

Generación de una proyección de la población española para el período 1996-2025, mediante un modelo de simulación estocástica

CASAS SÁNCHEZ, J.M., GUTIÉRREZ DE MESA, J.L. y NÚÑEZ VELÁZQUEZ, J.J.
Departamento de Estadística, Estructura Económica y O.E.I. Facultad de CC. Económicas y Empresariales. Universidad de Alcalá de Henares.

Facultad de CC. Empresariales. Antiguo Colegio de Mínimos. Plz de la Victoria, 2. 28802-Alcalá de H. Madrid. España. Tlf: 91 885 42 76. Fax: 91 885 42 01. E-mails: jmiguel.casas@yah.es; joseluis.gutierrez@uah.es; josej.nunez@uah.es

RESUMEN

Se trata de elaborar una proyección de la población española, mediante un modelo de simulación estocástica iterativo, a través de la selección de intervalos de crecimiento quinquenales para la esperanza de vida al nacer de las mujeres, así como para el Índice Sintético de Fecundidad (ISF) y el número de inmigrantes.

Palabras clave: Demografía, Simulación Estocástica, Proyección de Poblaciones.

ABSTRACT

The aim is the construction of a projection of the Spanish population, through a stochastic iterative simulation model, with the selection of growth intervals five-year for the life hope upon born of the women, as well as for the Total Fertility Rate (TFR) and the number of immigrants.

Key words: Demography, Stochastic Simulation, Populations forecasts.

Clasificación UNESCO: 5207.09, 5207.03, 1209.10

Clasificación J.E.L.: J11, C15.

1. INTRODUCCIÓN

La aproximación dominante entre las proyecciones de población a largo plazo la constituye el denominado “método de las componentes”, desarrollado por Whelpton y sus colaboradores, aunque el planteamiento inicial se remonta a finales del siglo XIX (Whelpton,

1947)¹. Este método combina distintas hipótesis sobre mortalidad, fecundidad y migrabilidad de la población en el periodo de proyección. No obstante, este método es determinista y obtiene un único resultado en función de las hipótesis adoptadas para las tasas vitales, por lo que los organismos y agencias que elaboran proyecciones pueden proponer hipótesis, más o menos optimistas sobre el comportamiento de las tasas vitales, dando lugar generalmente a tres escenarios, para los que se obtienen otras tantas proyecciones, que son interpretadas como un análogo de intervalo de confianza para la proyección media obtenida.

La versión estocástica del método de los componentes considera las tasas vitales como realizaciones de procesos estocásticos del tipo:

$$V(t+1) = R(t,t+1).V(t) + N(t+1), t = 1, 2, \dots$$

donde $V(t)$ es un vector que representa la población ordenada por edad en el instante t y $R(t,t+1)$ es una matriz que contiene las tasas vitales, ordenadas de manera adecuada, entre los años t y $t+1$ y, habitualmente, se le conoce como matriz de Leslie generalizada (Alho, 1990). A esta formulación se añaden los movimientos migratorios $N(t+1)$ (Cohen, 1986, Keyfitz, 1977). Sin embargo, estos modelos generan problemas relacionados con las correlaciones de las diferentes componentes y su medición a través de los datos históricos, así como con las autocorrelaciones de las propias tasas vitales (Alho, 1990). Por estas razones, los métodos de simulación se constituyen como una alternativa válida (Pflaumer, 1988), en la que se basa el modelo que se presenta en este trabajo.

Otra alternativa la constituyen los modelos de series temporales. Así, el uso de modelos ARIMA sobre series simples (Land y Cantor, 1983) no utiliza información de otras series de tiempo para construir modelos explicativos de la serie estudiada, de manera que las fluctuaciones estacionales solo pueden modelizarse a través de sus datos presentes y pasados. Así, esta aproximación puede considerarse como un paso preliminar en la construcción de modelos explicativos de series temporales estacionales bivariantes y multivariantes, de mucha mayor complejidad (Land, 1986).

La alternativa de la utilización de modelos paramétricos está relacionada con el uso de modelos socio-económicos que, típicamente, se incorporan como componente demográfica de modelos estructurales más amplios (Ahlburg, 1987). Entre ellos, pueden reseñarse los modelos Brookings y Wharton (Land, 1986, p. 897).

Los modelos de simulación estocástica ofrecen una alternativa a los modelos antes mencionados, estableciendo bandas de fluctuación sobre las predicciones del crecimiento demográfico mediante la consideración de las tasas vitales como variables aleatorias

1. Su planteamiento formal puede verse con detalle en Pflaumer, P. (1988), por ejemplo.

(Pflaumer, 1988). La simulación de la realidad permite que se puedan crear los escenarios más diversos y, mediante unas hipótesis previas, más o menos contrastadas con la realidad, comprobar cuál habría sido el resultado, sin necesidad de imponer un modelo paramétrico subyacente.

Los modelos de simulación demográfica se pueden clasificar en macrosimulaciones y microsimulaciones (Van Imhoff y Post, 1997). Las primeras suponen que algunas variables demográficas son aleatorias y tratan de determinar la población en un instante futuro, a base de generar itinerarios aleatorios para predecir cuál sería el número más probable de efectivos de población, de acuerdo con las hipótesis adoptadas y el método de proyección utilizado. La microsimulación, en lugar de generar realizaciones concretas de las variables demográficas, establece un itinerario para cada uno de los elementos de esa población aplicando, una a una, las diversas probabilidades de muerte, fecundación y migración, para determinar cuántos individuos llegan al final de la proyección. Existen diversos modelos de microsimulación como el DYNASIM (Orcutt et al. 1996), el modelo Frankfurt (Krupp et al. 1981), el DPMS (Van Imhoff y Post, 1997), el NEDYMAS (Nelissen, 1994) y MOSART (Brunborg y Keilman, 1995), en los que, tras numerosas encuestas y teniendo en cuenta situaciones socioeconómicas de los individuos, se les aplican distintas probabilidades de muerte, fecundidad y migrabilidad.

En este trabajo se presenta un modelo de simulación en el que se van a considerar las variables demográficas básicas como variables aleatorias. Así, en el punto 2 se presentará el modelo de simulación propuesto para pasar, posteriormente, a su aplicación al caso español y al análisis de los resultados.

2. MODELO DE SIMULACIÓN ESTOCÁSTICA

Se trata de desarrollar un modelo de proyección basado en el método de las componentes que, tomando valores de las tres variables demográficas básicas asociadas a los fenómenos de mortalidad, fecundidad y migración, proyecte la población a una fecha posterior de modo que se repita este proceso un número N suficientemente grande de veces, para determinar su comportamiento medio, así como los percentiles 5 y 95, con el fin de obtener una banda central donde se encuentre aproximadamente el 90% de los resultados.

2.1. Las variables del modelo

Para poder realizar una proyección basada en el método de las componentes se necesita conocer, en primer lugar, la población inicial o de partida, que se introducirá en el modelo como un parámetro conocido. Se necesitan, además, las tasas específicas de fecundidad por

edad de las mujeres en cada subperíodo de proyección², así como las probabilidades perspectivas de paso por edad o por grupos de edad, dependiendo del tipo de tabla de mortalidad disponible. También se precisan las tasas específicas de emigración y el total de inmigrantes esperados por edad y por subperíodo de proyección, para ajustar las migraciones.

Ahora bien, para simplificar el número de variables que actúan en el modelo, se utilizarán únicamente tres, una por cada uno de los fenómenos demográficos. Por tanto, se van a considerar las variables aleatorias:

- a) El “*Índice Sintético de Fecundidad*” (ISF), que es el número de hijos por mujer. Esto supone realizaciones aleatorias de las tasas de fecundidad.
- b) La “*Esperanza de Vida al Nacer de las Mujeres*” (e_0). Esto supone realizaciones aleatorias de las probabilidades perspectivas de paso, tanto de hombres como de mujeres de una población, puesto que las probabilidades de muerte de mujeres y varones están relacionadas a través de las medidas de sobremortalidad masculina.
- c) El “*Saldo Migratorio*” (SM), por subperíodo de proyección, para poder determinar sus posibles defunciones y nacimientos. Se considera que los inmigrantes adquieren automáticamente las mismas condiciones de fecundidad y mortalidad de los autóctonos.

El modelo se basa en el método de Montecarlo para generar números aleatorios que den lugar a las realizaciones de las tres variables aleatorias descritas.

Para ello, en el caso de la variable aleatoria “*ISF*”, conocida cada realización concreta para un subperíodo de proyección, se procede a obtener las diversas tasas específicas de fecundidad por edad f_k . Para ello, si se denomina X a la variable aleatoria “*Edad de la madre al nacimiento de un hijo*” y suponiendo que los grupos de edad son de amplitud h , siendo l la edad mínima y L la máxima, respectivamente, a la que las mujeres pueden tener hijos, se define la función de cuantía:

$$p_k = P[l + (k - 1)h < X \leq l + k.h] \quad \forall k = 1, \dots, (1/h).(L - l)$$

Por tanto, las tasas específicas por edad serán:

$$f_k = \text{ISF} \cdot p_k \quad \forall k = 1, \dots, (1/h).(L - l),$$

donde se asume que esta estructura de fecundidad por edades se mantiene a lo largo de todo el periodo de proyección, y se ha determinado previamente a través de los últimos datos conocidos.

2. El subperíodo de proyección será de un año o de cinco, dependiendo de que las tablas de mortalidad utilizadas sean completas o abreviadas.

Análogamente, en el caso de la variable aleatoria “*Esperanza de Vida Femenina al nacimiento*”, se obtendrá, en cada realización, un valor de la esperanza de vida de las mujeres para el subperíodo de proyección, y a partir de ella habría que determinar las distintas probabilidades de paso de hombres y mujeres. Este problema se puede resolver, bien empleando tablas proyectadas de esa población, o bien, utilizando tablas-tipo de mortalidad³. Para valores intermedios de las esperanzas de vida, se interpolan las tasas específicas de mortalidad asociadas a cada tabla, tal y como se recomienda en las proyecciones efectuadas por el Instituto de Demografía del CSIC y la Consejería de Economía de la Comunidad de Madrid en sus últimas proyecciones⁴. El problema de calcular las probabilidades perspectivas de paso de los varones, también queda solventado ya que estas tablas-tipo van ligadas, para cada nivel, una para hombres y otra para mujeres, para que no se produzcan diferencias anormales entre la esperanza de vida de las mujeres y de los varones.

Para el caso de los movimientos migratorios, se utiliza la variable aleatoria “Saldo Migratorio”. En cualquier caso, el modelo necesita conocer las proporciones por sexo y edad con el fin de conocer su número para cualquier realización de la variable aleatoria. Esta configuración se considera, obviamente, constante a lo largo de todo el periodo de proyección⁵.

2.2. Descripción del modelo.

El esquema de funcionamiento del modelo está representado en el Anexo 1. Consta de una serie de módulos que se describen a continuación.

El Módulo 1 sirve para introducir los valores iniciales del modelo, que son los parámetros externos, como la población inicial, el Índice Sintético de Fecundidad y la Esperanza de Vida al Nacer de las Mujeres en el instante anterior al momento inicial de la proyección. Otros parámetros que intervienen son N (número de iteraciones que se quieren realizar) y n (número de subperíodos necesarios para llegar al instante final), así como otros parámetros a , b , α , a' , b' , α' , a'' y b'' , que se describen a continuación y que sirven para crear diversos escenarios de proyección.

En el Módulo 2 se calculan e_0 , ISF y SM para cada subperíodo de proyección. El proceso consistiría, en principio, en tomar realizaciones aleatorias de estas variables en un intervalo fijado de antemano, pero íntimamente relacionado con las realizaciones del subperíodo anterior, o periodo inicial, en su caso. En concreto:

3. Las más utilizadas son las obtenidas por Coale, A. y Guo, G. (1989 y 1991).

4. Instituto de Demografía del CSIC (1994). Comunidad Autónoma de Madrid (1994).

5. Una posible extensión de este planteamiento requiere conocer suficientes datos históricos sobre esta variable, para establecer su patrón de evolución. Sin embargo, no existen fuentes de datos al respecto.

En el caso de la esperanza de vida al nacer de las mujeres, en lugar de elegir un valor e_0 entre dos valores extremos posibles, $e_0^{\text{máx}}$ y $e_0^{\text{mín}}$, se asumirá que la variación de e_0 de un periodo hasta el siguiente está acotada por una cierta cantidad x_t , para garantizar un comportamiento moderado y acorde con las variaciones observadas en periodos anteriores. Para ello, se eligen dos números positivos a y b , siendo $a < b$, que van a proporcionar el intervalo en el que se va a mover la variabilidad efectiva del periodo t (x_t) que, por tanto, será un número elegido al azar entre a y b . Con el fin de establecer una proporción entre las posibilidades de disminuir y de crecer, se elige, previamente, un número α , tal que $0 < \alpha < 1$, que indique la proporción de variabilidad total con la que e_0 podría decrecer. O sea, si se denota e_0^{t-1} al valor obtenido de la esperanza de vida en el periodo anterior, el valor que se obtendrá en el periodo siguiente, será cualquier número perteneciente al intervalo:

$$(e_0^{t-1} - \alpha \cdot x_t, e_0^{t-1} + (1-\alpha) \cdot x_t)$$

Análogamente, se elige un intervalo (a' , b') que contenga a la máxima variabilidad del ISF y una cantidad α' que indique la proporción de disminución. El nuevo ISF del periodo siguiente será un número seleccionado al azar del intervalo:

$$(\text{ISF}^{t-1} - \alpha' \cdot y_t, \text{ISF}^{t-1} + (1-\alpha') \cdot y_t)$$

donde y_t es la variabilidad total del periodo, obtenido al azar de entre a' y b' .

En el caso de las migraciones, no se hace ninguna hipótesis sobre su número y crecimiento. Por ello, se selecciona un número aleatorio entre el mínimo y el máximo de los valores que se consideran posibles, para cada subperíodo de proyección. La razón estriba en que el número de inmigrantes se comporta de modo más errático y no necesariamente sigue la pauta de periodos anteriores.

Obviamente, las descripciones anteriores implican la asunción de distribuciones de tipo uniforme para las variabilidades de ISF y e_0 , así como para el Saldo Migratorio. Podría plantearse la posibilidad de imponer otros modelos distribucionales alternativos, obtenidos a partir de las distribuciones empíricas correspondientes, pero esto requeriría series históricas suficientemente amplias, por una parte, y suponer que el comportamiento futuro será similar al observado en el pasado. Sin embargo, esta última suposición es difícil de asumir, sobre todo, en el caso de las tasas de fecundidad, por lo que resulta más conveniente utilizar la distribución uniforme por su carácter de máxima incertidumbre, lo que implicará obtener dispersiones más amplias que con modelos más estrictos.

Los módulos 3, 4 y 5 sirven para obtener las probabilidades perspectivas de paso, las tasas específicas de fecundidad y el saldo migratorio por sexo y edad de cada subperíodo de proyección, respectivamente, tal y como se ha indicado en la descripción de las variables del modelo.

El módulo 6 obtiene la población proyectada en un subperíodo aplicando el método de las componentes, ajustando las migraciones.

Este modelo de simulación hace que se repita el procedimiento un número elevado de veces, que serán otras tantas trayectorias del modelo estocástico descrito, que darán lugar a otros tantos resultados. En el módulo final, se determina el resultado medio y, calculando los percentiles 5 y 95, se puede obtener un intervalo que recoja el 90% central de las realizaciones. No será un intervalo de confianza, en sentido estricto, pero sí puede servir para indicar que, si se efectuasen un elevado número de proyecciones, el 90% de sus resultados estaría comprendido entre dichos extremos, aproximadamente.

2.3. Ventajas del modelo

Entre las ventajas de este planteamiento, destaca que, si se toman adecuadamente los parámetros necesarios, el intervalo resultante lleva incorporadas las dos hipótesis, alta y baja, para cada uno de los fenómenos vitales (mortalidad, fecundidad y migración), que en una proyección clásica dan lugar a $2^3 = 8$ proyecciones.

Otra ventaja es que permite hacer análisis de sensibilidad, ya que variando los parámetros iniciales en cierta proporción, se puede determinar como influyen en los resultados. Y, en fin, otra ventaja adicional de este modelo de simulación es que se puede programar utilizando el esquema que figura en el Anexo 1.

3. ADAPTACIÓN DEL MODELO AL CASO ESPAÑOL

Para adaptar el modelo de simulación descrito anteriormente a los datos de España, se ha confeccionado un programa cuyo esquema, en líneas generales, es el descrito en el Anexo 1 y que, mediante una rutina conveniente, permite realizar un número suficientemente grande de iteraciones, con cuyos resultados se obtienen los diferentes valores medios, varianzas y percentiles.

3.1. Hipótesis de partida

Se han adoptado las siguientes hipótesis iniciales:

- a) Para determinar las distintas probabilidades perspectivas de paso, se ha optado por utilizar las tablas de mortalidad calculadas en un trabajo anterior⁶ para la población española; en particular, las correspondientes a los años 1990-91, que proporcionan

6. Gutiérrez de Mesa, J.L. (2000).

una esperanza de vida al nacer de las mujeres de 80,62 años; las de 1995-96, que se corresponden con 82,01 años y las correspondientes al año 2020-25, que suponen una esperanza de vida de las mujeres de 85,95 años, ya que están más adaptadas al caso de España que las tablas-tipo de mortalidad de Coale y Guo para Europa Occidental.

- b) El esquema de distribución de los nacimientos por edad de la madre, que se mantiene constante durante todo el periodo de la proyección, coincide con el de 1995, último periodo observado antes de la misma. Por tanto, las variaciones en el ISF se asignan proporcionalmente según dicho esquema para obtener las distintas tasas de fecundidad.
- c) En el caso del Saldo Migratorio se ha aceptado que las emigraciones van a ser muy escasas y no van a desvirtuar las proyecciones, tal y como se ha asumido por diversos organismos que efectúan proyecciones de población para el caso de España, con lo que únicamente se van a tener en cuenta las inmigraciones cuya distribución por sexo y edad, que se va a mantener constante a lo largo de todo el periodo de proyección, es la utilizada por el Instituto de Demografía del CSIC (1994) a falta de unos estudios más recientes sobre esta composición y la escasez de datos al respecto.

3.2. Elección del escenario.

Como ya se ha indicado en la descripción del modelo, cada combinación de valores de los parámetros dará lugar a un escenario diferente, y se trata, pues, de elegir una combinación de valores que genere un escenario demográficamente aceptable. Ahora bien, la selección de los valores iniciales del modelo es subjetiva, de acuerdo con los razonamientos expresados en el epígrafe 2.2, y permite la generación de escenarios posibles en el horizonte de predicción. Dada la incertidumbre de éste, una de las grandes ventajas del método radica precisamente en la posibilidad de modificar dicho escenario, teniendo en cuenta los nuevos hechos demográficos que se van conociendo, sin más que alterar estos valores iniciales.

En este caso, el escenario propuesto consiste en suponer, como indican las fuentes y organismos oficiales en sus previsiones (INE 2001a; CSIC, 1994; Fernández Cordón, 1998), que el ISF cambiará su tendencia, aumentando progresivamente hasta situarse en torno a 1,5 hijos por mujer al final de la proyección y que la esperanza de vida al nacer de las mujeres seguirá aumentando, aunque más moderadamente, hasta situarse en torno a los 85 años. Estas consideraciones no pueden hacer olvidar la posibilidad de que ambas variables se puedan comportar de modo contrario al especificado y, por ello, el modelo permitirá disminuciones en algunas de las realizaciones de la simulación, aunque lógicamente con menor probabilidad que la tendencia general adoptada.

Para dar cumplimiento a lo expuesto, una selección idónea que, además, permite una variabilidad acorde con los datos históricos de ambas magnitudes⁷, es:

$$a = 1, b = 2,5; a' = 0, b' = 0,5; \alpha = \alpha' = 0,25$$

Para calcular el número de inmigrantes, se ha considerado que anualmente van a entrar en España alrededor de 160.000, de acuerdo con las más recientes previsiones del I.N.E. en su última revisión⁸ y, con el fin de dotar a esta cifra de una cierta variabilidad, se han elegido los valores extremos $a'' = 150.000$ y $b'' = 170.000$.

4. RESULTADOS OBTENIDOS EN LA PROYECCIÓN DE LA POBLACIÓN ESPAÑOLA PARA EL PERIODO 1996-2025

Con los datos iniciales expuestos, se ha realizado una simulación de 1000 iteraciones y, con sus resultados, se han calculado los valores medios, los percentiles 5 y 95, la amplitud del intervalo resultante y la desviación típica. Se han obtenido los resultados que figuran en la Tabla A1 del Anexo 2, donde están representados, por separado, los valores medios, intervalos truncados al 90%, amplitud de ese intervalo y la desviación típica, todas ellas para la población total, así como sus clasificaciones por sexo y edad, y complementados por las principales razones y proporciones para los años 2000, 2005, 2010, 2015, 2020 y 2025. Se presentan, a continuación unos cuadros resumen de valores medios y extremos superiores e inferiores de los intervalos⁹.

CUADRO 1. VALORES MEDIOS OBTENIDOS MEDIANTE SIMULACIÓN ESTOCÁSTICA

	2000	2005	2010	2015	2020	2025
Total	40837270	41905353	42723163	43247960	43599482	43979666
Varones	19955163	20454720	20827570	21053198	21188204	21326526
Mujeres	20882107	21450633	21895593	22194762	22411278	22653139
Jóvenes, de 0 a 14 años	6214616	6413120	6718331	6544543	6192966	5986942
Pot activos, de 15 a 64 años	27749640	28316675	28504134	28747308	28919778	28602419
Mayores	6873014	7175558	7500698	7956110	8486738	9390305
% jóvenes	48,865	48,812	48,750	48,679	48,596	48,489
% potencialmente activos	51,135	51,188	51,250	51,321	51,404	51,511
% mayores	15,217	15,299	15,713	15,113	14,178	13,579
Razón de dependencia	67,953	67,577	66,730	66,488	66,352	65,061
Razón Mayores/Pot. Activos	16,830	17,124	17,558	18,399	19,470	21,360

7. Pueden verse, por ejemplo, en Gutiérrez de Mesa (2000).

8. INE (2001a).

9. Se refiere a los intervalos (P_5 , P_{95}).

CUADRO 2. EXTREMOS INFERIORES OBTENIDOS MEDIANTE SIMULACIÓN ESTOCÁSTICA

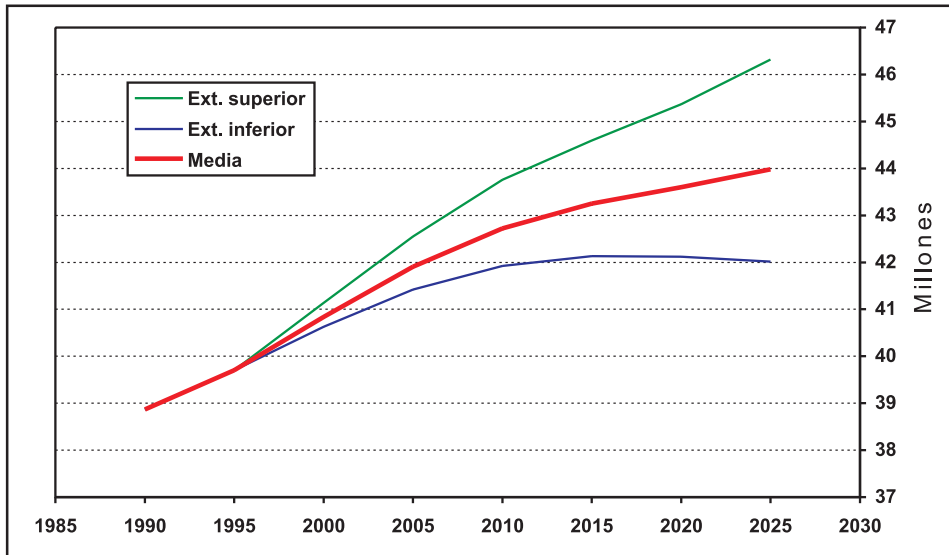
	2000	2005	2010	2015	2020	2025
Total	40630523	41421802	41922860	42134812	42120406	42012445
Varones	19850224	20206555	20417013	20481004	20422015	20313865
Mujeres	20781197	21209588	21504283	21642675	21689136	21692929
Jóvenes, de 0 a 14 años	6025563	5985927	6013963	5659757	5238625	4865930
Pot activos, de 15 a 64 años	27740406	28272548	28437307	28539264	28482705	27882369
Mayores	6801349	7000770	7225821	7592823	8020948	8799820
% jóvenes	48,850	48,777	48,694	48,597	48,489	48,349
% potencialmente activos	51,112	51,140	51,174	51,216	51,271	51,341
% mayores	14,827	14,447	14,312	13,391	12,326	11,500
Razón de dependencia	67,453	66,552	65,117	64,843	64,635	63,334
Razón Mayores/Pot. Activos	16,634	16,701	16,893	17,514	18,342	19,940

CUADRO 3. EXTREMOS SUPERIORES OBTENIDOS MEDIANTE SIMULACIÓN ESTOCÁSTICA

	2000	2005	2010	2015	2020	2025
Total	41136433	42551908	43760672	44594869	45369095	46317283
Varones	20111950	20791726	21360763	21757566	22106215	22525385
Mujeres	21028892	21766584	22392314	22854187	23267948	23788503
Jóvenes, de 0 a 14 años	6507299	7044547	7747344	7685234	7472572	7405813
Pot activos, de 15 a 64 años	27759261	28358446	28566314	29049450	29565562	29614722
Mayores	6959903	7366024	7813762	8381218	9035394	10099343
% jóvenes	48,888	48,860	48,825	48,783	48,729	48,658
% potencialmente activos	51,150	51,223	51,306	51,403	51,511	51,650
% mayores	15,827	16,570	17,692	17,294	16,506	15,943
Razón de dependencia	68,283	68,333	67,942	67,924	67,861	66,671
Razón Mayores/Pot. Activos	17,038	17,038	17,038	17,038	17,038	17,038

En estos resultados, se puede observar, como es lógico, que a lo largo de los años crece la amplitud de cada intervalo y, por tanto, la desviación típica. Además, debe tenerse en cuenta que no se ha intentado obtener un resultado final por grupos de edad y, por tanto, las clasificaciones no responden a una población concreta sino a los valores medios y percentiles de mil realizaciones simuladas, lo que indica que la suma de hombres y mujeres no será necesariamente la población total, y que la suma de los tres grandes grupos de edad podría asimismo no coincidir con ella. Estos valores tienden a estabilizarse al crecer el número de iteraciones, de acuerdo con la ley de los grandes números.

En el siguiente gráfico se puede observar la evolución de la población total media y sus extremos de variación donde se encuentra el 90% de las realizaciones.

Gráfico 1. Población proyectada de España y extremos de variación del 90%

5. COMPARACIÓN CON OTRAS PROYECCIONES.

Los resultados obtenidos mediante este modelo de simulación no son directamente comparables con otras proyecciones de población realizadas, ya que no coinciden exactamente las hipótesis de partida. No obstante, se van a comparar con las proyecciones de población de J.A. Fernández Cordon, publicadas por FEDEA¹⁰, las realizadas por el Instituto de Demografía del CSIC y la referencia obligada de las proyecciones oficiales del INE¹¹. Los resultados obtenidos por el Instituto de Demografía del CSIC se supone que están ya corregidos por las proyecciones de FEDEA, ya que su autor colaboró intensamente en aquellos y sigue su metodología.

El INE, en 1995, realiza una primera proyección de población, basada en el Censo de 1991, y aunque establece diversas hipótesis de mortalidad y fecundidad, únicamente publica los resultados medios y, en cuanto a las migraciones, supone que no hay emigración al extranjero y que los retornos de inmigrantes serán muy reducidos, estimándolos en 35.000 entradas anuales totales. De ahora en adelante esta proyección se denominará INE(1995). Recientemente, este Organismo ha publicado una revisión de la proyección anterior, corri-

10. Fernández Cordon, J.A. (1998).

11. INE (1995) , INE (2001a).

giendo ligeramente las hipótesis de mortalidad y fecundidad, pero, aunque establece varios escenarios en cuanto al número de inmigrantes, se decanta por la cifra de 160.000 anuales. A esta segunda proyección se le denominará INE(2001).

J.A Fernández Cordon, en FEDEA, presenta dos proyecciones. El Escenario A, aquí denominado Fedea A, establece una fecundidad creciente hasta 1,8 hijos por mujer en el año 2025 y una migración que consiste en admitir tantos inmigrantes como trabajadores se jubilen o mueran; en realidad, parte de 30.000 y, desde el año 2005, se incrementa hasta llegar a 220.000 en el año 2025. El Escenario B, aquí denominado Fedea B, supone un incremento de la fecundidad hasta 1,6 hijos por mujer en el año 2025 y una inmigración constante de 30.000 individuos.

En primer lugar, hay que tener en cuenta que las proyecciones presentadas en este trabajo están basadas en la población del Padrón Municipal del 1 de Mayo de 1996 y las otras se refieren a datos del Censo de 1991, actualizando los datos de mortalidad y fecundidad al momento que estuvieron disponibles.

Comparando los resultados globales obtenidos por el modelo de simulación y los obtenidos por las otras proyecciones, se presenta en primer lugar un cuadro donde se recogen los resultados de la población total.

CUADRO 8. POBLACIÓN TOTAL DE ESPAÑA CALCULADA AL 31 DE DICIEMBRE SEGÚN DIVERSAS PROYECCIONES

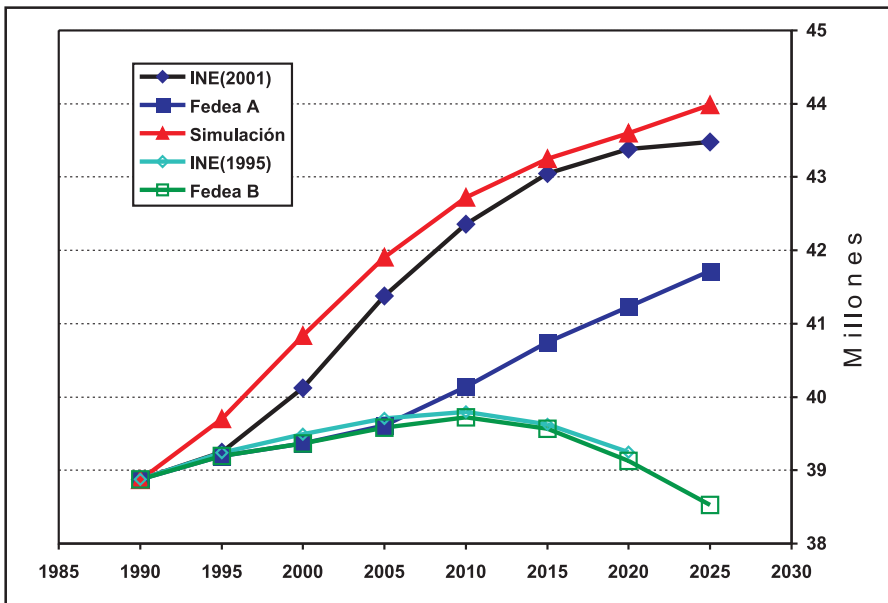
	1990	1995	2000	2005	2010	2015	2020	2025
INE(1995)	38874573	39241933	39489635	39709849	39798238	39624380	39247681	
Fedea B	38872268	39197000	39364000	39581000	39725000	39566000	39125000	38528000
Fedea A	38872268	39197000	39366000	39611000	40135000	40744000	41226000	41718000
INE(2001)	38874573	39249083	40121673	41378924	42359314	43048658	43378337	43480804
Simulación		39698230	40837270	41905353	42723163	43247960	43599482	43979666

En el gráfico 2, se recoge la evolución del número total de la población española según las diversas proyecciones hasta el año 2025.

Lo primero que cabe señalar es que los resultados de 1995 de las otras proyecciones, no están actualizados con los resultados del Padrón de 1996, como ya se ha indicado; por ello, son previsiones que se quedan cortas en unos quinientos mil individuos. También es de reseñar que las previsiones de Fedea A son cada vez mayores debido a la suposición de una inmigración cada vez más numerosa, hipótesis no recogida en otras proyecciones. Por otra parte, el número total de individuos obtenidos con el modelo propuesto de simulación es muy parecido al obtenido por el INE(2001), considerando los quinientos mil individuos de

error de partida en 1995, ya que se emplean hipótesis de partida muy similares, pero con la ventaja de que se obtiene, además, un intervalo donde aproximadamente se situará el 90% de las realizaciones.

Gráfico 2. Evolución de la población total de España hasta el 2025, según diversas proyecciones



6. CONCLUSIONES

En este trabajo, se ha presentado un modelo de simulación estocástica con el fin de poder obtener unos valores medios de la población futura y unos intervalos de fluctuación que recojan el 90% de la variabilidad total. El modelo se basa en suponer comportamientos aleatorios de los indicadores sintéticos de la fecundidad y de la mortalidad, como son el Índice Sintético de Fecundidad y la esperanza de vida al nacer de las mujeres, así como en asumir un saldo migratorio aleatorio que se distribuyen uniformemente en un determinado intervalo. Además de la explicación del modelo, se ha indicado su descomposición en módulos.

Este modelo se ha aplicado a la población de España, para obtener 1000 proyecciones, desde 1995 hasta el año 2025, de las cuales se han obtenido diversos valores e indicadores medios y los percentiles 5 y 95, para obtener un intervalo que contenga aproximadamente al 90% de las realizaciones. Se han presentado los diversos resultados obtenidos, se han

comparado con los de otras proyecciones y se ha comprobado que el modelo presenta resultados muy satisfactorios.

Además, se puede concluir que según las últimas cifras publicadas por el I.N.E.¹², la población total de España al 1 de enero del año 2000 es de 40.499.791 habitantes, cifra muy aproximada a la que se obtiene por interpolación geométrica entre la población al 31 de diciembre del año 1995 y la del año 2000 aquí calculada, que asciende a 40.606.877 individuos en media. Obviamente, esta cifra resulta superior a la oficial ya que aquí se ha supuesto una distribución uniforme de los inmigrantes a lo largo del primer quinquenio, en lugar de los 35.000 que ya se sabe que se recibieron en los primeros años de este periodo, que luego se incrementaron paulatinamente hasta redondear la cifra media de 160.000, que el INE aplica en su última revisión, de 2001.

Sería útil estudiar la velocidad de convergencia del algoritmo mediante el análisis del efecto del número de iteraciones sobre la solución final. En este sentido, las pruebas que se han realizado con un número inferior de iteraciones han dado como resultado valores muy parecidos, de donde se puede esperar que este factor no incida de manera notable sobre la estabilidad de los resultados, siempre que, obviamente, el número de iteraciones sea suficientemente elevado.

Finalmente, conviene señalar que, pese a los buenos resultados obtenidos, queda por efectuar un estudio profundo sobre su sensibilidad con respecto a los valores iniciales, por lo que se ha separado el primer módulo en el algoritmo. No obstante, esta segmentación permite dotar de un cierto dinamismo al modelo ya que, mediante una modificación de este módulo a la vista de la realidad observada, dota de resultados actualizados frente a escenarios concretos cambiantes.

Agradecimientos

Los autores agradecen las acertadas sugerencias emitidas por los evaluadores anónimos, que han contribuido a mejorar la versión final del artículo.

BIBLIOGRAFÍA

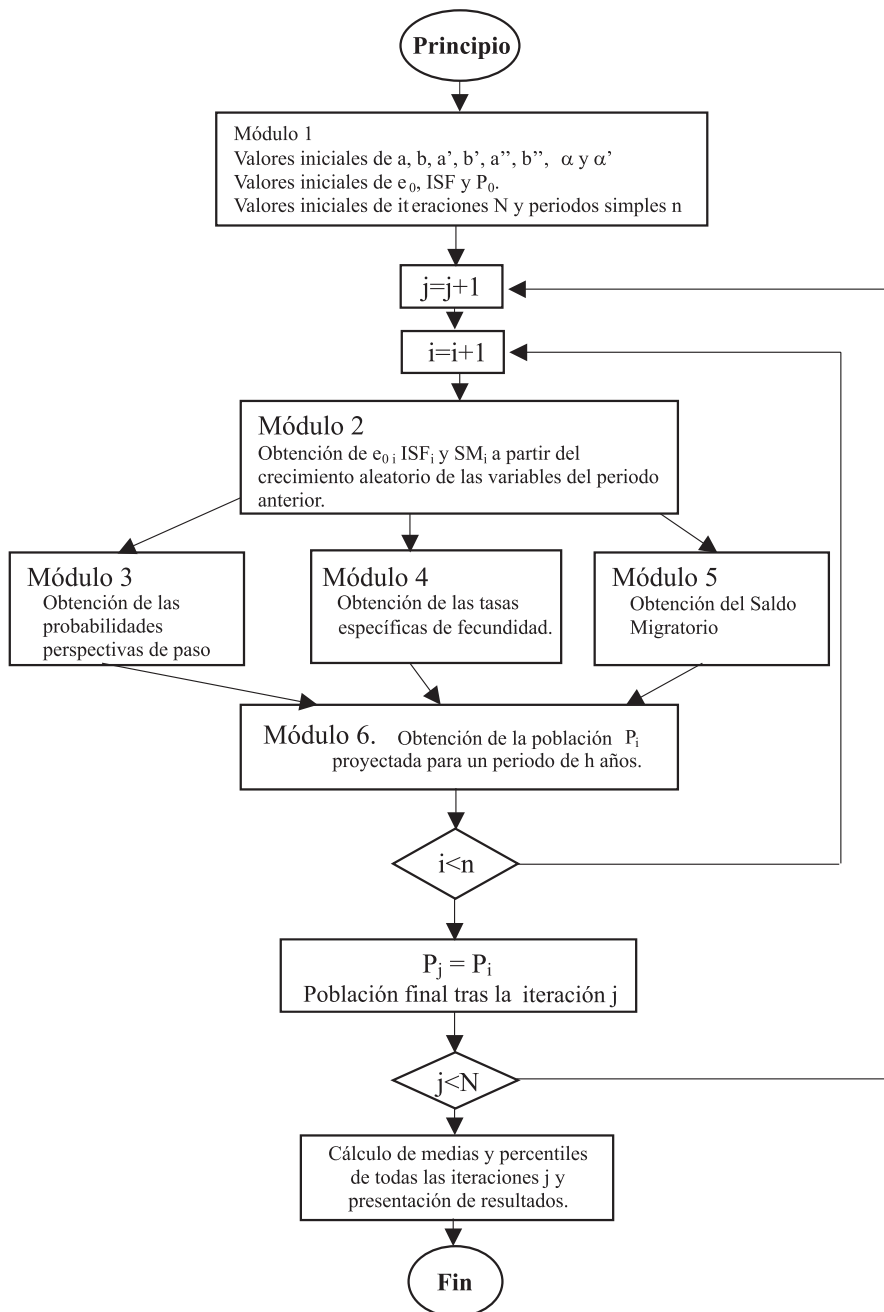
- AHLBURG, D.A. (1987). "Aggregate Economic-Demographic Models". En *Forecasting in Social and Natural Sciences* (K.C. Land y S.H. Schneider, eds.). D. Reidel. Boston.
- ALHO, J., (1990). "Stochastic methods in population forecasts", *International Journal of Forecasting*, 6, 521-530.
- BRUNBORG, H. y KEILMAN, N. (1995). "MOSART-H: A combined micro-macro model for simulation of household". En *Demography, economy and welfare, Scandinavian Population Studies*, vol. 10 (C. Lundh, ed.). Lund University Press. Lund. pp. 435-452.

12. INE(2001b).

- COALE, A. y GUO, G. (1989). "Revised regional model life tables at very low levels of mortality", Population Index, 55 nº4, pp. 613-643.
- COALE, A. y GUO, G. (1991).- "Utilización de nuevas tablas modelo de mortalidad para tasas de mortalidad muy bajas en proyecciones demográficas." Boletín de población de las Naciones Unidas nº 30. Naciones Unidas, Nueva York.
- COMUNIDAD AUTÓNOMA DE MADRID (1994). Proyecciones de población de la CAM al 2001. Detalle metodológico. Madrid
- COHEN, J.E., (1986). "Population forecasts and confidence intervals for Sweden: A comparison of model-based and empirical approaches", Demography, 23, 105-126.
- FERNANDEZ CORDON, J.A., (1998). "Proyección de la Población Española.", Documento de trabajo 98-11, FEDEA.
- GUTIÉRREZ DE MESA, J. L. (2000). "Crecimiento de la población, Tasas vitales y modelos de previsión". Facultad de Ciencias Económicas y Empresariales de la Universidad de Alcalá. Alcalá de Henares, Madrid.
- INSTITUTO DE DEMOGRAFIA DEL CSIC.(1994). Proyección de la Población Española de 1991 al 2026. Madrid.
- INSTITUTO NACIONAL DE ESTADISTICA (1995). Proyecciones de la población de España calculadas a partir del Censo de Población de 1991. Madrid.
- INSTITUTO NACIONAL DE ESTADISTICA (2001a). Proyecciones de la población de España calculadas a partir del Censo de Población de 1991. Evaluación y revisión. Madrid.
- INSTITUTO NACIONAL DE ESTADISTICA (2001b). Rectificación del Padrón de habitantes del año 2000. Madrid.
- KEYFITZ, N., (1977). Applied Mathematical Demography. Wiley and Sons. New York.
- KRUPP H.J., GALLER H.P., GROHMAN H., HAUSE R.y WAGNER G., (1981). Alternativen der Rentenreform '84, Campus Verlag, Frankfurt.
- LAND, K.C., (1986). "Methods for national population forecasts: A review", Journal of the American Statistical Association, 81, 888-901.
- LAND K.C. y CANTOR, D. (1983). "ARIMA Models of Seasonal Variation in U.S. Birth and Death Rates". Demography, 20, 541-568.
- LONG, J.F. y D. MCMILIEN, (1987). "A Survey of Census Bureau Projection methods". En Forecasting in the Natural and Social Sciences (K. Land y S. Schneider, eds.).D. Reidel. Boston, pp. 141-178.
- NELISSEN, J.H.M. (1994). Income redistribution and social security: An application of microsimulation, Chapman & Hall, London.
- ORCUTT, G., CALDWELL, S., y WERTHEIMER, R., III (1976). Policy Exploration Through Microanalytic Simulation. Urban Institute. Washington, D.C.
- PFLAUMER, P., (1988). "Confidence intervals for population projections based on Monte Carlo methods", International Journal of Forecasting, 4, 135-142.
- VAN IMHOFF, E. y POST, W. (1997). "Méthodes de micro-simulation pour des projection de population". Population, 4, 889-932.
- WHELPTON, P.K. (1947). Forecast of the Population of the United States, 1945-1975. U.S. Government Printing Office. Washington, DC.

ANEXO 1

ESQUEMA DEL PROCESO DE PROGRAMACIÓN DEL MODELO



ANEXO 2

TABLA A1. RESULTADOS DE LA PROYECCIÓN REALIZADA POR SIMULACIÓN.

Población total	2000	2005	2010	2015	2020	2025
Media	40837270	41905353	42723163	43247960	43599482	43979666
Límite inferior	40630523	41421802	41922860	42134812	42120406	42012445
Límite superior	41136433	42551908	43760672	44594869	45369095	46317283
Amplitud	505910	1130106	1837811	2460057	3248689	4304839
desviación típica	154101	348581	567585	787094	1025987	1314690
Varones	2000	2005	2010	2015	2020	2025
Media	19955163	20454720	20827570	21053198	21188204	21326526
Límite inferior	19850224	20206555	20417013	20481004	20422015	20313865
Límite superior	20111950	20791726	21360763	21757566	22106215	22525385
Amplitud	261725	585171	943751	1276563	1684200	2211521
desviación típica	79161	178997	291423	404006	526762	675456
Mujeres	2000	2005	2010	2015	2020	2025
Media	20882107	21450633	21895593	22194762	22411278	22653139
Límite inferior	20781197	21209588	21504283	21642675	21689136	21692929
Límite superior	21028892	21766584	22392314	22854187	23267948	23788503
Amplitud	247694	556996	888031	1211512	1578813	2095575
desviación típica	75038	169789	276487	383531	499745	639780
Jóvenes	2000	2005	2010	2015	2020	2025
Media	6214616	6413120	6718331	6544543	6192966	5986942
Límite inferior	6025563	5985927	6013963	5659757	5238625	4865930
Límite superior	6507299	7044547	7747344	7685234	7472572	7405813
Amplitud	481736	1058620	1733381	2025478	2233948	2539883
desviación típica	145864	327013	530545	630200	686097	756538
En edad activa	2000	2005	2010	2015	2020	2025
Media	27749640	28316675	28504134	28747308	28919778	28602419
Límite inferior	27740406	28272548	28437307	28539264	28482705	27882369
Límite superior	27759261	28358446	28566314	29049450	29565562	29614722
Amplitud	18855	85898	129007	510185	1082857	1732353
desviación típica	5814	27125	38907	154561	332781	535538
Mayores de 65	2000	2005	2010	2015	2020	2025
Media	6873014	7175558	7500698	7956110	8486738	9390305
Límite inferior	6801349	7000770	7225821	7592823	8020948	8799820
Límite superior	6959903	7366024	7813762	8381218	9035394	10099343
Amplitud	158554	365254	587942	788395	1014446	1299522
desviación típica	49946	109952	176626	240887	309047	389738

TABLA A1. RESULTADOS DE LA PROYECCIÓN REALIZADA POR SIMULACIÓN. (Cont.)

% de jóvenes	2000	2005	2010	2015	2020	2025
Media	15,217	15,299	15,713	15,113	14,178	13,579
Límite inferior	14,827	14,447	14,312	13,391	12,326	11,500
Límite superior	15,827	16,570	17,692	17,294	16,506	15,943
Amplitud	1,000	2,123	3,379	3,903	4,181	4,443
desviación típica	0,303	0,659	1,040	1,195	1,263	1,346
% de edad activa	2000	2005	2010	2015	2020	2025
Media	67,953	67,577	66,730	66,488	66,352	65,061
Límite inferior	67,453	66,552	65,117	64,843	64,635	63,334
Límite superior	68,283	68,333	67,942	67,924	67,861	66,671
Amplitud	0,830	1,781	2,825	3,081	3,226	3,336
desviación típica	0,251	0,547	0,863	0,971	0,993	1,003
% de mayores de 65 años	2000	2005	2010	2015	2020	2025
Media	16,830	17,124	17,558	18,399	19,470	21,360
Límite inferior	16,634	16,701	16,893	17,514	18,342	19,940
Límite superior	17,038	17,540	18,197	19,308	20,645	22,817
amplitud	0,404	0,839	1,305	1,794	2,302	2,877
desviación típica	0,118	0,253	0,398	0,536	0,683	0,870
Razón de dependencia.	2000	2005	2010	2015	2020	2025
Media	47,163	47,988	49,884	50,435	50,745	53,739
Límite inferior	46,445	46,338	47,171	47,212	47,351	49,937
Límite superior	48,233	50,254	53,531	54,154	54,705	57,881
amplitud	1,788	3,916	6,359	6,942	7,354	7,944
desviación típica	0,546	1,203	1,951	2,207	2,265	2,379
Dependencia de los mayores	2000	2005	2010	2015	2020	2025
Media	24,768	25,340	26,314	27,676	29,348	32,839
Límite inferior	24,518	24,741	25,378	26,401	27,693	30,611
Límite superior	25,072	26,008	27,381	29,083	31,208	35,324
amplitud	0,555	1,266	2,003	2,682	3,514	4,713
desviación típica	0,175	0,378	0,605	0,830	1,077	1,421