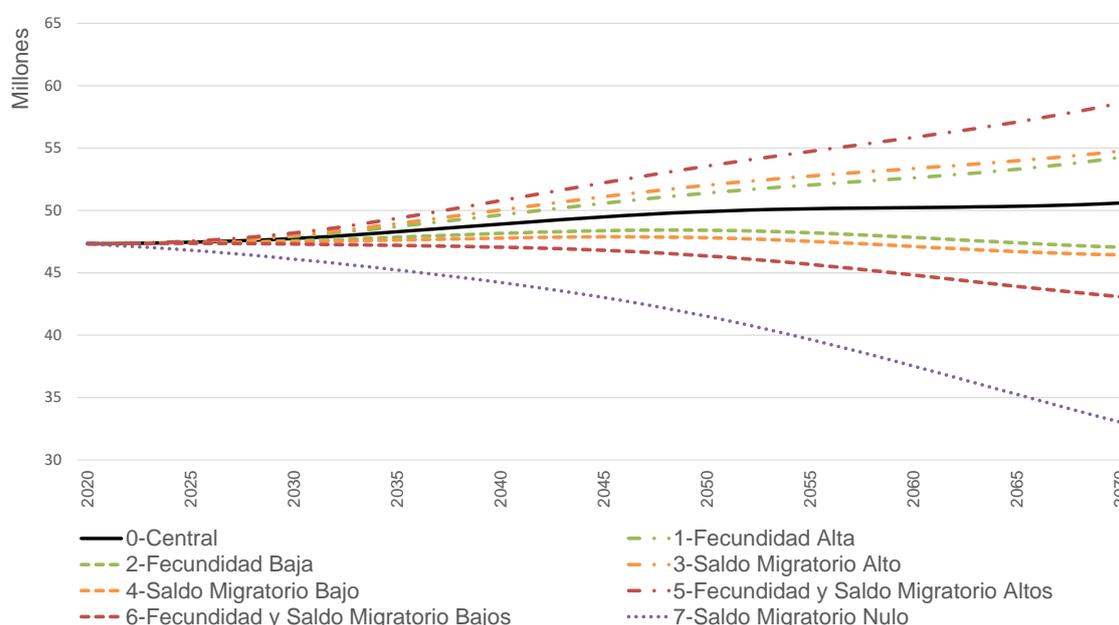
El objetivo de las **proyecciones** demográficas es proporcionar una simulación de la evolución futura de la población bajo ciertos supuestos, aunque estos no sean muy verosímiles. En muchas ocasiones sirven simplemente para comprender las consecuencias que tendría sobre la población el establecimiento de determinadas hipótesis.

Ante estas dificultades, en ocasiones se plantean varios conjuntos de hipótesis que dan lugar a distintos **escenarios** de evolución futura, basados en la combinación de diferentes supuestos sobre los fenómenos demográficos. Se pueden plantear un sinnúmero

de escenarios, pero lo más habitual es realizar unas proyecciones para un **escenario medio o central**, junto con distintos escenarios que varían entre un **escenario alto** y un **escenario bajo**, y que recogerían conjuntos de hipótesis más o menos favorables para el crecimiento de la población.

Las proyecciones de población del Instituto Nacional de Estadística no pretenden ser previsiones demográficas, sino proporcionar una herramienta de apoyo para la toma de decisiones. Consisten en una simulación estadística del rumbo demográfico que tomaría la población residente en España en los próximos años, bajo la hipótesis de que se mantuviesen las tendencias demográficas actuales. En las Proyecciones de Población 2020-2070 del INE, se establecieron los escenarios que se pueden ver en el siguiente gráfico.

Gráfico 6.1. Proyecciones de Población 2020-2070. Escenarios



Fuente: Instituto Nacional de Estadística

En el ámbito de las proyecciones de población se pueden definir una serie de conceptos:

- **Población base, inicial o de partida:** es la población, conocida, a partir de la cual se realiza el cálculo para proyectar la población en otros momentos.
- **Horizonte de proyección:** es el momento temporal más lejano en el tiempo para el cual se estima la población hipotética.
- **Periodo de proyección o proyectivo:** se denomina así al intervalo de años que se proyecta, desde el momento de la población base hasta el horizonte de proyección.

En función del momento en que se sitúe el horizonte de proyección, se distinguirán varios tipos de proyecciones:

- **Proyecciones a corto plazo:** cuando se contempla un horizonte de proyección hacia delante no superior a 5 años.
- **Proyecciones a medio plazo:** cuando se establece un horizonte proyectivo hacia delante de entre 5 y 20 años.
- **Proyecciones a largo plazo:** cuando el horizonte se aleja hacia delante más de 20 años.
- **Retroproyecciones:** en este caso el horizonte proyectivo se sitúa en el pasado, para estimar la población de años anteriores.

Por último, se definen las:

- **Proyecciones derivadas:** se realizan proyecciones derivadas cuando, dada una proyección de población, se realizan estimaciones adicionales de ciertas características, añadiendo variables a la proyección, como es el caso de las proyecciones de hogares.

El Instituto Nacional de Estadística publica proyecciones de población actualizadas cada dos años, con un horizonte proyectivo de 50 años para el total nacional y de 15 años para las comunidades autónomas y las provincias.

Un aspecto básico a la hora de realizar **proyecciones para diferentes niveles de desagregación territorial** es el de elegir una estrategia general para la proyección de los diferentes ámbitos. Los distintos modelos a la hora de realizar una proyección regional pueden ser:

- **Modelos desagregativos o top-down:** en estos modelos se parte de la realización de una proyección del área superior (por ejemplo la nacional), a la que posteriormente se le aplican unos coeficientes para distribuir la población proyectada entre las diferentes áreas inferiores en jerarquía (regiones, provincias...) contenidas en el área superior. Sus ventajas son la sencillez y coherencia entre ámbitos territoriales. Por el contrario, no considera los posibles diferenciales territoriales en los comportamientos demográficos ni la influencia que las diversas estructuras poblacionales tienen sobre ellos.
- **Modelos agregativos o bottom-up:** en este enfoque se proyecta cada una de las áreas inferiores independientemente para obtener, posteriormente, el ámbito superior por agregación. Su principal ventaja es que tienen en cuenta las especificidades territoriales, mientras que tienen varias desventajas: requieren formular un gran número de hipótesis y no se tiene control previo sobre los resultados de la región superior, pudiendo surgir incoherencias. Estas incoherencias provienen, fundamentalmente, de la no asignación de los movimientos migratorios entre las áreas menores.

- **Modelos mixtos:** combinan las ventajas del enfoque top-down (consistencia) con las características específicas de modelo bottom-up. Para cada región, el número de salidas es calculado independientemente, y el total de migrantes son obtenidos por suma dando lugar al *migrant pool* o *caja negra*. Los emigrantes en el *pool* son posteriormente distribuidos como inmigrantes entre las diferentes áreas.
- **Modelo multirregional** (Willekens, Drewe): surge para paliar las limitaciones de los anteriores modelos. Este modelo presenta la ventaja de considerar las especificidades demográficas de cada una de las regiones, al tiempo que se mantiene la coherencia con el nivel superior. Las interrelaciones entre las regiones vienen determinadas por los intercambios migratorios, que tienen en cuenta tanto la región de origen como la de destino. La especificidad de cada región viene reflejada por unos patrones de mortalidad, fecundidad y migraciones propios.

Así pues, los modelos multirregionales consisten en realizar una proyección superior e inferiores simultáneas, con un pequeño ajuste posterior para que cuadren entre sí. Esto asegura la consistencia entre los distintos niveles territoriales, así como entre stocks y flujos demográficos. Como inconvenientes se puede señalar la gran cantidad de requisitos de información y el mayor número de parámetros necesarios para proyectar. El Instituto Nacional de Estadística utiliza este método para elaborar sus proyecciones de población.

Existen diversas técnicas para estimar una población en un momento dado. Principalmente se engloban en dos tipos de métodos: **procedimientos matemáticos de estimación** y **método de los componentes**.

6.2. PROCEDIMIENTOS MATEMÁTICOS DE ESTIMACIÓN

Los primeros trabajos de proyección demográfica se basaron en el ajuste estadístico de diversas funciones matemáticas a las nubes de puntos que representaban la evolución temporal de una determinada población. Una vez determinado el modelo más adecuado, se estimaban las poblaciones futuras para momentos posteriores.

Así pues, los **procedimientos matemáticos o de extrapolación** fueron los más empleados en el pasado, por su simplicidad. Consideraban que el crecimiento poblacional presenta un patrón regular y, por tanto, que se puede expresar mediante un modelo matemático basado en las tasas anuales de crecimiento. Indirectamente, supone que las condiciones que afectan a dicho crecimiento se van a mantener constantes en el futuro, o variarán de forma muy gradual.

Su principal desventaja radica en que consideran que el crecimiento de la población es independiente de la evolución de cada uno de los fenómenos demográficos (nacimientos, defunciones y migraciones), despreciando esta información específica que afecta a la dinámica de la población. Además, en estos modelos, la proyección se realiza para la población total, sin desagregarla según la edad.

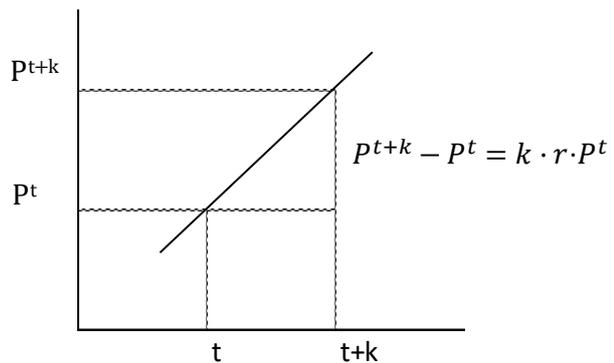
A continuación se presentan los principales modelos de extrapolación.

Modelo aritmético

Asume un comportamiento lineal de la población, es decir, que la población crece la misma cantidad cada año (también pueden ser décadas u otra unidad de tiempo). Por este motivo, es el modelo más simple de todos.

Si r es la tasa de crecimiento anual respecto de la población base P^t en el año t , entonces, cada año el crecimiento es de $r \cdot P^t$. De este modo, para un periodo genérico de k años, la fórmula para la población proyectada en el año $t+k$ será:

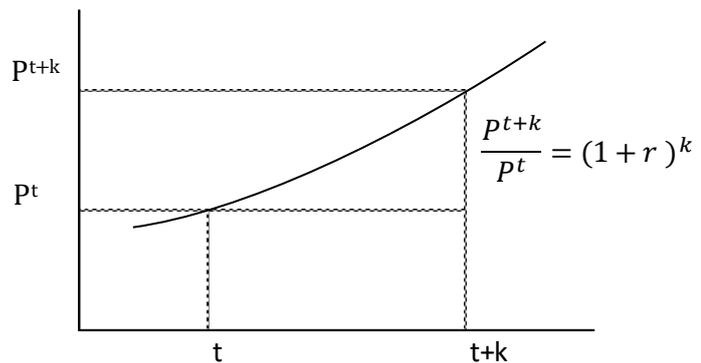
$$P^{t+k} = P^t \cdot (1 + k \cdot r)$$



Modelo geométrico

En el modelo geométrico se supone que la población se incrementa en la misma proporción cada año, a razón de una tasa de crecimiento constante r respecto de la población anterior. Este modelo viene representado por la siguiente fórmula:

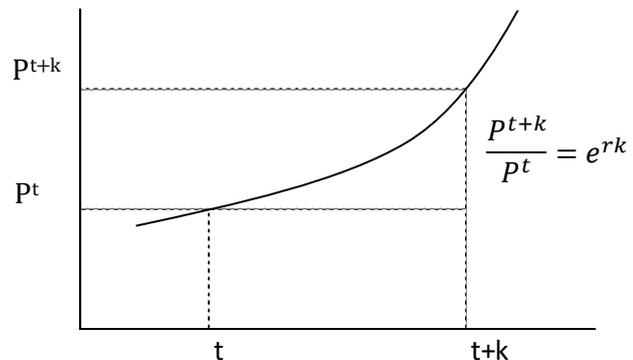
$$P^{t+k} = P^t \cdot (1 + r)^k$$



Modelo exponencial

La curva exponencial muestra una tasa de crecimiento constante pero continuo, a diferencia del modelo geométrico, que lo hace para cada unidad de tiempo. Gráficamente está representado por una línea curva hacia arriba que no tiene límite superior. Así, la población en $t+k$ vendría dada por:

$$p^{t+k} = p^t \cdot e^{r \cdot k}$$



Modelo logístico

Una de las posibles desventajas del modelo exponencial es que el crecimiento se puede disparar rápidamente. Para poner solución a este problema, surgió el modelo logístico, que introduce límites al crecimiento a través de dos asíntotas, superior e inferior. Este modelo supone, pues, que el crecimiento es lento al principio y al final, y rápido en el periodo central, siguiendo gráficamente la forma de una S. Se expresa matemáticamente como sigue:

$$p^t = \frac{P^\omega}{1 + e^{-r(t-t_0)}}$$

donde:

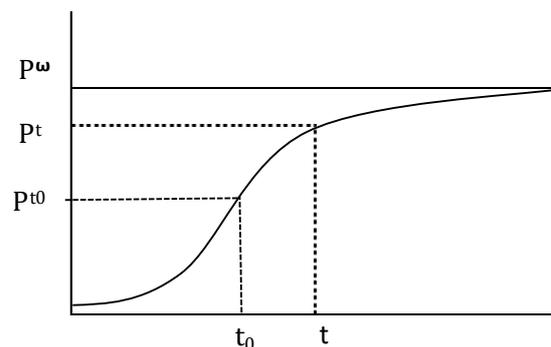
p^t es la población calculada para el año t

P^ω es el valor máximo alcanzable

r es la tasa de crecimiento anual

t_0 es el momento en que la población alcanza la mitad del valor máximo alcanzable:

$$p^{t_0} = \frac{P^\omega}{2}$$



6.3 EL MÉTODO DE LOS COMPONENTES

A diferencia de los anteriores, el método de los componentes tiene en cuenta los distintos fenómenos demográficos en el crecimiento poblacional, por lo que es el más empleado. La aplicación de este método responde al siguiente esquema: partiendo de la población residente en un cierto ámbito geográfico y de los datos observados para cada uno de los componentes demográficos básicos - la mortalidad, la fecundidad y la migración - se trata de obtener la población correspondiente a fechas posteriores bajo ciertas hipótesis sobre la evolución de esos tres fenómenos, que son los que determinan su crecimiento y su estructura por edades. Esto se realiza por medio de la **ecuación compensadora**:

$$P^{t+k} = P^t + N^{t,t+k} - D^{t,t+k} + SM^{t,t+k}$$

donde:

P^{t+k} es la población estimada k años después del año base t

P^t es la población en el año base t

$N^{t,t+k}$ es el número de nacimientos entre el año t y el año t+k

$D^{t,t+k}$ es el número de defunciones entre el año t y el año t+k

$SM^{t,t+k} = I^{t,t+k} - E^{t,t+k}$ es el saldo migratorio entre el año t y el año t+k, es decir, la diferencia entre las inmigraciones, $I^{t,t+k}$, y las emigraciones, $E^{t,t+k}$, de dicho periodo.

Por la naturaleza aritmética de este método, es posible descomponer la población en distintas subpoblaciones, como regiones, cohortes... y aplicarlo a cada una de ellas de forma independiente. De modo que, para una subpoblación dada i, su población en un momento t+k vendrá dado por:

$$P_i^{t+k} = P_i^t + N_i^{t,t+k} - D_i^{t,t+k} + SM_i^{t,t+k}$$

Además, si las subpoblaciones consideradas son áreas inferiores al total nacional (como las provincias), los movimientos migratorios se pueden descomponer en migraciones exteriores y migraciones interiores. El interés en separarlas radica en que suelen presentar comportamientos diferentes, de modo que las podemos tratar como dos fenómenos demográficos independientes para ayudar a establecer las hipótesis. Se reflejaría mediante la expresión:

$$P_i^{t+k} = P_i^t + N_i^{t,t+k} - D_i^{t,t+k} + SMext_i^{t,t+k} + SMint_i^{t,t+k}$$

donde:

$SMext_i^{t,t+k}$ es el saldo migratorio exterior de la subpoblación i con el extranjero.

$SMint_i^{t,t+k}$ es el saldo migratorio interior de la subpoblación i con el resto de regiones del país.

De este modo, la estimación de la población total en el horizonte proyectivo se obtendrá como la suma de todas las subpoblaciones que, suponiendo que sean n , vendrá dada por la expresión:

$$P^{t+k} = \sum_{i=1}^n P_i^{t+k}$$

Basta con observar la fórmula de la ecuación compensadora para darse cuenta de que la clave para la correcta aplicación del método de los componentes radica en un correcto planteamiento de las hipótesis sobre el comportamiento esperado de cada uno de los componentes demográficos a lo largo del periodo proyectivo.

Las **proyecciones de población que elabora el INE** se basan en el método de los componentes. El análisis retrospectivo de cada uno de los fenómenos demográficos básicos, haciendo uso de la información demográfica más actualizada disponible, ha permitido establecer hipótesis sobre la incidencia futura de los mismos en cada nivel territorial considerado para cada año del periodo proyectivo. Esta incidencia se cuantifica a través de tasas de fecundidad por generación, tasas de mortalidad por sexo y generación, tasas por sexo y generación de emigración, así como flujos de inmigración exterior para cada sexo y generación. Además, desde 2018, se han establecido hipótesis diferenciadas por lugar de nacimiento (España o extranjero), por tener ambos colectivos comportamientos y dinámicas demográficas distintas.

Por otro lado, la proyección de la población de cada sexo, edad y lugar de nacimiento residente en España y en cada una de sus comunidades autónomas y provincias, a 1 de enero de cada año del periodo proyectivo, se ha llevado a cabo de acuerdo a un modelo de proyección multirregional¹. De este modo, proporciona como resultados no solo las cifras de población, sino también las cifras proyectadas de nacimientos, defunciones y movimientos migratorios que tendrían lugar en cada uno de los años del periodo proyectivo, guardando todo ello la necesaria coherencia entre flujos y stocks demográficos y la debida consistencia interterritorial.

Para derivar la fórmula de proyección empleada, debemos tener en cuenta las definiciones de las distintas tasas por sexo s y edad x (a 1 de enero²). Si estimamos la

¹Willekens, F.J, "Demographic forecasting: state of the art and research needs", en Emerging Issues in Demographic Research, (ed) Hazeu y Frinking (1990), y Willekens, F.J. y Drewe, P. , "A multiregional model for regional demographic projection", en Heide, H. y Willekens, F.J. Demographic Research and Spatial Policy, (ed) Academic Press, Londres (1984).

² Nótese que a 1 de enero existe una relación biunívoca entre edad y año de nacimiento o generación, por lo que será equivalente hablar de generaciones o cohortes y edad a 1 de enero. En general, hablar en

población media del denominador como el promedio de las poblaciones a principio y final del periodo t, es decir, a 1 de enero de los años t y t+1 ($P_{s,x}^t$ y $P_{s,x+1}^{t+1}$), podemos expresar los distintos fenómenos en función de sus tasas, como sigue:

- $D_{s,x}^t = m_{s,x}^t \left(\frac{P_{s,x}^t + P_{s,x+1}^{t+1}}{2} \right)$ son las defunciones en el año t de edad x (a 1 de enero) y sexo s, donde $m_{s,x}^t$ son las tasas de mortalidad durante el año t para dicho sexo y edad.
- $N_x^t = f_x^t \left(\frac{P_{muj,x}^t + P_{muj,x+1}^{t+1}}{2} \right)$ son los nacimientos en el año t por edad de la madre x (a 1 de enero), donde f_x^t son las tasas de fecundidad durante el año t para las mujeres de dicha edad y $P_{muj,x}^t$ es la población de mujeres de edad x a 1 de enero del año t.
- $E_{s,x}^t = e_{s,x}^t \left(\frac{P_{s,x}^t + P_{s,x+1}^{t+1}}{2} \right)$ son las emigraciones en el año t de edad x (a 1 de enero) y sexo s, donde $e_{s,x}^t$ son las tasas de emigración durante el año t para dicho sexo y edad.

Sustituyendo estas expresiones en la ecuación compensadora y despejando la población proyectada, nos queda la siguiente expresión equivalente para proyectar la población un año después:

$$P_{s,x+1}^{t+1} = \frac{[1 - 0,5(m_{s,x}^t + e_{s,x}^t)] \cdot P_{s,x}^t + I_{s,x}^t}{[1 + 0,5(m_{s,x}^t + e_{s,x}^t)]}$$

donde $I_{s,x}^t$ es el flujo de inmigración procedente del extranjero en el año t de individuos de sexo s y edad x a 1 de enero del año t.

Para obtener la población proyectada para la última edad considerada, que en el INE es el grupo abierto de 100 y más años, y se denota por $P_{s,100+}^{t+1}$, se ha de sustituir en todos los componentes de la fórmula anterior el valor x por 99+, es decir, por 99 y más años.

El caso de la edad 0 también es especial, y se debe emplear la fórmula:

$$P_{s,0}^{t+1} = \frac{[1 - 0,5(m_{s,-1}^t + e_{s,-1}^t)] \cdot N_s^t + I_{s,-1}^t}{[1 + 0,5(m_{s,-1}^t + e_{s,-1}^t)]}$$

donde $m_{s,-1}^t$ es la tasa de mortalidad de la generación de individuos residentes en España, de sexo s, nacidos durante el año t; $e_{s,-1}^t$ la tasa de emigración al exterior de los individuos residentes en España, de sexo s, nacidos durante el año t; $I_{s,-1}^t$ es el flujo de inmigración procedente del extranjero de nacidos de sexo s durante el año t; y N_s^t son los nacidos en España de sexo s durante el año t, los cuales se derivan de la expresión siguiente:

términos de edad a 1 de enero es más práctico, por tener un rango de valores fijos para cualquiera que sea el año de referencia.

$$N_s^t = r \cdot \sum_{x=14}^{49} \frac{P_{muj,x}^t + P_{muj,x+1}^{t+1}}{2} \cdot f_x^t$$

siendo r el ratio de masculinidad al nacimiento proyectado en el caso de los varones; $P_{muj,x}^t$ la población de mujeres de edad x a 1 de enero del año t ; y f_x^t la tasa de fecundidad de la generación de mujeres residentes en España durante el año t que tienen edad x a 1 de enero de dicho año.

Para proyectar la población de varios años consecutivos se realiza este procedimiento de forma iterativa, tantas veces como años se deseen proyectar. La población proyectada de un año será la población de partida para la proyección del año siguiente.

Por último, una vez que tenemos la serie de poblaciones proyectadas, se pueden obtener los fenómenos demográficos proyectados asociados (nacimientos, defunciones y migraciones) a partir de las fórmulas que relacionan los fenómenos con sus tasas, expuestas previamente.

Por otro lado, el INE realiza proyecciones de población a nivel provincial (y de comunidades autónomas por agregación) para el horizonte proyectivo de 15 años. Como se emplea un modelo multirregional, han de establecerse hipótesis sobre la migración interior desde una provincia a otra, en forma de tasas de migración interior según provincia de origen y destino. En este caso basta introducir el componente de emigración interior, pues la inmigración interior es el fenómeno opuesto y se puede deducir de la primera. Esto da lugar a una fórmula de proyección que incorpora gran complejidad, pues para cada provincia la fórmula depende también del resto de provincias, es decir, están interconectadas, por lo que se deben proyectar de forma simultánea.

En este caso, además, se proyecta de forma independiente la población del total nacional. El cálculo de la proyección conlleva, por tanto, un proceso iterativo de comprobación de consistencia y ajuste de los resultados provinciales a los nacionales. Este proceso consiste en introducir sucesivos factores de corrección provinciales que modifican muy ligeramente, en el mismo grado para todas las provincias, las tasas de fecundidad, mortalidad y emigración al extranjero por generación (y por tanto sin modificar la posición relativa de cada provincia respecto a las demás en cuanto a la incidencia de cada fenómeno demográfico), hasta conseguir la completa consistencia interterritorial de stocks poblacionales y eventos demográficos proyectados.

6.4 PROYECCIÓN DE COMPONENTES

Para establecer hipótesis sobre la mortalidad, la fecundidad y las migraciones, lo más adecuado es dividir la población en subpoblaciones independientes con comportamientos diferenciados entre sí, y lo más homogéneos posible dentro de ellas. Dado que estos componentes varían, como mínimo, en función del sexo y la edad, este

será el nivel mínimo de desagregación para establecer dichas subpoblaciones a la hora de realizar la proyección. Cuantas más subpoblaciones se puedan proyectar independientemente, más precisa y correcta será la proyección.

Por este motivo, el primer paso para realizar la proyección será disponer de una **población base o de partida** desagregada al menos por sexo y edad (o grupos de edad). Además, dado que la generación o año de nacimiento tiene una correspondencia biunívoca con la edad a 1 de enero, interesa tratar con poblaciones a esta fecha.

A continuación, se debe realizar un estudio retrospectivo a partir del cual se elaborarán las hipótesis sobre la evolución de los fenómenos. En particular, será necesario expresar matemáticamente su **intensidad y calendario**. Si no se dispone de datos, es posible utilizar los de otras poblaciones con comportamientos similares.

Por último, a partir de la población de partida, **se proyecta cada subpoblación** de acuerdo a las hipótesis establecidas sobre los fenómenos. Se repite el proceso las veces que sean necesarias hasta alcanzar el horizonte proyectivo (la población proyectada es la población de partida de la siguiente iteración).

Existen muchas formas de establecer hipótesis para proyectar los componentes, desde cálculos sencillos hasta procedimientos tan complicados como se desee, pero hay que tener presente que un procedimiento complejo no tiene por qué ser necesariamente mejor. Es más importante conocer bien el comportamiento del componente que se desea proyectar, que la complejidad matemática del cálculo.

A continuación se presentan algunos modelos matemáticos que se pueden aplicar para la proyección de cada uno de los componentes. Generalmente se pueden plantear dos tipos de hipótesis: de **estabilidad o mantenimiento** y de **variabilidad**.

PROYECCIÓN DE LA MORTALIDAD

La mortalidad muestra bastante inercia en sus tendencias en ausencia de catástrofes naturales o guerras, por lo que es el componente que presenta menos problemas a la hora de establecer hipótesis. Además, su incidencia sobre las poblaciones calculadas es menor que la que produce la fecundidad o las migraciones, habiéndose comprobado que la estructura de la población puede influir más intensamente en las defunciones totales que las diferencias en las hipótesis de mortalidad.

Por este motivo, el envejecimiento de las poblaciones de los países más avanzados como consecuencia de un incremento en la esperanza de vida está dando lugar a un mayor interés en los demógrafos. En poblaciones en las que los ancianos son poco numerosos, las diferencias de intensidad de la mortalidad tienen poca repercusión, mientras que en poblaciones con mayor número de personas en edad avanzada, las hipótesis sobre mortalidad cobran mayor importancia.

Para proyectar la mortalidad, se comienza analizando la evolución de las tasas específicas de mortalidad por sexo y edad en años anteriores. A partir de ahí, y analizando las perspectivas futuras de mortalidad, se proyecta el comportamiento futuro de esas tasas, teniendo en cuenta también que pueden llegar a estabilizarse en algún momento. En general se pueden establecer dos tipos de hipótesis:

- **De mantenimiento:** se mantienen constantes a lo largo de la proyección los indicadores de intensidad de la mortalidad elegidos, ya sea de la última tabla de mortalidad conocida o de la tabla de mortalidad de otra población conocida, escogida por adecuarse al nivel que supuestamente tendrá en el futuro (esto ocasionaría una discontinuidad, por lo que este método sería cuestionable).
- **De variabilidad:** basándose en el análisis retrospectivo, se hace variar la mortalidad a lo largo del periodo proyectivo. Por ejemplo:
 - o Proyectar cada una de las tasas específicas por edad y sexo. Se puede emplear algún modelo matemático simple, como la extrapolación lineal o exponencial.
 - o Proyectar la esperanza de vida y, a partir de ahí, derivar las tasas de mortalidad. Para ello es necesario establecer un umbral máximo de longevidad humana, basado en propuestas de diversos autores especializados o en los niveles máximos de las tablas-tipo, para situarlo bastante después del horizonte proyectivo (por ejemplo, 100 años). También es necesario un umbral mínimo, que puede fijarse por tanteo.

En las **Proyecciones de Población 2020-2070 del INE**, aun cuando las hipótesis de proyección se incorporan separadas por el lugar de nacimiento en todos los fenómenos, en el caso de la mortalidad se emplean los mismos parámetros de proyección para los nacidos en España y para los nacidos en el extranjero. Esto es debido al bajo número de defunciones registrado entre la población nacida en el extranjero residente en España, que conlleva un escaso o nulo número de defunciones al desagregarlas por sexo, edad y provincia.

La metodología de proyección de la mortalidad en España se lleva a cabo a partir de una proyección basada en el nivel general, sintetizado por la esperanza de vida al nacimiento, y se derivan posteriormente tablas de mortalidad acordes con esos valores, mediante el uso de tablas tipo. A partir de esas tablas se derivarán las tasas de mortalidad por sexo y generación necesarias para la proyección. Se desarrolla en los siguientes pasos:

En primer lugar, se proyecta la esperanza de vida al nacimiento para cada uno de los años del periodo proyectivo, mediante una regresión lineal de una función logística frente al tiempo o año de calendario, hasta un máximo que se alcanzaría en un futuro teórico fijado en el infinito, esto es, muy alejado del año horizonte. Para ello, se utiliza la función logit recomendada por el Banco Mundial:

$$\text{Logit}(e_0^t) = \left(\frac{e_0^{\max} - e_0^t}{e_0^t - e_0^{\min}} \right)$$

Para determinar el valor máximo para la esperanza de vida, e_0^{\max} , en cada edición de las proyecciones (se renuevan cada dos años) se realiza una encuesta a expertos en demografía de toda España con el fin de recabar su opinión sobre la evolución futura esperable de los parámetros de los distintos fenómenos demográficos necesarios para la proyección. En particular, en lo relativo a la mortalidad se les pregunta qué valor consideran ellos que alcanzaría la esperanza de vida al nacimiento para los hombres y las mujeres residentes en España, separadamente, dentro de 50 años. Así, el valor máximo de la esperanza de vida al nacimiento, se elige como aquel valor que permite que en el último año del periodo proyectivo a largo plazo considerado, la esperanza de vida al nacimiento sea igual a la media aritmética de las respuestas dadas por los expertos.

El valor mínimo de la esperanza de vida al nacimiento, e_0^{\min} , necesario para introducir en la función logit, será el que proporcione el mejor ajuste al ser asociado con el valor máximo que se considere como límite.

Con los logits correspondientes a los años observados desde 1991, se calculan los elementos de la recta de regresión sobre el tiempo, y con ellos se calculan los logits correspondientes para toda la serie proyectada. Estos logits se transforman en esperanzas de vida deshaciendo la fórmula anterior.

Dado que es conocido que la esperanza de vida al nacimiento es mayor en mujeres que en hombres, se hace una distribución progresiva, en 20 años para las mujeres, y en 40 años para los hombres, de la diferencia obtenida entre la esperanza de vida al nacimiento observada y estimada para el último año observado.

Por último, se proyecta la tabla de mortalidad de cada uno de los años del periodo proyectivo, adecuada al nivel de esperanza de vida al nacimiento proyectado a partir de la regresión logit, mediante el uso de tablas tipo de mortalidad.

La proyección de la incidencia de la mortalidad en las provincias se desarrolla a partir de una metodología relacional siguiendo el método de los *logits de Brass*³, que relaciona la incidencia de la mortalidad de cada provincia con la del total nacional.

PROYECCIÓN DE LA FECUNDIDAD

Los cambios socioeconómicos y demográficos pueden producir cambios de intensidad y calendario en la fecundidad de una población.

De forma análoga a la proyección de la mortalidad, el primer paso sería realizar un análisis retrospectivo del comportamiento de las tasas específicas de fecundidad en

³ William Brass, (1975), *Methods for estimating fertility and mortality from limited and defective data*.

años anteriores y del Indicador Coyuntural (o Sintético) de Fecundidad (ICF). A la hora de establecer hipótesis para el futuro, se supone que el ICF llegará a estabilizarse en cierto momento, debiéndose fijar cierto valor hacia el cual hacerlo tender. Puede ser tentador elegir este nivel como el de reemplazo, que asegura la reproducción de las generaciones femeninas. Sin embargo, en poblaciones como la española, que están muy por debajo de este valor, se obtendrían resultados ilógicos, al menos en un plazo breve tras el inicio de la proyección.

Un método sencillo consiste en prolongar razonablemente la evolución de las tasas específicas por edad de la fecundidad, de posibles dos formas:

- Proyectar cada una de las tasas de fecundidad. Mediante este método no se tiene control sobre el nivel global de fecundidad, sintetizado por la suma de las tasas, el ICF, pues se obtiene de forma derivada.
- Proyectar el ICF y derivar las tasas a posteriori. A través de este método se consigue el control sobre la proyección del nivel global de fecundidad, y a la vez proporciona la flexibilidad suficiente para establecer su calendario adaptado a las hipótesis que hayamos planteado.

Asimismo, al igual que en el caso de la mortalidad, es necesario establecer un umbral, alejado del horizonte de proyección, hacia el cual tenderán nuestros patrones de fecundidad proyectados. Para ello, se deben fijar el ICF, la Edad Media a la Maternidad y su varianza, parámetros necesarios para aplicar después alguna función matemática que nos permita derivar la serie de tasas de fecundidad por edad que se corresponden con los parámetros establecidos.

Por último, habría que obtener las tasas para los años intermedios. Para ello se puede aplicar algún tipo de interpolación, o bien se puede utilizar el conocimiento del fenómeno para establecer el ritmo del cambio y los momentos en que, de producirse, se invierten las tendencias, así como los valores mínimos y máximos que alcanzarán las tasas proyectadas.

Hay otras alternativas para la proyección de la fecundidad, desde las más simples, como mantener el último nivel conocido, hasta las más complejas, como proyectar por separado la fecundidad por estado civil de la madre o por rango de nacimiento.

En las **Proyecciones de Población 2020-2070 del INE**, el método general de proyección de la fecundidad de las mujeres residentes en el territorio español se basa en establecer hipótesis sobre la evolución futura del nivel general de fecundidad, sintetizado por el Indicador Coyuntural de Fecundidad (ICF) o número medio de hijos por mujer, y de los parámetros que describen su distribución por edad: la Edad Media a la Maternidad (EMM) y la Varianza de la Edad Media a la Maternidad (Var(EMM)).

Estas hipótesis se plantean distinguiendo el lugar de nacimiento de la madre, para tener en cuenta así el distinto comportamiento de las mujeres nacidas en España y de las nacidas en el extranjero.

Las curvas de fecundidad por edad observadas en años anteriores son continuas, acampanadas, unimodales y ligeramente desviadas a la derecha, por lo que una de las distribuciones de probabilidad que mejor puede ajustar el calendario de la fecundidad observada en España es la distribución de probabilidad Beta, corregida por los límites de edad inferior y superior. Así pues, las tasas de fecundidad por edad para cada año del periodo proyectivo se obtienen mediante el ajuste de una distribución de probabilidad Beta de parámetros ICF, EMM y Var(EMM). Por este motivo, necesitamos proyectar dichos parámetros para cada uno de los años del periodo proyectivo.

Las tasas por edad así obtenidas se transforman posteriormente en tasas por generación, a partir de la semisuma de las tasas correspondientes a las dos edades consecutivas que intervienen en cada generación, bajo la hipótesis de distribución uniforme de los cumpleaños a lo largo del año.

Para determinar el ICF y la EMM se utilizan los datos de la encuesta que se realiza a expertos en demografía mencionada anteriormente. En particular, en lo relativo a la fecundidad se les pregunta el número medio de hijos por mujer y la edad media a la maternidad esperados dentro de 15 y 50 años, distinguiendo entre mujeres nacidas en España y en el extranjero. Para los años intermedios se aplica una interpolación lineal entre el último valor observado (en el caso de las proyecciones 2020-2070 son los datos provisionales de 2019) y la media aritmética de los valores dados por los expertos en la encuesta.

En cuanto a la varianza de la edad media a la maternidad de cada uno de los años del periodo proyectivo se deja constante e igual al valor del último año observado, que en la última edición correspondieron a los datos provisionales de 2019.

La proyección de la evolución de la fecundidad en cada una de las provincias de España se lleva a cabo a partir de una simulación del comportamiento diferencial de la fecundidad en cada provincia respecto al total nacional, para lo que se emplea el *modelo de Brass-Gompertz Relacional*, siguiendo la metodología propuesta por Zeng y otros (2001)⁴.

PROYECCIÓN DE LA MIGRACIÓN

La proyección de las migraciones es la más complicada de todos los componentes, pues es el más volátil y difícil de prever. Esta dificultad viene motivada por la propia naturaleza del fenómeno y sus múltiples causas, así como la falta de disponibilidad de series de datos suficientemente largas. Por este motivo, a veces se opta por realizar proyecciones en ausencia de migraciones, es decir, considerando poblaciones cerradas, o bien se plantean hipótesis sencillas como suponer una migración constante o decreciente en su magnitud, basada en lo observado en el pasado reciente.

⁴ Zeng Yi, Wang Zhenglian, Ma Zhongdong y Chen Chunjun. 2000. "A simple method for projecting or estimating α and β : An extension of the Brass Relational Gompertz Fertility Model", Population Research and Policy Review 19:525–549.

Al igual que en el resto de componentes, hay múltiples métodos para proyectar las migraciones. Una opción adecuada consiste en considerar por separado la emigración y la inmigración, por tratarse de fenómenos con comportamientos diferenciados, aunque en ocasiones no hay más remedio que proyectar directamente el saldo migratorio si no se dispone de información diferenciada.

Para la emigración se deben establecer unas tasas específicas de emigración por generación, cuya suma proporciona el Índice Sintético de Emigración (ISE). Para realizar la proyección se fija un valor del ISE para cada periodo de la proyección, de acuerdo con las hipótesis establecidas sobre la evolución futura de las migraciones. A continuación se le aplicaría un calendario para desagregarlo en las tasas correspondientes. Para establecer un calendario coherente con cada intensidad migratoria existen modelos teóricos como el de Rogers y Castro (1982), o también se puede optar por una estructura conocida, como, por ejemplo, el calendario emigratorio observado en los periodos recientes.

Para la inmigración no podemos establecer una tasa de primera categoría análoga a la de emigración, por resultar inviable (habría que poner en el denominador la población del resto del mundo). Por este motivo, lo que se hace es, en primer lugar, proyectar el nivel global de la inmigración medido en forma de flujos, es decir, en valor absoluto. Para ello se podría emplear, por ejemplo, un método de regresión a partir de datos retrospectivos observados. En segundo lugar, se deben desagregar estos flujos por edad y sexo aplicando una estructura, que podría ser una estructura conocida a partir de los datos de periodos recientes, o bien de alguna población similar.

En las **Proyecciones de Población 2020-2070 del INE**, se proyecta por separado la inmigración y la emigración exterior y, a su vez, cada uno de los fenómenos se ha dividido en nacidos en España y en el extranjero, como así lo aconseja el hecho de tratarse de migraciones de naturaleza y dinámica muy distinta.

La **inmigración exterior** se introduce en el proceso de proyección a través de los flujos proyectados para cada año del periodo proyectivo por sexo, generación, lugar de nacimiento y, para los 15 primeros años, también por provincia.

En primer lugar, se establece una intensidad de inmigración global para cada año del periodo proyectivo, distinguiendo entre nacidos en España y en el extranjero. Estos flujos se reparten a su vez por sexo, generación y, en su caso, por provincia, aplicando distribuciones promedio obtenidas a partir de las correspondientes a los últimos cinco años observados en la Estadística de Migraciones, publicada por el INE, para así evitar la variabilidad propia de un mayor nivel de detalle para un conjunto de datos menor. Estas distribuciones proyectadas se mantienen constantes a lo largo de la proyección.

La proyección de los flujos totales de inmigración de nacidos en España y fuera de España para la última edición de las Proyecciones de Población se obtiene diferenciando tres tramos del periodo proyectivo.

- Tramo 1: en general, está formado por una estimación now-cast o del momento para el primer año de la proyección, partiendo de la información parcial disponible en el momento de elaboración de las proyecciones, seguido de algunos años en que se recoge la tendencia observada en los últimos años. No obstante, en las últimas Proyecciones de Población, afectadas en su año de partida por la incidencia del COVID-19, este tramo únicamente constó de dos años. Para el segundo año se hizo una estimación específica que no recogía la tendencia.
- Tramo 2: interpolación lineal entre el valor obtenido para el último año del tramo 1 y el valor asignado para dentro de 15 años según la encuesta a los expertos, obtenido como la media aritmética de sus respuestas.
- Tramo 3: interpolación lineal entre los valores asignados para el último año del tramo 2 y el valor asignado para dentro de 50 años según la encuesta a los expertos, obtenidos como la media aritmética de sus respuestas.

En el proceso de proyección de la población, la **emigración exterior** entra en forma de tasas de emigración por generación para cada sexo y lugar de nacimiento, y para cada año del periodo proyectivo. Además, para los años en que la población se proyecta a nivel provincial, también son necesarias las tasas específicas de cada provincia.

De forma análoga a la inmigración exterior, para su proyección se establece una intensidad de emigración global para cada año del periodo proyectivo, sintetizada en este caso por el Índice Sintético de Emigración (ISE), distinguiendo entre nacidos en España y en el extranjero. Estos valores se distribuyen a su vez por sexo, generación y, en su caso, por provincia, a través de calendarios por edad y diferenciales por sexo y provincia promedio, obtenidos a partir de los correspondientes a los últimos cinco años observados de la Estadística de Migraciones. Estas distribuciones proyectadas se mantienen constantes a lo largo de la proyección.

Sin embargo, dado que el planteamiento de la intensidad emigratoria vía índices sintéticos (suma de tasas) es poco intuitiva, se ha establecido dicha hipótesis mediante flujos, de forma similar a la inmigración, para así también poder incluir las opiniones de los expertos recabadas por la encuesta. De este modo, para proyectar los flujos de emigración de nacidos en España y en el extranjero, se ha dividido el periodo proyectivo en los mismos tramos que en la proyección de la inmigración exterior, y se ha aplicado el mismo método para obtener los flujos proyectados para cada año.

Después, los flujos de emigración al extranjero y el calendario de emigración proyectados son sometidos a un proceso iterativo en la ejecución del ejercicio proyectivo a nivel nacional que permite derivar, a partir de una solución de partida, un ISE para cada año del periodo proyectivo coherente con los flujos y calendarios proyectados. A partir de esos ISE proyectados y de los calendarios y diferenciales proyectados, se obtienen las tasas de emigración necesarias para el cálculo de la proyección.

Como el INE elabora Proyecciones de Población a nivel de provincia para los primeros 15 años, es necesario también establecer hipótesis sobre la **migración interprovincial** que reflejen los intercambios proyectados de población entre provincias.

Para proyectar la población a nivel provincial es necesario proporcionar las tasas de migración interior por generación para cada sexo, lugar de nacimiento, provincia de origen y provincia de destino para cada año del periodo proyectivo. Para ello se procede de forma similar al resto de fenómenos: proyectando en primer lugar una medida global del fenómeno, en este caso el Índice Sintético de Emigración interprovincial (ISEint) de cada provincia al resto de España, y a su vez unas distribuciones para desagregar por el resto de variables: diferenciales por sexo, calendarios por generación y coeficientes de reparto de la emigración según provincia de destino.

Dado que la incidencia de la migración interior es bastante estable, sus tasas se proyectan constantes para todo el periodo, proyectando, también de forma constante, cada uno de sus componentes. La intensidad de la migración interior (ISEint) se obtiene a partir de los datos de migración interior de los últimos tres años de la Estadística de Migraciones, tratando de recoger la tendencia más reciente, pero a la vez buscando evitar la variabilidad en los datos propia de las áreas geográficas más pequeñas. Para el resto de componentes, cuya estabilidad es mayor, se emplean datos de los últimos cinco años de la Estadística de Migraciones.

6.5 ESTIMACIONES INTERCENSALES DE POBLACIÓN

El Censo de Población es la operación destinada a medir, cada diez años, la población residente en España y en cada uno de sus municipios, de acuerdo a las definiciones internacionalmente reconocidas.

Los efectivos de población por sexo y edad resultantes en cada censo se emplean como población de partida para el establecimiento de proyecciones o estimaciones de población, por el método de componentes, en la mayoría de los países del contexto occidental. A su vez, las poblaciones estimadas o proyectadas pueden compararse a posteriori con las censales en la fecha de referencia correspondiente.

Las diferencias entre las poblaciones estimadas y las observadas producen saltos o rupturas en las series de población, haciendo necesario enlazar o corregir las estimaciones de población realizadas de acuerdo con los resultados de los sucesivos censos, es decir, obligan a realizar **Estimaciones Intercensales de Población**.

Las diferencias que habitualmente se encuentran entre las poblaciones calculadas y las observadas tienen su origen, bien en errores en las poblaciones censales de partida y/o de llegada, o bien en las hipótesis realizadas respecto a todos o alguno de los componentes proyectados (mortalidad, fecundidad y migración). Pero dada la imposibilidad de distinguir la parte de error que proviene de diferencias en el subregistro o la sobreenumeración en dos censos sucesivos, de la parte debida a la estimación incorrecta de los componentes, el procedimiento general consiste en hacer la hipótesis

de que el error final obtenido se distribuye de forma lineal a lo largo del período intercensal.

Con ese objetivo, se han ido calculando las Estimaciones Intercensales de Población en España, de acuerdo con los sucesivos Censos de 1970, 1981, 1991, 2001 y 2011, haciendo uso de los datos sobre nacimientos, defunciones y migraciones disponibles para los correspondientes periodos intercensales.

Las series de estimaciones intercensales de población alcanzan hasta el 1 de enero de 2012 y enlazan tanto con los censos intermedios (1981, 1991, 2001 y 2011) como con las Cifras de Población que se publican a partir de 2013 para el periodo posterior al Censo de 2011. Todo ello constituye, junto a los censos anteriores a 1970, la serie estadística histórica de Cifras de Población de España.

Existen métodos de interpolación que solo son útiles para determinar las poblaciones intermedias totales, como la interpolación geométrica, que supone una tasa de crecimiento anual constante, mostrando como clara desventaja el no contemplar la dinámica de la población dada por los fenómenos demográficos ocurridos en cada año, además de no permitir la desagregación de subpoblaciones.

Sin embargo, lo que más interesa es obtener una serie de población desagregada al menos por las variables demográficas básicas, como son el sexo y la edad, y que además refleje el ritmo de crecimiento derivado de los fenómenos demográficos ocurridos en cada año del periodo. Para este objetivo, el INE emplea el método de los componentes.

Las cifras de población residente en España, en cada comunidad autónoma y en cada provincia (e isla, en las provincias insulares) según sexo, año de nacimiento, país de nacionalidad y país de nacimiento en las fechas de referencia censales proporcionadas por los respectivos Censos de Población, se toman como puntos de referencia.

La metodología general de cálculo de las Estimaciones Intercensales de Población 1971-2011 del INE consiste, básicamente, en:

- 1- Llevar a cabo una contabilidad de los sucesos demográficos acontecidos (nacimientos, defunciones, migraciones y adquisiciones de nacionalidad española) durante el periodo intercensal de manera que, partiendo de las cifras de población residente en la fecha de referencia del Censo de Población inicial, se obtenga como resultado una cifra de población residente en la fecha de referencia del siguiente Censo y en todos los años intermedios.
- 2- Ajuste de los resultados del paso anterior, mediante un reparto uniforme a lo largo de cada año del periodo considerado, de la diferencia observada entre el Censo de Población final y la población estimada para su misma fecha de referencia.

Por otro lado, la contabilización de sucesos demográficos se lleva a cabo a partir de las siguientes fuentes de información:

- **Nacimientos:** resultados definitivos de las Estadísticas del Movimiento Natural de la Población (MNP) de todos los años desde 1971.
- **Defunciones:** resultados definitivos de las Estadísticas del Movimiento Natural de la Población (MNP) de todos los años desde 1971.
- **Migraciones:** desde 2001 se utilizan datos elaborados a partir de la información registrada en el Padrón Municipal, sometida a un tratamiento estadístico. Los datos obtenidos para el periodo 2008-2011 son coincidentes con los resultados de la Estadística de Migraciones. Para los años anteriores a 2001 no se han contabilizado las migraciones, pues para ese periodo de tiempo no existen fuentes de información fiables.
- **Adquisiciones de nacionalidad española:** al proporcionarse datos por nacionalidad, es necesario tener en cuenta también las nacionalizaciones (y en teoría también las pérdidas de nacionalidad española, pero son despreciables). Desde 2001 hasta 2011, se aproximan a partir de los datos sobre concesiones de nacionalidad española otorgadas por el Ministerio de Justicia. Para los años anteriores a 2001 no se ha realizado una contabilización de las adquisiciones de nacionalidad, puesto que para ese periodo de tiempo no existen fuentes de información fiables.