

Proyección estocástica de la mortalidad mexicana por medio del método de Lee-Carter*

Víctor Manuel García Guerrero**

Manuel Ordorica Mellado***

En este trabajo se aplica el método de Lee-Carter para proyectar la mortalidad de México. Dicho método supone una relación lineal entre el logaritmo natural de las tasas centrales de mortalidad y la edad y el tiempo; una de sus ventajas es que es parsimonioso, pues combina un modelo demográfico con pocos supuestos con el análisis de series de tiempo; otra ventaja es que es posible medir la incertidumbre asociada al evento mortalidad y así calcular intervalos de confianza. Los pronósticos de la esperanza de vida al nacer de la población femenina para 2050 prevén con un 95% de confianza que podría variar entre 83.53 y 86.40 años, y en el caso de los hombres oscilaría entre 78.53 y 80.53 años.

Palabras clave: mortalidad mexicana, proyección de la mortalidad, métodos demográficos, pronóstico estocástico, método de Lee-Carter.

Fecha de recepción: 7 de julio de 2011.

Fecha de aceptación: 16 de noviembre de 2011.

Stochastic Projection of Mexican Mortality through the Lee-Carter Method

This study uses the Lee-Carter method to project mortality in Mexico. This method assumes a linear connection between the natural logarithm of the central rates of mortality and age and time. One of its advantages is that it is parsimonious, since it combines a demographic model with few assumptions with the analysis of time series. Another advantage is that it is possible to measure the uncertainty associated with the mortality event and thus calculate confidence intervals. The forecast of life expectancy at birth for female population to 2050 foresee, with 95% of confidence that it could vary between 83.53 and 86.40 years and in the case of men, would oscillate between 78.53 and 80.53 years.

Key words: Mexican mortality, projection of mortality, demographic methods, stochastic forecast, Lee-Carter method.

* Los autores agradecen los comentarios y sugerencias que dos árbitros anónimos hicieron a este artículo.

** Profesor investigador del Centro de Estudios Demográficos, Urbanos y Ambientales de El Colegio de México. Correo electrónico: <vmgarcia@colmex.mx>.

*** El Colegio de México. Correo electrónico: <mordori@colmex.mx>.

Introducción

La elaboración de pronósticos sobre los niveles de mortalidad es una tarea fundamental para la definición de las políticas de salud y para realizar proyecciones demográficas, las cuales a su vez son un instrumento básico para la definición de la política de población. Para el año 2011 se estima que México tiene una esperanza de vida al nacer de 77 años, siete menos que la de Japón para este mismo año. En medio siglo nuestro país ha incrementado en 17 años la esperanza de vida al nacimiento. Según las estimaciones de Naciones Unidas, a mitad del siglo XXI México alcanzará el nivel que tiene hoy Japón, mientras que ese país llegará casi a los 90 años en 2050. Esta cifra es todavía pequeña si la comparamos con la edad a la que murió la francesa Jeanne Calment –122 años– en 1997. Las proyecciones de la mortalidad realizadas en los años setenta del siglo XX indicaban que en el año 2000 México llegaría a 70 años en la esperanza de vida al nacer, pero la realidad mostró una cifra cuatro años mayor que la proyectada. Esto muestra una enorme falla en la proyección de la esperanza de vida a la edad cero, de casi 25%, si bien entonces se desconocía que los avances de la ciencia médica contribuirían a que la esperanza de vida fuera superior a lo proyectado.

Existen diversos métodos para proyectar los niveles de la mortalidad por edad, aunque al final todos suponen una caída muy parecida a la evolución de una función logística; es decir, una disminución lenta al principio, luego una declinación rápida, y finalmente de nuevo una disminución lenta. Esta dinámica corresponde con lo que se espera de acuerdo con la transición demográfica.

Entre la gran diversidad de métodos de proyección se encuentran algunos relativamente sencillos que utilizan escala semilogarítmica; se ubican las tasas de mortalidad por edad o las probabilidades de muerte en el eje de las *Y* y el tiempo en el eje de las *X*. Las tasas se extrapolan siguiendo una línea recta. Otros métodos utilizan funciones matemáticas, como las exponenciales modificadas, Gompertz y Makeham, entre otras. Para tratar prácticamente estas funciones es necesario hacer un cambio de variable que permita transformarlas en una recta y de esta manera estimar los parámetros mediante el método de regresión lineal simple. Todas ellas se caracterizan por tener un valor asintótico al que tiende la mortalidad por edad a medida que pasa el tiempo. Ello impide tener valores de las tasas de mortalidad por edades que incluso pudieran ser negativos.

Existen también métodos que utilizan las funciones de tipo *logito* o *probit*, estableciendo una relación lineal en diferentes momentos en el tiempo entre dicho logito de las funciones de supervivencia observadas y uno estándar. Luego se proyectan los valores de los parámetros y se obtiene un nuevo logito proyectado, el cual permite obtener los nuevos valores de la función de supervivencia.

Cuando se tiene poca información o ésta es deficiente, se aplican métodos que primero proyectan la esperanza de vida al nacer mediante alguna función parecida a una logística ascendente y luego utilizan tablas modelo de mortalidad para establecer la estructura por edades. En casi todos los casos descritos, directa o indirectamente, los parámetros de los diferentes modelos son variables aleatorias.

El objetivo del presente trabajo es aplicar la metodología propuesta por Lee y Carter (1992) a la mortalidad mexicana por sexo y comparar los resultados con las últimas proyecciones realizadas por el Consejo Nacional de Población (Conapo, 2006). El método de Lee-Carter supone una relación lineal entre el logaritmo natural de las tasas centrales de mortalidad y dos elementos que la explican: la edad y el tiempo. Una de las ventajas del método de Lee-Carter es que es parsimonioso en el sentido de que combina un modelo demográfico con pocos supuestos con los métodos de las series de tiempo. Esto permite aprovechar la teoría que subyace a este método. Además es posible tener intervalos de confianza de las estimaciones.

Antecedentes

Las proyecciones demográficas han cautivado a los demógrafos desde hace muchos años. En el aspecto metodológico ha habido algunos intentos por tomar en cuenta la incertidumbre inherente a toda prefiguración del futuro. Entre los principales trabajos sobre proyecciones estocásticas se encuentran los de Sykes (1969), Le Bras (1971), Saboia (1974 y 1977), Heyde y Cohen (1985), Cohen (1986), Alho y Spencer (2005), Alho (1985 y 1990), Lee (1993 y 1974), Lee y Carter (1992), Lee y Tuljapurkar (1994) y, recientemente, Hyndman y Booth (2008).

Los trabajos de Sykes (1969), Le Bras (1971), Cohen (1986) y Heyde y Cohen (1985) se caracterizan por ampliar la formulación de las matrices de proyección propuestas por Leslie al campo probabilístico. Sykes propone algunas medidas probabilísticas para determinar

la precisión de las proyecciones de población, de manera que a partir del modelo de proyección determinista prueba tres modelos estocásticos para la proyección del crecimiento poblacional, resultando ser el más eficiente aquel que supone que las matrices de transición son variables aleatorias. Por otro lado, Le Bras establece relaciones aleatorias entre las distintas variables demográficas considerando la varianza y la covarianza en el modelo de proyección. Finalmente, en los trabajos de Heyde y Cohen se extiende la formulación de Leslie a una matriz con tasas vitales aleatorias; para ello demuestran que la probabilidad conjunta de cualquier número finito de matrices de Leslie es invariante respecto a los cambios en el tiempo; también suponen que la secuencia de las matrices de Leslie es ergódica si para cada evento converge casi seguramente a la probabilidad del evento, y que las tasas de la matriz de Leslie se vuelven cada vez más independientes conforme el tiempo entre ellas se aleja. Con lo anterior pronostican la población sueca y su tasa de crecimiento, y construyen sus correspondientes intervalos de confianza.

Estos trabajos sirvieron de base para desarrollos metodológicos ulteriores. Pero los que tuvieron más impacto en los métodos actualmente utilizados fueron los trabajos de Saboia (1974 y 1977) y Lee (1974). En ellos, los autores utilizan la metodología de series de tiempo propuesta por Box y Jenkins (1976) para pronosticar la población total de Suecia, los nacimientos en Noruega y los nacimientos en Estados Unidos, respectivamente.

Por otro lado, Alho y Spencer (2005), y Alho (1985 y 1990) proponen modelar el tamaño de los grupos por edad y sexo de acuerdo con funciones no-lineales de tasas vitales aleatorias; asimismo proponen una medida de volatilidad para la fecundidad. En trabajos más recientes, Alho ha propuesto una forma de pronosticar los componentes demográficos basándose en el uso de distribuciones de probabilidad conocidas así como de métodos bayesianos. Sus trabajos han influido en la forma en que se calculan algunas de las proyecciones demográficas en Europa y Asia.

Los trabajos de Lee (1993), Lee y Carter (1992) y Lee y Tuljapurkar (1994) se basan en el trabajo de Gómez de León (1990). En ellos se desarrolló uno de los modelos que más impacto han tenido en los últimos tiempos en la modelación y pronóstico de la mortalidad. Con este modelo, Lee pronostica las tasas específicas de natalidad y las centrales de mortalidad utilizando el análisis estadístico de series de tiempo. El método que se presentará más adelante se basa en las pro-

puestas metodológicas de Lee, por lo que se explicarán con más detalle en las secciones siguientes.

El último y más reciente exponente es el trabajo de Hyndman y Booth (2008). Estos autores proponen un método muy innovador para pronosticar las poblaciones basado en *modelos de datos funcionales*. A grandes rasgos, su método trata de modelar las tasas vitales por edad y sexo de acuerdo con modelos funcionales con coeficientes estimados con series de tiempo. La migración es estimada como la diferencia entre los datos poblacionales históricos y los pronosticados para cada año. Para eliminar el error de pronóstico se utiliza la transformación de Box-Cox. Después, con el fin de hacer más robusto el pronóstico, se asume que los nacimientos y las defunciones siguen una distribución de *Poisson*. Finalmente, se utiliza el método de simulación de monte-carlo para pronosticar la población total. Esta nueva técnica parece ser muy robusta y sería muy interesante que se desarrollara en trabajos futuros con el fin de comparar los pronósticos aquí propuestos.

Para México, sólo unos cuantos demógrafos han intentado incursionar en el ámbito de los pronósticos poblacionales estocásticos. Entre ellos se encuentran los trabajos de Ordorica (2004 y 1995) y Kesseli y Galindo (2007). Ordorica (2004) realiza una proyección del número total de defunciones de México por medio de la metodología de series de tiempo propuesta por Box y Jenkins, es decir, siguiendo el modelo autorregresivo integrado de promedios móviles (ARIMA). La intención del autor es aplicar nuevas metodologías estadísticas para realizar pronósticos demográficos. Una limitante de esta metodología es, como bien reconoce el autor, que no se considera el efecto de la estructura por edad, aunque sigue siendo útil para el pronóstico de corto plazo. De este análisis resulta que entre los años 2000 y 2010 se esperan 420 mil muertes anuales. Sin embargo, de acuerdo con las estimaciones del Conapo para el año 2000 habría 456 mil defunciones, 474 mil para 2005 y 509 mil para 2010. Las diferencias radican en que el modelo aplicado no consideró la estructura por edad de la población.

En otro trabajo, Ordorica (1995) realiza una aplicación novedosa del filtro de Kalman para la estimación de la población mexicana de todo el país y en el análisis de políticas de población. La metodología propuesta por Ordorica es de gran utilidad para el análisis de series de tiempo con parámetros dependientes del tiempo. Un elemento muy importante de este trabajo es que el autor analiza el efecto en las proyecciones de población en el caso de que se adopten diferentes niveles de las varianzas de los errores. Cabe mencionar que este trabajo de

investigación es un hito en el campo de las proyecciones poblacionales, pues es el primero realizado para México en el que se ha adoptado un enfoque estocástico. Finalmente Ordorica muestra una aplicación del Filtro de Kalman para la evaluación y redefinición de objetivos en materia de políticas públicas. Aunque no se lleva a cabo una comprobación empírica, queda el modelo teórico para análisis posteriores.

Por último, en el trabajo de Kesseli y Galindo (2007) se elabora una proyección demográfica de tipo probabilístico utilizando el programa *Program on Error Propagation* (PEP) elaborado por Juha Alho,¹ y se realiza un análisis de la incertidumbre que rodea a la información demográfica de México. El trabajo de Kesseli y Galindo se centra en realizar algunas reflexiones sobre la calidad de las fuentes de información y sobre los múltiples problemas que pudieron influir para que se incrementara la incertidumbre de los resultados obtenidos en el Censo Nacional de Población realizado en 2005. De acuerdo con los autores es muy poco probable que la población contada en ese año hubiese sido inferior a 103.1 millones de habitantes. Una cuestión que sobresale en este trabajo es que muestra que también es posible realizar supuestos subjetivos respecto a la evolución futura de los componentes demográficos al utilizar el modelo de Alho (Alho y Spencer, 2005).

Modelo

El modelo de Lee y Carter (1992) (LC) es un modelo estadístico y demográfico que permite extrapolar las tasas centrales de mortalidad de acuerdo con su información histórica. No contiene variables exógenas, más bien se basa en el análisis estadístico de series de tiempo, es decir, combina un modelo demográfico con pocos supuestos, con métodos estadísticos de series de tiempo. Por ello este modelo provee una base estocástica al pronóstico de la mortalidad, así como intervalos de predicción.

El cambio en el nivel de la mortalidad se expresa en el modelo de LC por medio de un solo índice. Así, permite a cada tasa central de mortalidad decrecer exponencialmente, con lo que la evolución de la esperanza de vida se produce sin agregar algún supuesto adicional al de la tendencia observada históricamente.

¹ <<http://joyx.joensuu.fi/~ek/pep/pepstart.htm>>.

Existen dos variantes del modelo: una para series históricas completas (Lee y Carter, 1992) y otra para datos incompletos (Li, Lee y Tuljapurkar, 2004). En este trabajo se utilizará la versión original, ya que se cuenta con la serie histórica de las defunciones y la población expuesta al riesgo de 1960 a 2005, estimada en la conciliación demográfica de 2006 (INEGI, Conapo y El Colegio de México, 2006). González y Guerrero (2007) aplicaron el modelo de LC para datos incompletos.

El modelo de LC parte del supuesto de que existe una relación lineal entre el logaritmo natural de las tasas centrales de mortalidad $m(x, t)$ y dos factores que las determinan: la edad x y el tiempo t . El modelo matemático es expresado como sigue,

$$\ln(m_{x,t}) = a_x + b_x k_t + \varepsilon_{x,t}, \tag{1}$$

donde los parámetros de forma, $\{a_x\}$, y de cambio, $\{b_x\}$, varían con cada edad, y el índice de mortalidad, $\{k_t\}$, varía con el tiempo. El factor $\exp(a_x)$ es el patrón general por edad de la mortalidad. El parámetro b_x señala cuáles tasas disminuyen más rápido y cuáles más lento en función de los cambios en el índice de mortalidad. Esto significa que $d\ln(m_{x,t})/dt = b_x dk_t/dt$. Es decir, b_x es el ritmo al que disminuye la mortalidad a cada edad ante los cambios en k_t , y también puede interpretarse como la intensidad del cambio del índice de mortalidad para cada edad a lo largo del tiempo (González y Guerrero, 2007). Para ciertas edades b_x puede ser negativo indicando con ello que la mortalidad en esas edades tiende a incrementarse cuando disminuye en las otras edades. Si k_t es lineal en el tiempo, la mortalidad a cada edad cambia conforme a su propia tasa constante. Conforme k_t tiende a menos infinito, cada tasa central tiende a cero; debido a que las tasas son modeladas logarítmicamente no es posible obtener tasas negativas.

El primer paso para el pronóstico de la mortalidad de acuerdo con el modelo LC es ajustar los datos brutos al modelo. Para ello Lee y Carter sugieren restringir el modelo [1] a que $\sum_{x=0}^{\omega+} b_x = 1$ y $\sum_{t=1}^n k_t = 0$, con el fin de establecer la unicidad en la solución del sistema. Estas restricciones implican que

$$a_x = \frac{\sum_{t=1}^n \ln(m_{x,t})}{n}, \tag{2}$$

es decir, que el parámetro a_x se estima como el promedio aritmético simple del logaritmo de las tasas centrales a lo largo del tiempo. De esta manera se tiene que

$$\ln(m_{x,t}) - a_x = b_x k_t + \varepsilon_{x,t}. \quad [3]$$

Para obtener los vectores b_x y k_t Lee y Carter proponen utilizar la descomposición en valores singulares (DVS).² Formalmente, la factorización DVS indica que para toda $\mathbf{A} \in \mathbb{R}^{m \times n}$ de rango r existen matrices ortogonales $\mathbf{U}_{m \times m}$ y $\mathbf{V}_{n \times n}$ y una matriz diagonal $\mathbf{D}_{r \times r} = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$ tales que:

$$\mathbf{A} = \mathbf{U} \begin{pmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}_{m \times n} \mathbf{V}^t \text{ con } \sigma_1 \leq \sigma_2 \leq \dots \leq \sigma_r.$$

Las σ_i 's son los valores singulares de \mathbf{A} . Cuando $r < p = \min\{m, n\}$ se dice que \mathbf{A} tiene $p - r$ valores singulares iguales a cero. En la factorización anterior, a las columnas de \mathbf{U} y \mathbf{V} se les denomina vectores singulares de \mathbf{A} izquierdos y derechos, respectivamente.

En este contexto, si se aplica la DVS al lado izquierdo de la ecuación [3] los valores de b_x y k_t se obtienen de la siguiente manera: primero se toma el primer componente de la matriz \mathbf{U} (el primer vector singular o primera columna), y para obtener el índice de la mortalidad, k_t , se toma el primer componente de la matriz resultante del producto $\mathbf{D}\mathbf{V}^t$, es decir, el primer renglón de dicha matriz.

Otra forma de estimar el vector k_t es sumando cada una de las columnas de la matriz $\ln(m_{x,t}) - a_x$, por ejemplo sumar sobre todas las edades para cada t . En otros términos, para cada $t \in [1, n]$,

$$k_t = \sum_{x=0}^{\omega+} (\ln(m_{x,t}) - a_x).$$

Luego, para estimar b_x se resuelve la ecuación [3] por mínimos cuadrados ordinarios. Es decir, $b_x = \sum_{t=1}^n k_t (\ln(m_{x,t}) - a_x) / \sum_{t=1}^n k_t^2$ para cada $x \in [0, \omega+]$.

Existen otras maneras de estimar los parámetros del modelo de LC. Dos de las más eficientes fueron propuestas primero por Wilmoth (1993) y luego por Booth *et al.* (2002). El primero propone estimar los pará-

² La DVS se ha descubierto muchas veces independientemente. Entre aquellos a quienes se atribuyen los primeros desarrollos de este método destacan: Eugenio Beltrami (1835-1899) en 1873, M.E. Camille Jordan (1838-1922) en 1875, James J. Sylvester (1814-1897) en 1889, L. Autonne en 1913, y C. Eckart y G. Young en 1936. Para más información al respecto véase Meyer, 2000.

metros a_x, b_x y k_t minimizando la función $\sum_{x=0}^{\omega+} \sum_{t=1}^n D_{x,t} [\ln(m_{x,t}) - a_x - b_x k_t]^2$ por medio del método de mínimos cuadrados ponderados o por máxima verosimilitud, donde $D_{x,t}$ son las defunciones totales de los individuos a la edad x ocurridas durante el año t . El problema con este método es que no se obtienen ganancias sustanciales respecto de los otros modelos y la complicación algorítmica es ligeramente mayor.

Por otro lado, Heather Booth propuso ajustar k_t de tal manera que las defunciones anuales a cada edad son estimadas por medio de un modelo de regresión de Poisson. Para ello, a_x y b_x son los parámetros estimados por el modelo de LC original. El modelo de Poisson es:

$$\ln(D'_{x,t}) = \ln(E_{x,t} m'_{x,t}) + \varepsilon'_{x,t} = \ln(E_{x,t}) + \ln(m'_{x,t}) + \varepsilon'_{x,t}$$

donde $D'_{x,t}$ son las defunciones anuales estimadas a cada edad y $E_{x,t}$ es la población expuesta al riesgo de morir (población a mitad de año) a cada edad. El $\ln(m'_{x,t}) = a_x + b_x k'_t$ donde k'_t es el ajuste de k_t con este modelo y $\varepsilon'_{x,t}$ son los residuales obtenidos después del ajuste de k_t .

El criterio de minimización apropiado lo da el cálculo de la devianza,

$$\text{devianza}_t = 2 \sum_{x=0}^{\omega+} \left\{ D_{x,t} \ln \left(\frac{D_{x,t}}{D'_{x,t}} \right) - (D_{x,t} - D'_{x,t}) \right\}$$

La devianza es similar al cálculo de la prueba ji-cuadrada para la bondad de ajuste de $D'_{x,t}$ y además su cálculo es más sencillo, matemáticamente,

$$\chi^2 = \sum_{t=1}^n \sum_{x=0}^{\omega+} \left[\frac{(D_{x,t} - D'_{x,t})^2}{D'_{x,t}} \right]. \tag{4}$$

La ecuación [4] es utilizada para comparar la bondad de ajuste entre los modelos; el criterio de discriminación es: aquel modelo que presente la χ^2 menor es el que se ajusta mejor a los datos de las defunciones.

Adicionalmente, se calcula el coeficiente de determinación, R^2 , para medir el nivel de variabilidad explicado por la aproximación de primer orden definido como,

$$R^2 = 1 - \frac{\sum_{t=1}^n \sum_{x=0}^{\omega+} \varepsilon_{x,t}^2}{\sum_{t=1}^n \sum_{x=0}^{\omega+} [\ln(m_{x,t}) - a_x]^2}. \tag{5}$$

En virtud de lo anterior, en este trabajo se calculan y comparan las aproximaciones propuestas por Lee y Carter, y por Booth y sus colaboradores, que se muestran después.

Método de pronóstico

Después de haber elegido el modelo que se ajusta mejor a las defunciones observadas, el paso que sigue es realizar el pronóstico de los índices de la mortalidad. Esto se lleva a cabo modelando a k_t como un proceso estocástico de series de tiempo. En el artículo de Lee y Carter se propone que el índice sea modelado como una caminata aleatoria con deriva, aunque es necesario probar una buena cantidad de modelos para elegir el más indicado. En la actualidad este procedimiento ya no es complicado; en muchos paquetes estadísticos dichas pruebas se realizan automáticamente y en un tiempo computacional muy corto.

Si k_t se comporta de acuerdo con una caminata aleatoria con deriva, entonces se modela como sigue:

$$k_t = c + k_{t-1} + e_t, \quad [6]$$

donde c es el término constante de deriva (la tendencia “determinística” del modelo) y e_t es el error que se distribuye como una variable aleatoria normal con media 0 y varianza, σ^2 , constante. Si se utiliza el operador diferencia, ∇ , entonces el modelo [6] se puede simplificar de la siguiente manera: $\nabla k_t = k_t - k_{t-1} = c + e_t$. Observe que por las propiedades de los errores, ∇k_t se distribuye como una variable aleatoria normal con media c y varianza σ^2 .

Ahora bien, si k_t se comporta de acuerdo con un proceso autorregresivo integrado de promedios móviles (ARIMA) de orden (p, d, q) entonces se modela de la siguiente manera,³

$$\phi(B) = (1 - B)^d k_t = \theta(B) e_t, \quad [7]$$

donde $\phi(B)$ y $\theta(B)$ son polinomios de grados p y q , respectivamente;⁴ a B se le conoce como operador de retardo y cumple con que $B^j k_t = k_{t-j}$;

³ Antes de modelar es necesario identificar el modelo que mejor se ajusta a los datos. Para ello se ha de seguir una serie de pasos, como observar gráficamente la serie y calcular y analizar las funciones de autocorrelación total y parcial. Para más detalles véase Chatfield, 1995 y Box y Jenkins, 1976.

⁴ De aquí se sigue que $(1 - B)^d k_t = \phi_1(1 - B)^d k_{t-1} + \dots + \phi_p(1 - B)^d k_{t-p} + \zeta_t + \dots + \theta_q \zeta_{t-q}$ (Chatfield, 1995: 42 y 43).

e_t se define como un proceso puramente aleatorio⁵ con media cero y varianza constante σ_e^2 .

Box y Jenkins (1976: 91 y 92) recomiendan considerar la tendencia en el modelo [7] como una función determinística del tiempo. Esto se debe a que al incluir dicha constante automáticamente se permite una tendencia determinística polinomial.⁶ Por ejemplo, si en el modelo [7] se incluye un término constante, θ_0 , y $d=1$, es posible estimar la tendencia determinística en presencia de ruido no estacionario. Esto es importante, ya que como se verá más adelante, considerar al término de tendencia o deriva es de gran utilidad en el pronóstico de los indicadores demográficos. Así, el modelo [7] puede expresarse en general de la siguiente forma,

$$\phi(B)(1-B)^d k_t = \theta_0 + \theta(B)e_t. \tag{8}$$

Ahora bien, el término constante de la caminata aleatoria con deriva se estima de la siguiente manera: si \hat{e}_{t+1} son los residuales del modelo, entonces

$$\hat{c} = \frac{\sum_{t=2}^n \hat{e}_t}{n-1} = \frac{\sum_{t=2}^n \nabla k_t}{n-1} = \frac{\sum_{t=2}^n (k_t - k_{t-1})}{n-1}$$

la cual tiene un error estándar

$$S.E.(\hat{c}) = \frac{\hat{\sigma}_e}{\sqrt{\dim(\hat{e}_t)}}$$

donde $\dim(\hat{e}_t)$ es la dimensión del vector de residuales y

$$\hat{\sigma}_e = \sqrt{\sum_{t=2}^n (\hat{e}_t - \hat{c})^2 / (n-2)}$$

es la desviación estándar del ajuste.

La estimación de los parámetros del modelo ARIMA (p,d,q) no es sencilla. Para ello se requiere utilizar la función de verosimilitud condicional del proceso. Con dicha función se minimiza la suma de cuadrados de la función por medio de un algoritmo recursivo. La explicación de dicha metodología excede los propósitos de este trabajo,

⁵ Un proceso puramente aleatorio consiste en una secuencia de variables aleatorias, independientes e idénticamente distribuidas (Chatfield, 1995: 31).

⁶ Considere a $z_t = (L+Kl) + M \sum_{j=1}^{t-1} e_j + e_t$, entonces $\nabla z_t = K + (M-1)e_{t-1} + e_t$; por lo tanto $K = \theta_0$ y $M-1 = \theta_1$. Aquí se trata de un modelo ARIMA (0,1,1).

además de que es una técnica ampliamente utilizada de la que existe una amplia bibliografía,⁷ y varios paquetes estadísticos ya contienen librerías que automatizan los cálculos.

Así, el pronóstico del índice de mortalidad se realiza aplicando la ecuación [6] o la [7], según sea el caso, a partir del último valor de k_p para $h > 0$ periodos en el futuro, por ejemplo hasta obtener el valor de k_{t+h} . A partir del pronóstico del índice de mortalidad se calcula el valor de las tasas centrales de mortalidad futuras aplicando la ecuación [1] manteniendo constantes los valores estimados de a_x y b_x . Después se calculan las distintas funciones que componen la tabla de mortalidad para cada año, cuyo producto final es la esperanza de vida al nacimiento,⁸ aunque para fines del pronóstico de la población total es más importante el cálculo de las probabilidades de sobrevivencia, ya que con ellas y junto con el pronóstico de los nacimientos se estructura la matriz de Leslie, como se explicará más adelante.

Aplicación al caso de México

En este apartado se aplican las técnicas de estimación, ajuste y pronóstico explicadas en el apartado anterior relativo a los datos de las defunciones estimadas por INEGI, Conapo y El Colegio de México (2006). El periodo histórico comprendido en dicha estimación o conciliación comprende los años 1960 a 2005, aunque en este trabajo sólo se utilizan los años 1960 a 2000 con el fin de comparar los pronósticos obtenidos con las estimaciones de los años 2001, 2002, 2003, 2004 y 2005. En virtud de que el fin último de este pronóstico es elaborar uno para la población total, el método LC es aplicado para mujeres y hombres por separado.

Mortalidad femenina

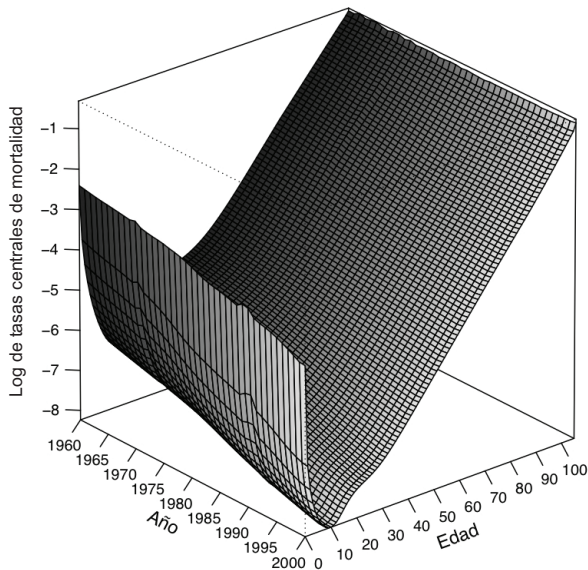
De acuerdo con las estimaciones de INEGI, Conapo y El Colegio de México la mortalidad femenina ha descendido durante los últimos 40 años del siglo XX. En 1960 hubo casi 254 mil defunciones de mujeres de todas las edades (48% del total de las muertes), y en el año 2000 fueron casi 215 mil

⁷ Para más información al respecto véase el capítulo 7 de Box y Jenkins, 1976.

⁸ El procedimiento que se utiliza en esta investigación sigue la forma de construir tablas de mortalidad propuesta por Preston *et al.*, 2000, capítulo 3.

GRÁFICA 1

Dinámica de las tasas centrales de mortalidad femenina, 1960-2000



FUENTE: Cálculos propios con base en INEGI, Conapo y El Colegio de México, 2006.

(45% del total de las muertes). Esto es, dado que en 1960 había 18.6 millones de mujeres y 49.7 millones en 2000, en 1960 morían 14 mujeres por cada mil, mientras que en 2000 murieron cuatro por cada mil. La mortalidad infantil de niñas pasó de 41 por cada mil nacimientos de ambos sexos a 11 por cada mil. Por otro lado, de 1960 a 2000 la esperanza de vida femenina al nacimiento pasó de 59.3 a 77 años, es decir, las mujeres ganaron casi 18 años más vida. La evolución de las tasas centrales de mortalidad a través del tiempo se muestra en la gráfica 1, en donde se observa que la menor probabilidad de fallecer en 1960 era alrededor de los 15 años, mientras que en 2000 se hallaba entre los 5 y los 10 años.

Ahora bien, a partir de la ecuación [1] se aplicaron los distintos métodos descritos arriba para estimar los parámetros a_x y b_x y la variable k_t . El mejor ajuste se obtuvo utilizando la DVS (con $R^2 = 0.9936$ y $\chi^2 = 17,183$). En la gráfica 2 se muestra cada uno de los parámetros obtenidos.

Una vez estimados los parámetros básicos, el paso que sigue es pronosticar el índice de mortalidad de acuerdo con el modelo de series

de tiempo que mejor se ajuste. Después de probar varios modelos, el que ofreció un mejor ajuste fue un ARIMA (1,1,1) con deriva, es decir,⁹

$$\begin{aligned}
 k_t &= \theta_0 + k_{t-1} + \phi_1(k_{t-1} - k_{t-2}) + e_t + \theta_1 e_{t-1} \\
 &= -0.286 + k_{t-1} + 0.874(k_{t-1} - k_{t-2}) + e_t - 0.64e_{t-1} \quad [9] \\
 &\quad (0.037) \quad (0.132) \quad (0.181)
 \end{aligned}$$

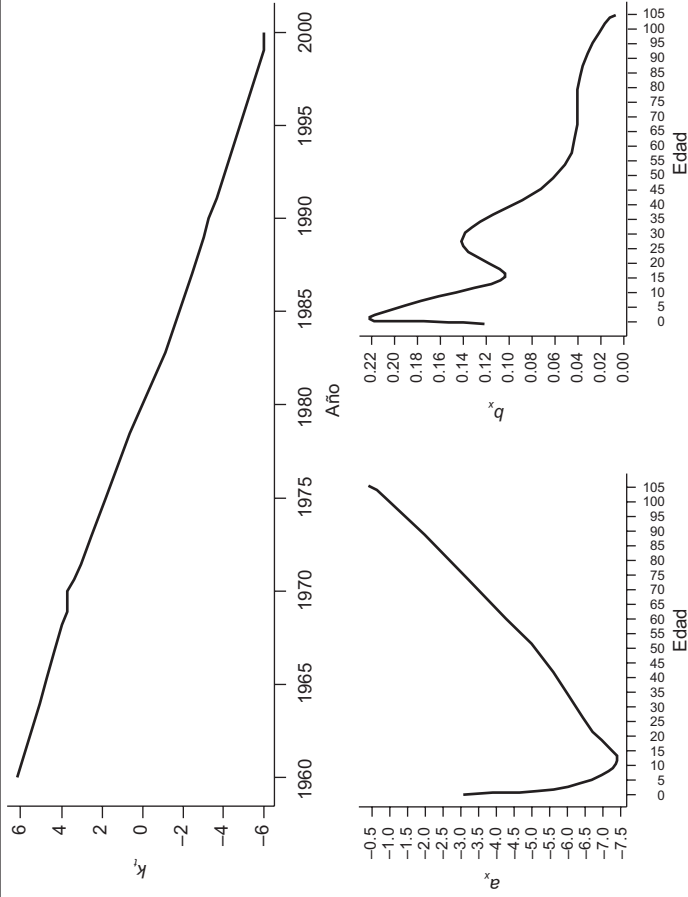
y la varianza de los residuales es $\sigma_e^2 = 0.00664$. Con estos valores es posible pronosticar, por métodos analíticos o por simulación, el índice de la mortalidad femenina hasta el año 2050. La gráfica 3 muestra los valores pronosticados y los intervalos de predicción de 67 y 95 por ciento.

Como se mencionó, el parámetro a_x en la ecuación [1] se refiere a la forma que tiene el logaritmo de la mortalidad respecto a la edad. Como se puede observar en la gráfica 2, este parámetro indica que el logaritmo de la mortalidad tiene forma de “U”, como usualmente se encuentra en cualquier población humana. El parámetro b_x se refiere a la intensidad de la mortalidad a cada edad. Contra lo que la intuición pudiera indicar, en la gráfica 2 a menor valor mayor intensidad de la mortalidad en esas edades. Esto se debe a que al multiplicar las matrices \mathbf{D} y \mathbf{V}^t , en la DVS, ambas pueden tener signo positivo o negativo, lo cual no cambia los resultados, sólo su presentación. En este caso ambas tienen signo negativo para seguir la forma en que Lee y Carter estimaron dichos parámetros. Por ello, en el caso de la mortalidad femenina en México, este parámetro indica que la mortalidad será más intensa en el primer año de vida, en los grupos etarios correspondientes a las mujeres adolescentes y jóvenes (menores de 30 años) y, lógicamente, mucho más intensa en los grupos de mujeres en edades avanzadas.

El parámetro de tendencia, llamado *índice de mortalidad*, denotado por k_t , es el que se extrapola hacia el futuro. Como antes se mencionó, para el caso de las mujeres el mejor modelo de pronóstico corresponde a un ARIMA (1,1,1), lo que indica que, una vez eliminada la tendencia de la serie, el índice de la mortalidad desestacionarizado es causado tanto por el índice de la mortalidad del periodo inmediato anterior como por la incertidumbre atribuible a estos dos periodos. Es decir, difícilmente la mortalidad femenina podrá cambiar su tendencia general, y el hecho de que se mantenga en descenso por dos periodos consecutivos refuerza el descenso del siguiente periodo, sujeto a la

⁹ Los números entre paréntesis son los errores estándar de cada parámetro estimado.

GRÁFICA 2
Parámetros estimados del modelo de Lee y Carter para la mortalidad femenina

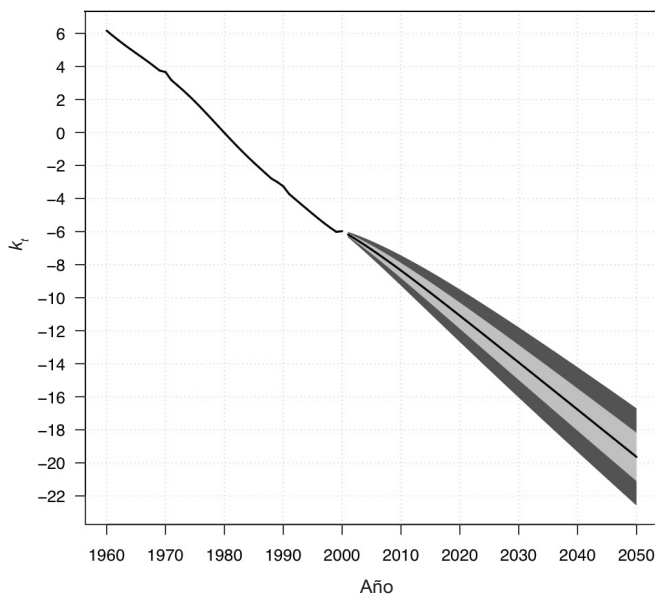


FUENTE: Cálculos propios con base en INEGI, Conapo y El Colegio de México, 2006.

variabilidad de la incertidumbre asociada. Los términos referentes a los procesos AR y MA tienen un efecto en el aumento o disminución del ritmo que lleva la tendencia de la mortalidad dependiendo del valor de los parámetros estimados. Por ejemplo, en el caso del modelo [9], el valor $\theta_0 = -0.286$ señala que en ausencia de incertidumbre y controlando la misma variable en el tiempo, la tendencia de la mortalidad femenina siempre disminuiría a ese ritmo, y no es factible que cambie de signo dado el valor de su error estándar. Ahora, el parámetro del término autorregresivo $\phi_1 = 0.874$ revela una aceleración de la tendencia que lleva el índice, y de hecho, muestra que el valor obtenido en el periodo previo al pronosticado tiene más peso que el de dos periodos anteriores. Además muestra que el periodo inmediato anterior acelera la tendencia y el de dos periodos previos la frena. El parámetro del proceso MA, $\theta_1 = -0.64$, aunado al valor intrínseco de 1 del térmi-

GRÁFICA 3

Pronóstico del índice de la mortalidad femenina, 2000-2050.
Las franjas claras representan el intervalo de 67% de predicción,
las oscuras el intervalo de 95% de predicción



FUENTE: Cálculos propios con base en INEGI, Conapo y El Colegio de México, 2006.

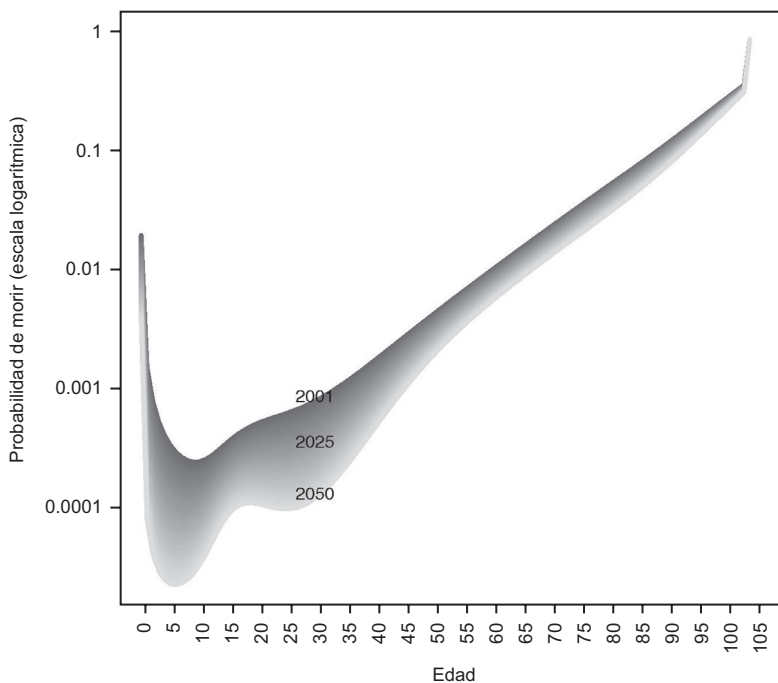
no de error e_t indica que, dado su rango de variabilidad, este proceso frena la tendencia de la mortalidad –es decir, frena al índice de la mortalidad– en la mayoría de los casos, ya que todo depende del juego de valores que pueden obtener dichas variables aleatorias. Además, el que el proceso de promedios móviles sea de orden uno significa que la tendencia futura no será lineal en el primer periodo pronosticado, originando así una ligera curvatura. Por ello se puede observar en la gráfica 3 que el cambio de tendencia del índice de mortalidad en el periodo 1999-2000 ocasiona que este índice no descienda linealmente a partir del primer periodo pronosticado, 2000-2001, sino que haga una pequeña curva para después comenzar a descender linealmente.

Una vez pronosticado el índice de mortalidad, el paso siguiente es calcular las funciones que conforman la tabla de mortalidad para el pronóstico medio, así como para uno de los intervalos de predicción. Como en este trabajo se utiliza la desagregación por edad simple los cálculos son muy sencillos (véase Preston *et al.*, 2000). En la gráfica 4 se presenta el pronóstico medio de la probabilidad de morir, q_x , de 2001 a 2050, y en la gráfica 5 se muestra el pronóstico de esta misma función para algunos años seleccionados, incluyendo sus respectivos intervalos de predicción.

En la gráfica 4 se puede observar que se espera que la probabilidad de fallecer de las mujeres siga disminuyendo, aunque la forma que adquiere el patrón de esta disminución se aparta un poco de lo que se ha observado en otros contextos. Es decir, la forma del patrón de la probabilidad de morir en México muestra un desfase en las mujeres entre 17 y 40 años. Cabe advertir que en el año 2050 se incrementa la probabilidad de fallecer de las mujeres entre 18 y 28 años respecto de las que están entre 29 y 40 años. Usualmente se esperaría que conforme avance la edad también se incremente la probabilidad de fallecer. Una posible explicación de esto, y que de hecho es materia de políticas públicas, es que en el futuro habrá de incrementarse la “masculinización” de las mujeres en las edades jóvenes. Al utilizar el término “masculinización” se hace referencia a la exposición de los individuos a riesgos que exponen su integridad física con el fin de reforzar un conjunto de atributos dentro de la sociedad. En el presente caso es posible que las mujeres en dichas edades se encuentren en un proceso de transformación de sus roles tradicionales que las lleva a imitar ciertas conductas –fumar y beber en exceso, uso de drogas, ejercer violencia, etc.– que menguan su calidad de vida y por ende incrementan su probabilidad de fallecer.

GRÁFICA 4

Pronóstico medio de la probabilidad de fallecer. Mujeres, 2001-2050



FUENTE: Cálculos propios con base en INEGI, Conapo y El Colegio de México, 2006.

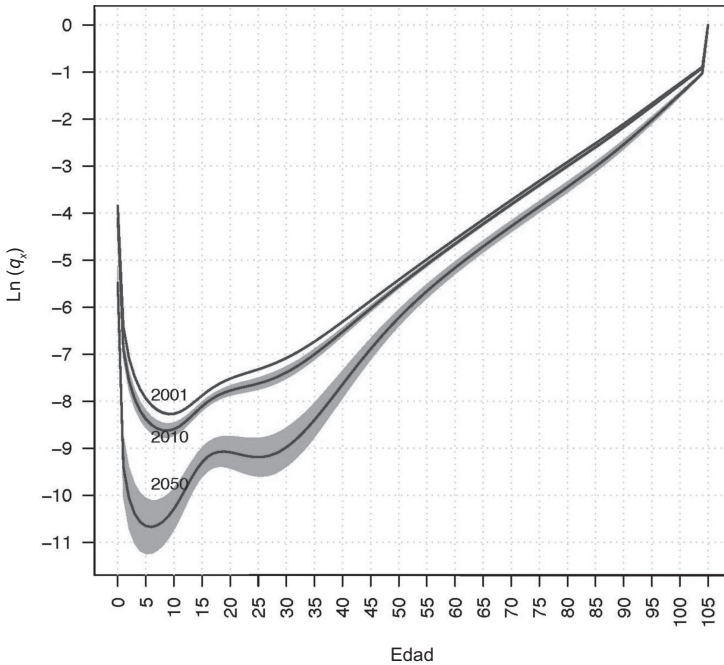
En la gráfica 5 se muestra el patrón por edad de la probabilidad de fallecer para tres años seleccionados, cada uno de los cuales muestra el intervalo de 95% de predicción. Observe que conforme se incrementa el horizonte de pronóstico, también lo hace la incertidumbre.

Así como la incertidumbre va aumentando con el paso del tiempo, también la edad representa un factor que la incrementa. Como se puede observar en la gráfica 6, a menor edad, mayor incertidumbre. En efecto, se espera que las ganancias en esperanza de vida para las edades superiores a 65 años sean más ciertas que las de las primeras edades. Lo anterior se debe a que el incremento en la esperanza de vida se debe principalmente a la disminución de la mortalidad en las primeras edades.

Otro aspecto interesante de la gráfica 6 es que muestra que hasta poco antes de 1990 la esperanza de vida al nacimiento era menor que

GRÁFICA 5

Pronóstico de la probabilidad de fallecer con intervalos de predicción de 95%. Mujeres, 2001, 2010, 2050



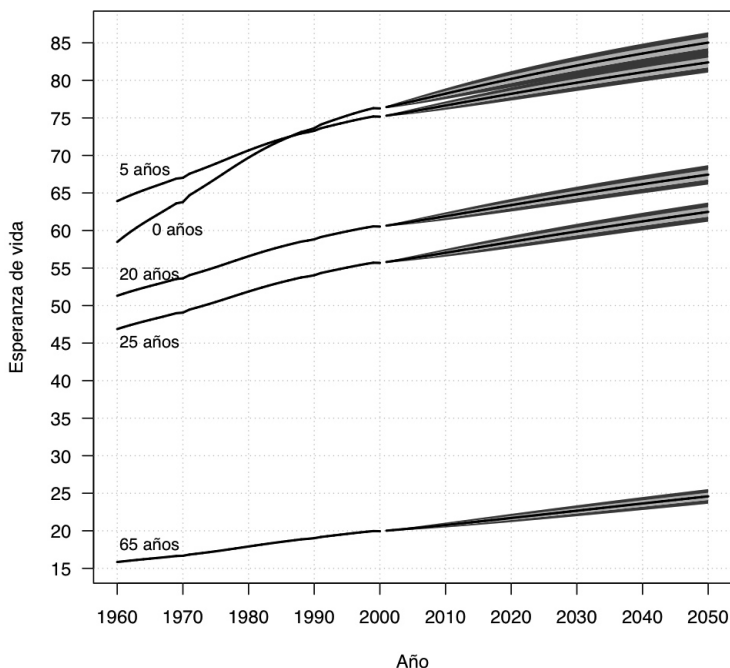
FUENTE: Cálculos propios con base en INEGI, Conapo y El Colegio de México, 2006.

la esperanza de vida de los niños de cinco años. Después hay un cruce en donde la primera supera a la segunda y se mantiene así a lo largo del horizonte de proyección; aunque se puede observar un cruce, casi tangencial, entre los intervalos de 95% de predicción, lo que significa que existe una posibilidad de que la esperanza de vida entre estas dos edades se vuelva a cruzar en el futuro; sin embargo tal posibilidad es mínima.

Ahora bien, al comparar el pronóstico aquí realizado con la última proyección del Conapo, se observa en la gráfica 7 que las dos variantes de proyección determinística caen dentro del intervalo de predicción. Comúnmente el Conapo formula una sola hipótesis para la tendencia de la mortalidad, sin embargo en el ejercicio proyectivo de 2005 considera dos opciones para extrapolar la dinámica de la mortalidad: la

GRÁFICA 6

Esperanza de vida con intervalos de predicción de 67% (franja clara) y de 95% (franja oscura). Mujeres, varias edades



FUENTE: Cálculos propios con base en INEGI, Conapo y El Colegio de México, 2006.

primera consiste en disminuir las tasas específicas por edad de menores de 65 años a lo largo de los 45 años de la proyección y posponer hasta 2015 el descenso de las tasas de los adultos mayores; la segunda, llamada hipótesis alternativa, fue realizada para analizar la sensibilidad del riesgo de morir durante la proyección; en dicha hipótesis se retiene el descenso previsto para la mortalidad antes de los 65 años de edad, pero también se incluye el descenso en la senectud a lo largo del horizonte de proyección, es decir, no lo posterga diez años como en el primer escenario.

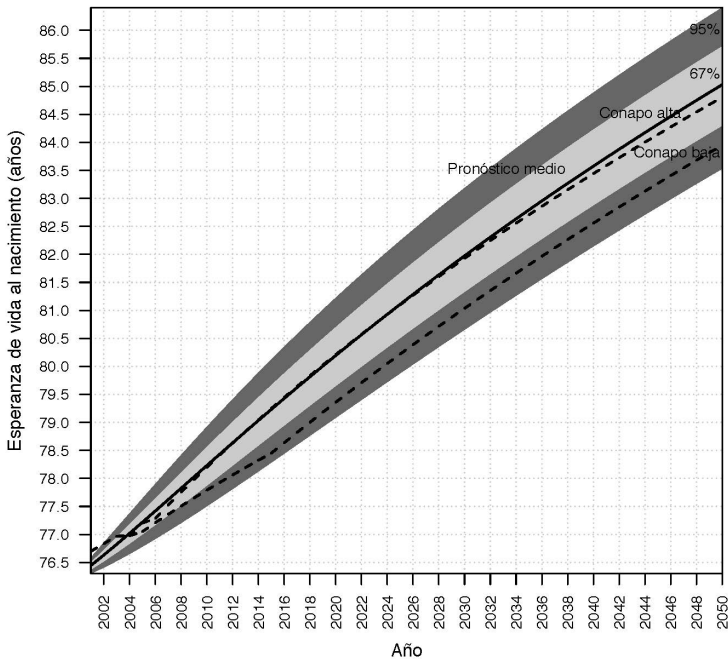
Cabe mencionar que el método de extrapolación empleado por el Conapo es muy similar al que se utiliza en este trabajo, $\ln(q_x) = \alpha_x + \beta_x B_t$, aunque no considera la aleatoriedad del modelo ni extrapola considerando la serie histórica. Este modelo se utiliza para determinar la es-

estructura por edad de la mortalidad. Después se emplea una función logística para modelar la dinámica de las esperanzas de vida parcial de 0 a 64 años y de 5 a 64 años.

De esta manera, en la gráfica 7 se puede observar que para el caso de las mujeres, la variante que aquí se denomina “alta” –y que de hecho es la que se considera dentro del escenario “probable”, ya que se encuentra determinada por supuestos de mortalidad baja– es muy similar al pronóstico estocástico medio. La variante alternativa no cae dentro del intervalo de 65% pero sí dentro del de 95%. Esto no significa que exista una probabilidad alta de que ocurra cualquiera de las dos variantes, sino que constituyen tan sólo una posibilidad dentro del intervalo de predicción y, dado que no fueron construidas considerando la aleatoriedad asociada, no se puede decir que el área comprendida

GRÁFICA 7

Comparativo entre la esperanza de vida pronosticada y las proyecciones del Conapo. Mujeres, 2001-2050



FUENTE: Cálculos propios con base en INEGI, Conapo y El Colegio de México, 2006.

entre ellas representa un espacio de posibilidades de ocurrencia, sino que son dos eventos aislados o, en otras palabras, son dos conjuntos de medida cero.

Al comparar el pronóstico estocástico medio con la variante de proyección considerada por el Conapo como probable (cuadro 1), se observa que las diferencias son muy pequeñas, menores a un mes en la mayoría de los casos, ya que se incrementan con el tiempo; para 2050 tal diferencia es de tres meses aproximadamente. En casi todo el horizonte de proyección el pronóstico es ligeramente superior a la proyección.

La División de Población de Naciones Unidas (ONU) proyectó una sola variante de mortalidad en su ejercicio prospectivo de 2006. Al comparar el pronóstico con la proyección de la ONU se observa que las diferencias son más grandes que con las del Conapo. En efecto, la proyección de la ONU es mayor que el pronóstico durante casi la primera mitad del periodo pronosticado: para 2005 la diferencia es de alrededor de dos meses, en 2010 es de cuatro meses, en 2015 de cinco meses; entre 2020 y 2030 las respectivas curvas de la esperanza de vida se intersectan y el pronóstico se vuelve mayor que la proyección, llegando a tener una diferencia de un año y medio para 2050.

Por último, en la gráfica 8 se muestra un diagrama de superficies estructurado para representar el diagrama de Lexis de la evolución histórica de la mortalidad, así como el pronóstico mediano calculado en este trabajo. En el eje horizontal se encuentran los años desde 1960 hasta 2050 y en el eje vertical las edades. Las superficies representan la intensidad de la mortalidad, medida a través de las tasas centrales estimadas y pronosticadas. Como se puede advertir, realizar una gráfica de esta índole con intervalos de confianza es muy complicado, por lo que para ejemplificar la evolución de la mortalidad histórica y la probable se empleó el pronóstico mediano.

Este tipo de gráficas es muy interesante porque se pueden hacer análisis por edad, periodo o cohorte. Por ejemplo, en 1960 morían entre 1 y 5 por cada 100 mujeres de entre 50 y 55 años, pero es probable que para el año 2050 disminuya la mortalidad en estas edades de tal forma que mueran entre 1 y 5 por cada mil. Un análisis de periodo es observar la estructura de la mortalidad para determinado grupo de años; por ejemplo, observe la diferencia de la estructura de la mortalidad entre los años 1975-1980 y 2025-2030, en la cual claramente se ve que con excepción del grupo entre 60 y 65 años, el resto presenta una reducción en su mortalidad.

CUADRO 1
Comparativo de esperanzas de vida al nacimiento en mujeres. Varios autores

Año	Pronóstico estocástico					Conapo 2005			NU 2006	Diferencias**	
	Baja		Alta			Alta***	Baja***	Única	Conapo alta	NU	
	95%	67%	Mediana	67%	95%						
2000*	76.26	76.26	76.26	76.26	76.26	76.47	76.47	76.07	-0.21	0.19	
2001	76.30	76.37	76.44	76.51	76.58	76.70	76.70	n.d.	-0.25	n.d.	
2002	76.41	76.52	76.63	76.74	76.85	76.83	76.83	n.d.	-0.20	n.d.	
2003	76.53	76.68	76.82	76.97	77.11	76.97	76.97	n.d.	-0.15	n.d.	
2004	76.65	76.84	77.02	77.20	77.37	76.98	76.98	n.d.	0.04	n.d.	
2005	76.78	77.00	77.22	77.43	77.64	77.05	77.20	77.36	0.17	-0.14	
2010	77.50	77.87	78.23	78.58	78.92	77.79	78.20	78.61	0.44	-0.38	
2015	78.29	78.77	79.23	79.68	80.12	78.45	79.25	79.67	0.78	-0.44	
2020	79.08	79.65	80.19	80.71	81.22	79.36	80.21	80.55	0.84	-0.36	
2025	79.88	80.51	81.11	81.68	82.24	80.22	81.10	81.29	0.89	-0.18	
2030	80.66	81.33	81.97	82.59	83.18	81.04	81.93	81.91	0.94	0.06	
2040	82.14	82.88	83.58	84.24	84.89	82.56	83.44	82.84	1.02	0.74	
2050	83.53	84.30	85.03	85.73	86.40	83.95	84.81	83.47	1.08	1.56	

* Estimaciones.

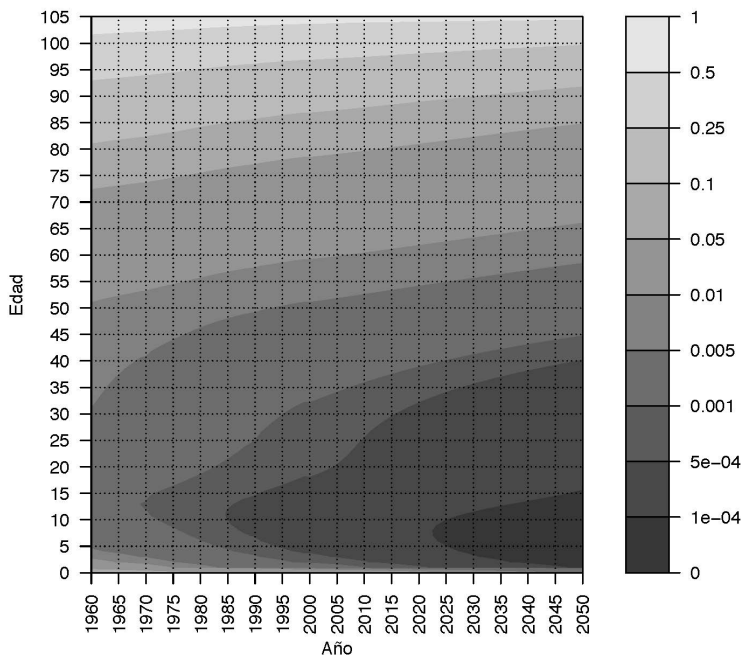
** Diferencias respecto al pronóstico mediano.

*** Los términos "alta" y "baja" se refieren al valor de la variable, mas no al supuesto de mortalidad alta o baja.
n.d. = datos no disponibles.

FUENTE: Cálculos propios y proyecciones de población del Conapo, 2005 y de Naciones Unidas, 2006.

GRÁFICA 8

Superficie de Lexis para la mortalidad femenina, 1960-2050



FUENTE: Cálculos propios con base en INEGI, Conapo y El Colegio de México, 2006.

Finalmente, para hacer un análisis de cohorte es necesario trazar –en este caso hay que imaginar– líneas de 45 grados comenzando en cualquier año calendario; suponga que se comienza en el año 1990, es decir, se considera a las mexicanas que nacieron en el año 1990. En ese año morían entre 10 y 50 niñas por cada mil niñas entre 0 y 1 años de edad –nótese que no se trata de la mortalidad infantil, ya que no se consideran los nacimientos–; luego, en el año 2000 cuando dichas niñas llegaron a los 10 años, morían entre 1 y 5 por cada 10 000 niñas entre 10 y 11 años. Dicha tasa de mortalidad se mantendrá hasta que esas niñas lleguen a tener entre 35 y 40 años. En cambio, las mujeres que nacieron en 1965 y que tuvieron entre 35 y 40 años entre los años 2000 y 2005 tuvieron una tasa de mortalidad de entre 10 y 50 por cada 10 000 mujeres entre esas edades.

Mortalidad masculina

De 1960 al 2000 la mortalidad masculina descendió, aunque en términos absolutos se ha mantenido. En 1960 murieron 274 mil hombres (52% del total de defunciones), mientras que en el año 2000 murieron 265 mil hombres (55% del total). En 1960 había 19 millones de varones, mientras que en el año 2000 eran 49 millones. Esto significa que en 1960 morían 15 hombres por cada mil y en 2000 morían 5 por cada mil. La mortalidad infantil de varones pasó de 50 defunciones de menores de un año por cada mil nacimientos a 11 por cada mil nacimientos.

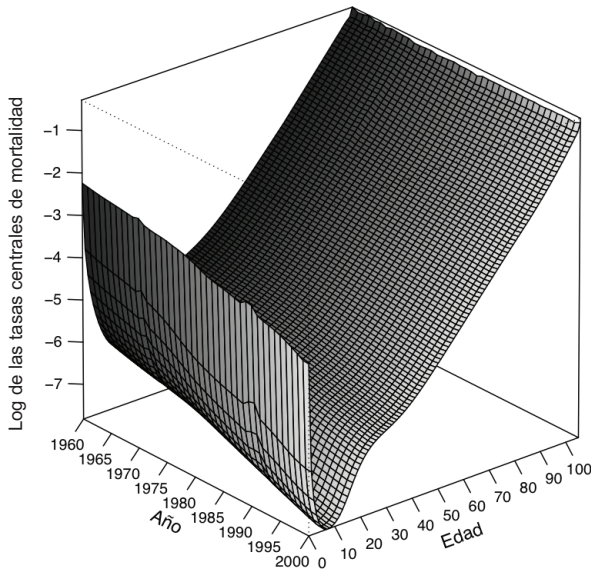
Biológicamente los hombres tienden a vivir menos que las mujeres. Existen muchas explicaciones de este fenómeno, que se ha encontrado en distintas sociedades a lo largo de la historia. En primer lugar se encuentran las explicaciones biológicas que establecen que es una condición de la naturaleza. Otras explicaciones, que principalmente provienen de la sociología, argumentan que dado que los hombres se exponen a una gran cantidad de riesgos a lo largo de su vida y principalmente durante el periodo en que se construye la masculinidad, los impactos violentos sobre su cuerpo se ven reflejados en la disminución de los años que viven.

En demografía, cuando la mortalidad y la fecundidad son altas, es decir, en la etapa previa a la transición demográfica y durante la primera etapa de ésta, la esperanza de vida entre los dos sexos es muy similar, aunque ligeramente mayor la femenina. Conforme la mortalidad y la fecundidad disminuyen, la diferencia tiende a ampliarse (Caldwell, 2006: 157-180). Aún hoy en día existe un debate inconcluso sobre la duración de la vida y si la diferenciación por sexos puede reducirse (véase Oeppen y Vaupel, 2002 y Vaupel, 2004).

Respecto al caso mexicano, la esperanza de vida de los varones se ha incrementado 15 años de 1960 a 2000, mientras que la de las mujeres aumentó 17 años. En 1960 se esperaba que los varones recién nacidos vivieran hasta los 56 años en promedio; para el año 2000 se esperaba que vivieran 76 años. En la gráfica 9 se muestra la dinámica de las tasas centrales de mortalidad masculina –de 1960 a 2000– del patrón por edad de la probabilidad de fallecer q_x en escala logarítmica. En esta gráfica se observa que la disminución de la mortalidad entre los varones se debe principalmente a la disminución de la mortalidad en los niños, mientras que la disminución en los adolescentes y adultos, si bien es considerable, no representa la principal fuerza que origina dicha disminución.

GRÁFICA 9

Dinámica de las tasas centrales de mortalidad masculina, 1960-2000



FUENTE: Cálculos propios con base en INEGI, Conapo y El Colegio de México, 2006.

Ahora bien, a partir de la ecuación [1] se estimaron los parámetros a_x , b_x y la variable k_t . El mejor ajuste se obtuvo utilizando la DVS ($R^2 = 0.986$ y $\chi^2 = 14,956$). En la gráfica 10 se muestra cada uno de los parámetros obtenidos.

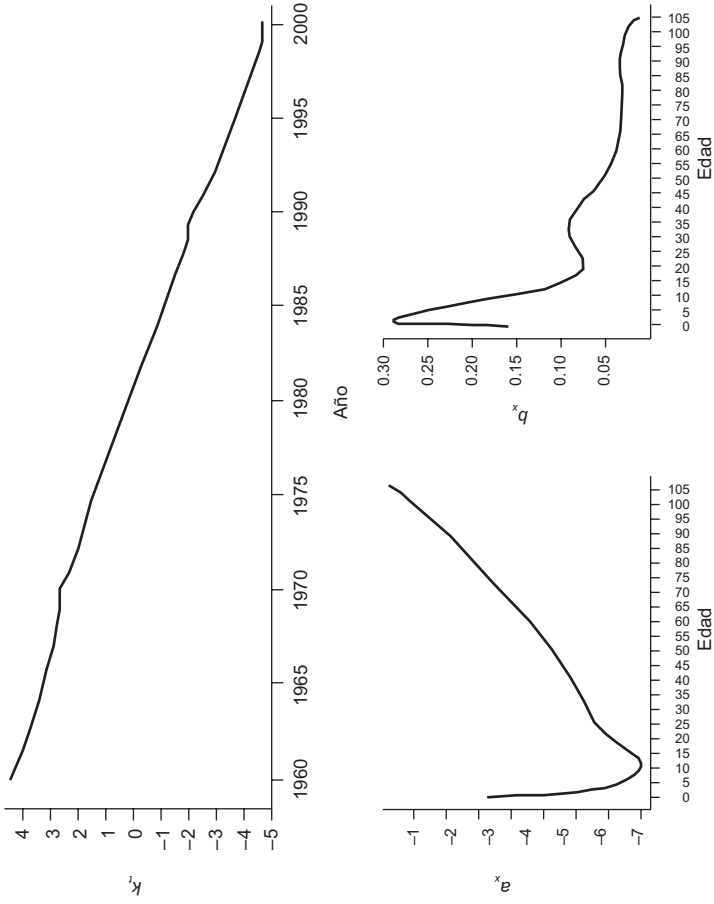
Una vez estimados los parámetros básicos, el paso que sigue es pronosticar el índice de mortalidad de acuerdo con el modelo de series de tiempo que mejor se ajuste. Después de probar distintos modelos, el que ofreció un mejor ajuste fue un modelo ARIMA (0,1,0) con deriva o una caminata aleatoria con deriva, es decir,¹⁰

$$\begin{aligned}
 k_t &= \theta_0 + k_{t-1} + e_t \\
 &= -0.2272 + k_{t-1} + e_t \\
 &\quad (0.0129)
 \end{aligned}
 \tag{10}$$

y la desviación estándar de los residuales es $\sigma_e = 0.08186$. Con estos valores es posible pronosticar, por métodos analíticos o por simulación,

¹⁰ Los números entre paréntesis son los errores estándar de cada parámetro estimado.

GRÁFICA 10
Parámetros estimados del modelo de Lee y Carter para la mortalidad masculina



FUENTE: Cálculos propios con base en INEGI, Conapo y El Colegio de México, 2006.

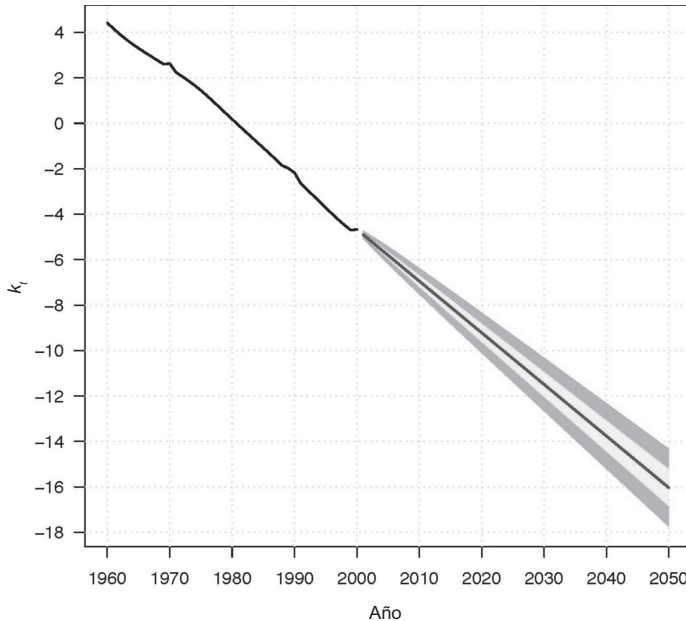
el índice de la mortalidad masculina hasta el año 2050. La gráfica 11 muestra los valores pronosticados y los intervalos de predicción de 67 y 95 por ciento.

Cabe mencionar que, a diferencia del caso de las mujeres, en la gráfica 10 el parámetro de intensidad, b_x , muestra un efecto menos fuerte de lo que fue denominado “masculinización”. Observe que la intensidad de la mortalidad en los grupos etarios de adolescentes y de adultos jóvenes no es mayor que la estimada para los grupos de adultos maduros –alrededor de 40 años–, como se observó en el caso de las mujeres. Esto sólo muestra que el patrón de “masculinización” en los hombres está muy consolidado y no se espera que haya cambios en el futuro; el efecto de los riesgos a los que se sujetan los individuos en las edades jóvenes determina o se intensifica respecto al

GRÁFICA 11

Pronóstico del índice de la mortalidad masculina, 2000-2050.

Las franjas claras representan el intervalo de 67% de predicción, las oscuras el intervalo de 95% de predicción



FUENTE: Cálculos propios con base en INEGI, Conapo y El Colegio de México, 2006.

de los de las edades maduras. Lo anterior también es materia de políticas públicas, ya que en lugar de incentivar a los jóvenes mexicanos para que lleven una vida libre de riesgos innecesarios, se observa que el patrón predominante de reafirmación de valores seguirá siendo el masculino. Sería importante hacer un esfuerzo para tratar de revertir tales patrones.

Respecto al parámetro que mide la tendencia de la mortalidad, el índice de mortalidad k_t , se observa una tendencia casi lineal a lo largo del tiempo. De acuerdo con el modelo de pronóstico, el índice desestacionarizado dependerá únicamente de la media histórica y de un factor inherente de la incertidumbre. Es decir, al ser una caminata aleatoria con deriva, el valor futuro del indicador dependerá solamente del valor que tuvo un periodo antes, de la tendencia histórica y de un factor de incertidumbre asociado al periodo pronosticado. Observe que el valor del parámetro de tendencia θ_0 es negativo, por lo que la mortalidad masculina continuará disminuyendo. Esto indica que al no tener ningún factor que acelere o disminuya la tendencia histórica, la mortalidad masculina seguirá al mismo ritmo promedio de descenso que ha llevado durante los 40 años previos al inicio del pronóstico.

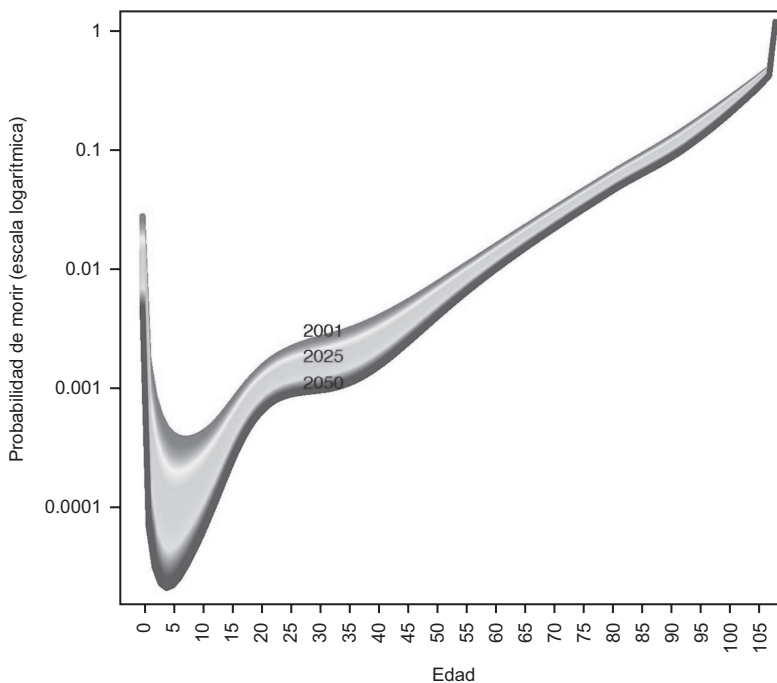
En las gráficas 12 y 13 se muestran el pronóstico medio de la probabilidad de morir, q_x , de 2001 a 2050, y el pronóstico de esta misma función para algunos años seleccionados incluyendo sus respectivos intervalos de predicción. En la gráfica 12 se puede observar que se espera que la probabilidad de fallecer siga disminuyendo en los hombres. Como ya se mencionó, la disminución de la mortalidad masculina se encuentra principalmente influida por la disminución de la mortalidad en los primeros años.

En la gráfica 13 se presenta el patrón por edad de la probabilidad de fallecer para tres años seleccionados, cada uno de los cuales muestra el intervalo de 95% de predicción. Observe que conforme se incrementa el horizonte de pronóstico, también lo hace la incertidumbre, la cual se manifiesta principalmente en las edades que disminuyen más su probabilidad de fallecer.

Así como la incertidumbre se incrementa con el paso del tiempo, también la edad representa un factor que coadyuva a su incremento en el sentido de que, como se puede observar en la gráfica 14, a menor edad se observa una mayor incertidumbre. Es decir, se espera que las ganancias en esperanza de vida para las edades superiores a 65 años sean más ciertas que las de las primeras edades. Lo anterior implica

GRÁFICA 12

Pronóstico medio de la probabilidad de fallecer. Hombres, 2001-2050



FUENTE: Cálculos propios con base en INEGI, Conapo y El Colegio de México, 2006.

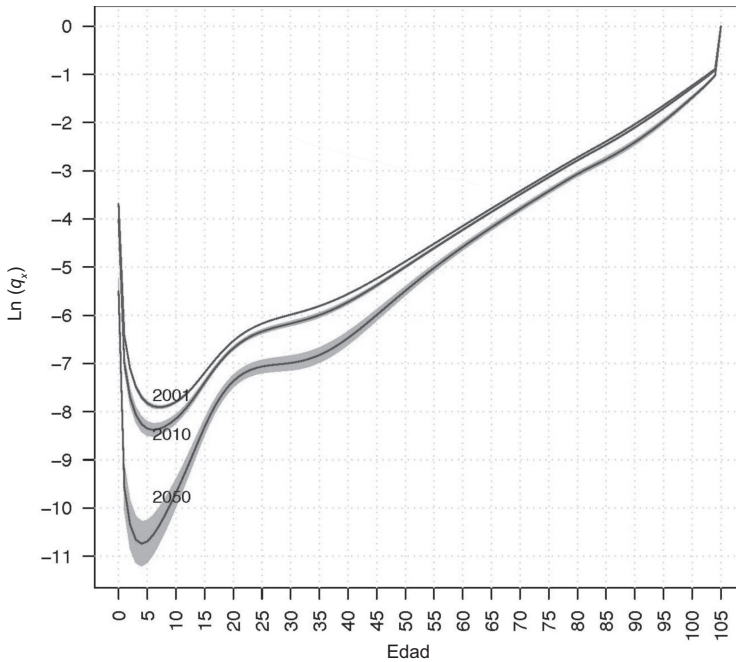
que aún hace falta realizar esfuerzos más significativos para disminuir la mortalidad infantil masculina, de tal suerte que para el futuro los esfuerzos se concentren en las edades adultas.

Por otro lado, en la gráfica 14 se muestra que hasta poco antes de 1990 la esperanza de vida al nacimiento en los varones era menor que la esperanza de vida de los niños de cinco años. Después hay un cruce en donde la primera supera a la segunda y se mantiene así a lo largo del horizonte de proyección. A diferencia del caso femenino, no se puede observar un cruce entre los intervalos de 95% de predicción.

Comparando los dos escenarios proyectados por el Conapo con el pronóstico medio y sus intervalos de predicción, en la gráfica 15 se muestra que la variante alternativa –variante baja– se acerca más al pronóstico medio, y la variante considerada como “probable” no sólo

GRÁFICA 13

Pronóstico de la probabilidad de fallecer con intervalos de predicción de 95%. Hombres, 2001, 2010, 2050



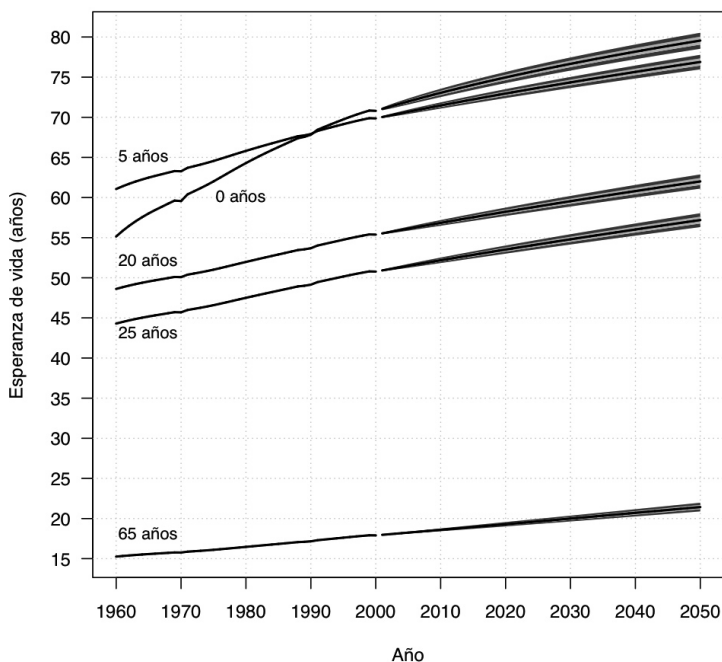
FUENTE: Cálculos propios con base en INEGI, Conapo y El Colegio de México, 2006.

está en el límite superior del intervalo de predicción de 95%, sino que lo sobrepasa después de 2025. Esto muestra que es posible que las proyecciones oficiales estén sobrevaluando la disminución de la mortalidad masculina.

En el cuadro 2 se muestran los valores de la esperanza de vida al nacimiento para distintos años. En las primeras cinco columnas se presentan el pronóstico medio y los intervalos de predicción calculados con el modelo de LC; en las siguientes dos columnas se muestran las variantes proyectadas por el Conapo, en la columna siguiente está la proyección única elaborada por la ONU, y en las dos últimas columnas se muestran, primero, las diferencias entre el pronóstico medio y la variante “probable”, y al final las diferencias entre el pronóstico medio y la proyección de la ONU. Comparando el pronóstico estocástico con la variante del Conapo, se puede observar que a lo largo del horizonte

GRÁFICA 14

Esperanza de vida con intervalos de predicción de 67% (franjas claras) y de 95% (franjas oscuras). Hombres, varias edades



FUENTE: Cálculos propios con base en INEGI, Conapo y El Colegio de México, 2006.

de pronóstico la proyección se encuentra sobre el valor pronosticado. En el mejor de los casos la diferencia es de tres meses, pero en general la proyección se encuentra por arriba del pronóstico en más de cinco meses. Para 2050 la diferencia es de poco más de un año. En cambio, al comparar el pronóstico estocástico con la proyección de la ONU se observa que esta última también se encuentra por arriba del pronóstico medio en casi todo tiempo –las diferencias llegan a ser de más de nueve meses para los años 2015 y 2020–, salvo en los cinco últimos años, en donde las esperanzas de vida proyectadas son menores que las pronosticadas –en más de siete meses para 2050.

Ahora bien, en la gráfica 16 se muestra el diagrama de Lexis correspondiente a las tasas de mortalidad masculinas históricas y pronosticadas; estas últimas se refieren al pronóstico mediano. Observe que a diferencia del diagrama de las tasas de mortalidad femeninas, éste

CUADRO 2
Comparativo de esperanza de vida al nacimiento en hombres. Varios autores

Año	Pronóstico estocástico					Conapo 2005		NU 2006	Diferencias**	
	Baja		Mediana	Alta		Alta***	Baja***	Única	Conapo	
	95 %	67 %		67 %	95 %				Alta	NU
2000*	70.80	70.80	70.80	70.80	70.80	71.30	71.30	71.26	-0.50	-0.46
2001	70.87	70.96	71.05	71.13	71.22	71.59	71.59	n.d.	-0.54	n.d.
2002	71.04	71.16	71.28	71.40	71.52	71.78	71.78	n.d.	-0.49	n.d.
2003	71.22	71.37	71.52	71.66	71.81	71.98	71.98	n.d.	-0.47	n.d.
2004	71.40	71.58	71.75	71.91	72.08	72.05	72.05	n.d.	-0.30	n.d.
2005	71.59	71.78	71.97	72.16	72.34	72.21	72.26	72.43	-0.24	-0.46
2010	72.52	72.79	73.05	73.30	73.55	73.08	73.41	73.74	-0.03	-0.69
2015	73.41	73.73	74.04	74.34	74.64	73.87	74.52	74.85	0.17	-0.81
2020	74.26	74.62	74.97	75.31	75.65	74.84	75.54	75.78	0.13	-0.81
2025	75.06	75.46	75.84	76.21	76.58	75.76	76.51	76.57	0.08	-0.73
2030	75.82	76.25	76.66	77.06	77.45	76.65	77.41	77.23	0.01	-0.57
2040	77.23	77.71	78.17	78.62	79.06	78.31	79.08	78.22	-0.14	-0.05
2050	78.53	79.06	79.56	80.05	80.53	79.86	80.60	78.90	-0.30	0.66

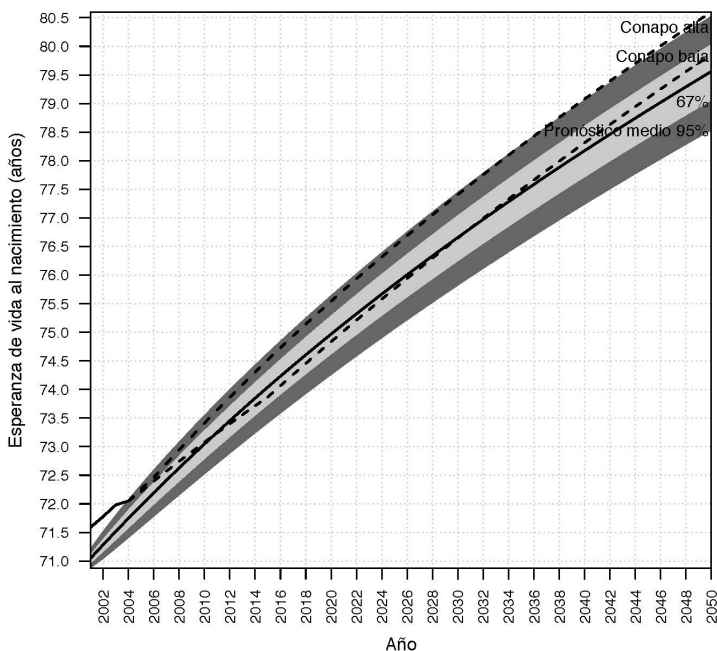
* Estimaciones.

** Diferencias respecto al pronóstico mediano.

*** Los términos "alta" y "baja" se refieren al valor de la variable, mas no al supuesto de mortalidad alta o baja.
n.d. = datos no disponibles.

FUENTE: Cálculos propios y proyecciones de población del Conapo, 2005 y de Naciones Unidas, 2006.

GRÁFICA 15
Comparativo entre las esperanzas de vida pronosticadas
y las proyecciones del Conapo. Hombres, 2001-2050

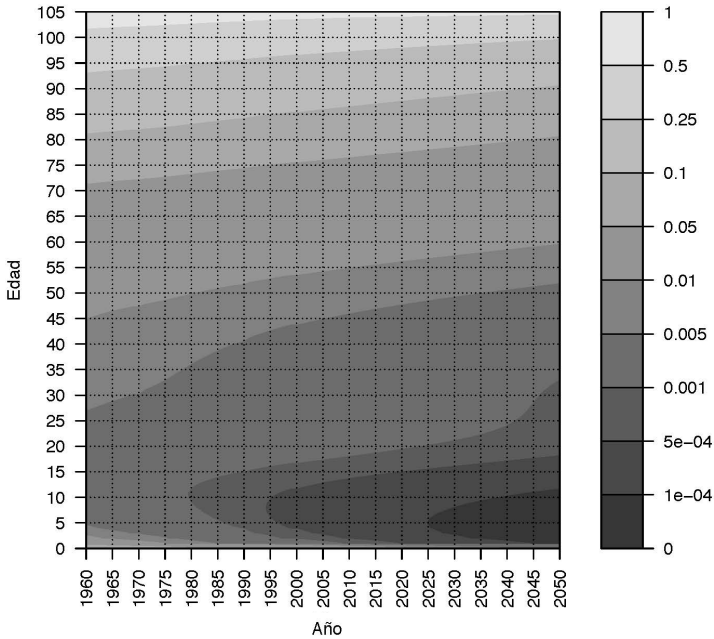


FUENTE: Cálculos propios con base en INEGI, Conapo y El Colegio de México, 2006.

revela que se espera una menor pérdida de mortalidad en el futuro. Las tasas menores de 1 por cada mil difícilmente rebasarán a los mayores de 25 años, y únicamente lo harán al final del horizonte de proyección. Un caso interesante que se puede observar en la gráfica mencionada es que la cohorte de los nacidos en 1960 mantiene una tasa de mortalidad de entre 1 y 5 por cada mil casi toda su vida, y sólo comienza a incrementarse después del año 2005, cuando los miembros de esa cohorte pasan de 45 años.

Por último, en la gráfica 17 se presentan los pronósticos de la esperanza de vida al nacimiento para mujeres y hombres con sus respectivos intervalos de predicción. Se advierte que el patrón biológico que diferencia a ambos sexos se mantiene con el tiempo e incluso no permite que se crucen los respectivos intervalos de predicción. Asimismo en esta gráfica se muestra más claramente que la incertidumbre

GRÁFICA 16
Superficie de Lexis para la mortalidad masculina, 1960-2050



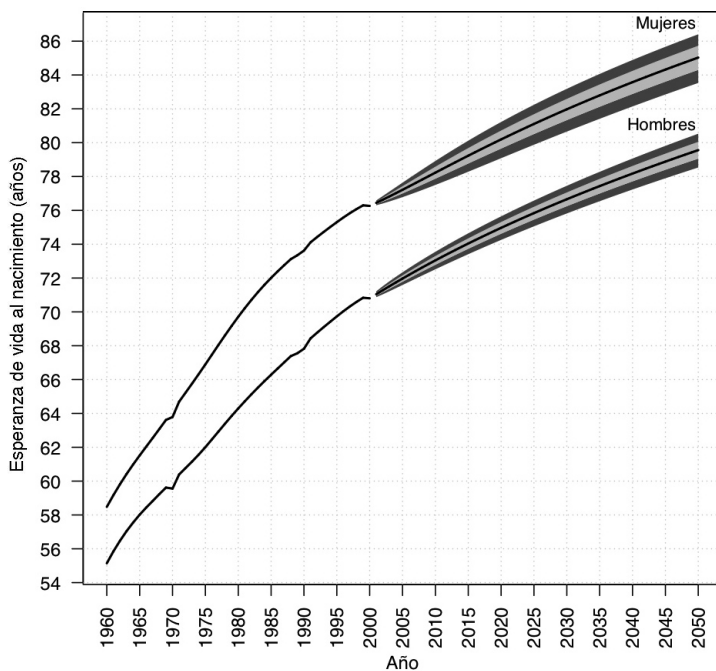
FUENTE: Cálculos propios con base en INEGI, Conapo y El Colegio de México, 2006.

asociada al pronóstico de la esperanza de vida al nacer de las mujeres es mayor que la de los hombres. Esto en virtud de que los cambios en la mortalidad femenina han mostrado mayor dinamismo a lo largo de la historia.

Conclusiones

Al analizar los pronósticos de la esperanza de vida al nacer de la población femenina para el 2050 se observa que los resultados varían para el 95% de confianza, entre 83.53 años en la hipótesis baja y 86.40 en la hipótesis alta, con una mediana de 85.03; en el caso de los hombres dicha variación oscila entre 78.53 y 80.53 años para esos mismos niveles de confiabilidad, con una mediana de 79.56 años. Es importante

GRÁFICA 17

Esperanza de vida al nacimiento. Hombres y mujeres, 2001-2050

FUENTE: Cálculos propios con base en INEGI, Conapo y El Colegio de México, 2006.

destacar que existe una mayor diferencia en las hipótesis extremas para el caso de las mujeres, de 2.87 y de dos años para los hombres. Esto muestra una mayor incertidumbre en los pronósticos de la mortalidad femenina como resultado de la mayor varianza en la evolución en los niveles de la mortalidad de esta población.

Un hecho importante a destacar es que, independientemente de la hipótesis que se utilice en la dinámica de la mortalidad, los resultados en los montos de la población no serán muy diferentes en los diversos escenarios de este componente demográfico. También significa que dada la historia de la evolución de la mortalidad no podemos aspirar a tener niveles que rompan el comportamiento observado a lo largo de varias décadas; esto es, los niveles de la mortalidad tienen una inercia que impide que ocurran cambios bruscos en la tendencia. Es importante también hacer notar que el método de pronóstico usado

en este trabajo se encuentra muy vinculado al pasado al estar construido con base en la técnica de series de tiempo. A partir de la experiencia observada en la ciencia médica durante el siglo XX, es importante resaltar que habría sido difícil anticiparse a ciertos avances de la medicina que hicieron posible abatir bruscamente los niveles de la mortalidad, como fue el caso, por ejemplo, del descubrimiento de los antibióticos y microbicidas. Asimismo han sido centrales las acciones de expansión de la cobertura de la escolaridad formal, el saneamiento ambiental y la mejora nutricional, entre otras.

En trabajos posteriores podría incorporarse en las proyecciones de este componente, el análisis de las causas de muerte en el método de proyección. Esto permitiría hacer estimaciones futuras considerando parámetros para cada causa en particular.

Finalmente es importante resaltar la virtud del método presentado, al permitir tener intervalos de confianza de las estimaciones de la mortalidad. El enfoque estocástico reconoce que el comportamiento demográfico no sigue leyes determinadas, sino que la demografía, que al final de todo estudia conjuntos humanos, está sujeta a las leyes del azar. Además, el método permite realizar ajustes continuos a los pronósticos, pensando también en el número de muertes que se podrían evitar para alcanzar una meta en la esperanza de vida.

Bibliografía

- Alho, J. (1985), "Uncertain Population Forecasting", *Journal of the American Statistical Association*, vol. 80, núm. 390, pp. 306-314.
- Alho, J. (1990), "Stochastic Methods in Population Forecasting", *International Journal of Forecasting*, núm. 6, pp. 521-530.
- Alho, J. y B. Spencer (2005), *Statistical Demography and Forecasting*, Nueva York, Springer (Springer Series in Statistics).
- Booth, H., J. Maindonald y L. Smith (2002), "Age-time Interactions in Mortality Projection: Applying Lee-Carter to Australia", *Working Papers in Demography*, núm. 85. Canberra, Demography and Sociology Program, Research School of Social Sciences, The Australian National University.
- Box, G. y G. Jenkins (1976), *Time Series Analysis, Forecasting and Control*, San Francisco, Holden-Day.
- Caldwell, J.C. (2006), *Demographic Transition Theory*, Dordrecht, Springer.
- Chatfield, C. (1995), *The Analysis of Time Series. An Introduction*, Nueva York, Chapman and Hall.
- Cohen, J.E. (1986), "Population Forecasts and Confidence Intervals for Sweden:

- A Comparison of Model-Based and Empirical Approaches”, *Demography*, vol. 23, núm. 1, pp. 105-126.
- Conapo (2006), *Proyecciones de la población de México, 2005-2050*, México, Consejo Nacional de Población.
- Gómez de León, José (1990), “Empirical DEA Models to Fit and Project Time Series of Age-Specific Mortality Rates”, Oslo, Central Bureau of Statistics (inédito).
- González Pérez, C. y V. Guerrero Guzmán (2007), “Pronósticos estadísticos de mortalidad y su impacto sobre el sistema de pensiones de México”, Comisión Nacional del Sistema de Ahorro para el Retiro <http://www.consar.gob.mx/premio_pensiones/premio_2007.shtml> [trabajo ganador del Premio de Pensiones 2007].
- Heyde, C.C. y J.E. Cohen (1985), “Confidence Intervals for Demographic Projections Based on Products of Random Matrices”, *Theoretical Population Biology*, vol. 27, núm. 2, pp. 120-153.
- Hyndman, R.J. y H. Booth (2008), “Stochastic Population Forecasts Using Functional Data Models for Mortality, Fertility and Migration”, *International Journal of Forecasting*, vol. 24, núm. 3, pp. 323-342.
- INEGI, Conapo y El Colegio de México (2006), *Conciliación de cifras sobre la dinámica demográfica de México*, México, Instituto Nacional de Estadística y Geografía / Consejo Nacional de Población / El Colegio de México <<http://www.conapo.gob.mx>> (27 de mayo de 2007).
- Kesseli, K. y C. Galindo (2007), “The Many Mexicos. Stochastic Forecast 2001-2050”, *Papeles de Población*, núm. 51, pp. 147-185.
- Le Bras, H. (1971), “Un modèle de projections aleatoires”, documento presentado en la conferencia de la IUSSP, Ciudad de México (mimeo).
- Lee, R. (1974), “Forecasting Births in Post-Transition Populations: Stochastic Renewal with Serially Correlated Fertility”, *Journal of the American Statistical Association*, vol. 69, núm. 347, pp. 607-617.
- Lee, R. (1993), “Modeling and Forecasting the Time Series of US Fertility: Age Distribution, Range, and Ultimate Level”, *International Journal of Forecasting*, vol. 9, pp. 187-202.
- Lee, R. y L. Carter (1992), “Modeling and Forecasting U.S. Mortality”, *Journal of the American Statistical Association*, vol. 87, núm. 419, suplemento “Aging, Health, and Public Policy”, pp. 659-671.
- Lee, R. y T. Miller (2001), “Evaluating the Performance of the Lee-Carter Method for Forecasting Mortality”, *Demography*, vol. 38, núm. 4, pp. 537-549.
- Lee, R. y S. Tuljapurkar (1994), “Stochastic Population Forecast for the United States: Beyond High, Medium and Low”, *Journal of the American Statistical Association*, vol. 89, núm. 428 pp. 1175-1189.
- Li, N., R. Lee y S. Tuljapurkar (2004), “Using the Lee-Carter Method to Forecast Mortality for Populations with Limited Data”, *International Statistical Review*, vol. 72, pp. 19-36.

- Meyer, C. (2000), *Matrix Analysis and Applied Linear Algebra*, Filadelfia, Society for Industrial and Applied Mathematics.
- Oeppel, J. y J. Vaupel (2002), "Broken Limits to Life Expectancy", *Science*, vol. 296, pp. 1029-1031.
- Ordorica Mellado, M. (1995), "El filtro de Kalman en la planeación demográfica", tesis de doctorado, México, Facultad de Ingeniería, UNAM.
- Ordorica Mellado, M. (2004), "Pronóstico de las defunciones por medio de los modelos autorregresivos integrados de promedios móviles", *Papeles de Población*, núm. 42, pp. 249-264.
- ONU (2007), *World Population Prospects. The 2006 Revision*. Nueva York, Population Division, Department of Economic and Social Affairs, Organización de las Naciones Unidas.
- Partida, V. (2003), *Proyecciones de la población de México, de las entidades federativas, de los municipios y de las localidades, 2000-2050*, México, Consejo Nacional de Población.
- Partida, V. (2008), *Proyecciones de la población de México, de las entidades federativas, de los municipios y de las localidades, 2005-2050. Documento metodológico*, México, Consejo Nacional de Población.
- Preston, S., P. Heuveline y M. Guillot (2000), *Demography: Measuring and Modeling Population Processes*, Londres, Blackwell Publishers.
- R Development Core Team (2005), *R: A Language and Environment for Statistical Computing*, Viena, R Foundation for Statistical Computing <<http://www.R-project.org>>.
- Saboia, J.L. (1974), "Modeling and Forecasting Populations by Time Series: The Swedish Case", *Demography*, núm. 11, pp. 483-492.
- Saboia, J.L. (1977), "Autoregressive Integrated Moving Average (ARIMA) Models for Birth Forecasting", *Journal of the American Statistical Association*, vol. 72, núm. 358, pp. 264-270.
- Sykes, Z.M. (1969), "Some Stochastic Versions of the Matrix Model for Population Dynamics", *Journal of the American Statistical Association*, vol. 44, pp. 111-130.
- Vaupel, J. (2004), "The Biodemography of Aging", *Population and Development Review*, vol. 30, suplemento "Aging, Health, and Public Policy", pp. 48-62.
- Wilmoth, J. (1993), "Computational Methods for Fitting and Extrapolating the Lee-Carter Model of Mortality Change", *Technical Report*, Berkeley, Department of Demography, University of California.

Acerca de los autores

Víctor Manuel García Guerrero estudió Actuaría y la maestría en Investigación de Operaciones en la UNAM. Cuenta con un diplomado en Econometría Avanzada en la Facultad de Economía de dicha institución. Es doctor en Estudios de Población por el Centro de Estudios Demográficos, Urbanos y Ambientales de El Colegio de México. Ha sido profesor de asignatura en la Facultad de Ciencias de la UNAM y profesor investigador en la Facultad Latinoamericana de Ciencias Sociales, sede México; ha fungido como asesor en métodos demográficos en el Consejo Nacional de Población y es miembro del Consejo Directivo de la Somede, 2011-2013. Sus temas de investigación son: dinámica demográfica, demografía formal, estimaciones y proyecciones demográficas y políticas de población. Actualmente es profesor investigador del Centro de Estudios Demográficos, Urbanos y Ambientales de El Colegio de México.

Manuel Ordorica Mellado es actuuario por la Facultad de Ciencias de la UNAM, maestro en Demografía por El Colegio de México, y doctor en Ingeniería con especialidad en Investigación de Operaciones por la UNAM. Su tesis de doctorado, “El filtro de Kalman en la planeación demográfica”, obtuvo mención honorífica. Fue jefe del Departamento de Evaluación y Análisis Demográfico en la Dirección General de Estadística; director de Estudios de Población en el Consejo Nacional de Población; consultor en Educación en Población de la UNESCO; coordinador de la maestría en Demografía y del doctorado en Estudios de Población en El Colegio de México, y fungió como director del Centro de Estudios Demográficos y de Desarrollo Urbano de dicha institución. Forma parte del consejo editorial de la revista *Population* (INED, París). Pertenece al Sistema Nacional de Investigadores y recibió el Premio Nacional de Demografía 1998. Trabaja en el campo de la demografía matemática. Actualmente es secretario general de El Colegio de México.