

“Uso de variables sintomáticas para el ajuste de proyecciones de poblaciones menores. El caso del Gran Córdoba (Argentina), 2001-2016”

Autor: Leandro M. González¹

Correo electrónico: leandrogonzalez@yahoo.com.ar

RESUMEN:

Se propone un procedimiento de ajuste de proyecciones de poblaciones menores, combinando las técnicas de proyecciones por componentes y variables sintomáticas. Si se dispone de dos o más proyecciones con diferentes hipótesis de migración, las variables sintomáticas permiten seleccionar la proyección más cercana o definir una nueva. La migración poscensal se deriva de la diferencia proporcional entre la población estimada con variables sintomáticas y las proyecciones más próximas.

Esta metodología resulta particularmente útil en poblaciones con importantes movimientos migratorios. Se ilustra el caso de las localidades del Gran Córdoba (Argentina), con una tasa anual de migración de 18,4 ‰ entre 1991-2001. Se efectuaron estimaciones con registros de nacimientos, defunciones, alumnos primarios y electores. Cuando se compara con un conjunto de proyecciones se observa que, al año 2005, estas localidades mantienen una migración semejante a la década anterior (19,6 ‰). Si se mantiene esta tendencia, la población crecería un 55 por ciento hacia el año 2016.

PALABRAS CLAVE: PROYECCIÓN DE POBLACIÓN – VARIABLES SINTOMÁTICAS – POBLACIONES MENORES



Uso de variables sintomáticas para el ajuste de proyecciones de poblaciones menores por [González, Leandro M.](#) se distribuye bajo una [Licencia Creative Commons Atribución-NoComercial-SinDerivar 4.0 Internacional](#).

¹ Demógrafo. Investigador del Consejo Nacional de Investigaciones Científicas y Tecnológicas (CONICET) y de la Universidad Nacional de Córdoba (Argentina).

1. INTRODUCCIÓN

Este artículo presenta una propuesta metodológica tendiente a ajustar proyecciones de poblaciones menores² con variables sintomáticas, para años posteriores al último censo disponible. Se analiza aquí la utilidad del procedimiento especialmente en áreas de alto movimiento migratorio, como el caso de áreas receptoras de la población que emigra de las grandes ciudades hacia los suburbios (migración intrametropolitana)³. A diferencia de las poblaciones nacionales o provinciales, la migración puede representar el principal factor de cambio de las poblaciones menores.

Las técnicas de proyección demográfica permiten la actualización de la información poblacional con importantes grados de aproximación. Partiendo de hipótesis científicamente respaldadas, es posible contar con estimaciones de la magnitud y composición de las poblaciones, con desagregaciones que respondan al interés del investigador y de los usuarios de la información.

Las estimaciones prospectivas de población para áreas menores permiten una multiplicidad de usos, como insumo de la gestión pública y la actividad económica privada. A partir de ellas se pueden deducir potenciales demandas de obras y servicios públicos, para la cual es posible planificar las correspondientes políticas de ingresos y erogaciones. Asimismo ofrecen una guía adecuada para la promoción de actividades conjuntas entre los municipios y las entidades privadas, tendientes a fomentar el desarrollo local.

La migración representa el factor dinámico menos predecible de la dinámica demográfica, especialmente en poblaciones de menor escala. Mientras la fecundidad y la mortalidad pueden ser extrapoladas a partir de funciones matemáticas, la migración responde a múltiples factores socioeconómicos difícilmente predecibles. Por esta razón se propone una metodología que permite derivar indirectamente el nivel de la migración vigente en una población menor, en años posteriores al último censo disponible.

En primer lugar, se describen las implicancias teóricas de los procedimientos utilizados con variables sintomáticas, analizando sus supuestos y alcances metodológicos. En segundo lugar, se propone la utilización de estas técnicas como procedimiento de ajuste de proyecciones de población efectuadas con el método de los componentes.

Para ilustrar el procedimiento propuesto se presentan los resultados de su aplicación a las localidades del aglomerado Gran Córdoba, pertenecientes al departamento Colón (provincia de Córdoba). La principal razón por la que se elige a las localidades del Gran Córdoba es su alto crecimiento, basado en un importante aporte migratorio.

² El término "poblaciones menores" se emplea aquí para denominar a los residentes de distintas áreas subprovinciales que componen una jurisdicción política mayor, sin establecer un número máximo de habitantes.

³ Para profundizar sobre el fenómeno de la migración interurbana y sus implicancias sociodemográficas recomendamos consultar los trabajos de Pinto da Cunha (2002), Rodríguez Vignoli (2004) y Arriagada Luco-Rodríguez Vignoli (2003).

2- NOCIONES TEÓRICAS

Una proyección demográfica es un procedimiento de cálculo de la evolución futura de una población, partiendo de cursos hipotéticos de fecundidad, mortalidad y migración. Por lo general se trata de cálculos formales que muestran los efectos de los supuestos adoptados sobre una población conocida (IUSPP, 1985:115).

El procedimiento más frecuentemente usado es el método de los componentes, que realiza una simulación del cambio de una población de acuerdo a sus componentes de crecimiento. Este método sigue a cada cohorte de personas de la misma edad a través de su vida tomando en cuenta que está expuesta a la mortalidad, fecundidad y migración. Empezando con la población base por edad y sexo, la población en cada edad específica está expuesta a las posibilidades de muerte, reproducción y migración. El procedimiento completo se repite para cada año del periodo de proyección, dando como resultado la población proyectada por edad y sexo por año calendario (Arriaga, 2001:309-10).

Otros procedimientos de cálculo prospectivo de la población están basados en funciones matemáticas. La población futura de un área se basa en los resultados de los últimos censos de población y en alguna función matemática (generalmente exponencial o logística). Arriaga advierte que estos métodos no producen *proyecciones* propiamente dichas, sino *estimaciones* de población. La principal diferencia radica en que las estimaciones no proporcionan la estructura de edad sino el total de una población, aunque la composición etaria puede ser estimada por otros procedimientos (Arriaga, 2001:412-13).

Una variante de los métodos matemáticos de estimación demográfica son las llamadas “variables sintomáticas”. Siguiendo los conceptos propuestos por Andrew Howe (2004:3) y CELADE (1998:78-79), podemos definir a las variables sintomáticas como un conjunto o serie de datos que muestran alta correlación con los cambios en el tamaño de una población. Se trata, por lo general, de registros estadísticos que se encuentran asociados al volumen y cambio de una población.

Especialistas del CELADE afirman que:

Los métodos de variables sintomáticas intentan enfrentar el desafío de la actualización encontrando variables asociadas al cambio de población y sobre las cuales se cuenta con información permanente y de buena calidad. Esta búsqueda implica encontrar variables que satisfagan dos requisitos: i) presentar una alta correlación con el tamaño y la evolución de la población; ii) contar con registros permanentes. A decir verdad, existen listados de actualización periódica para numerosas variables, tales como los registros de consumidores de servicios básicos (electricidad y agua, por ejemplo), las estadísticas de matrícula escolar, los registros electorales, los registros de hechos vitales, estadísticas de construcción habitacional, etc. (CELADE, 1998:78-79).

Una cantidad no despreciable de estas variables guarda una relación con el cambio demográfico aunque este vínculo no es preciso. A causa de lo anterior es necesario encontrar alguna fórmula o algoritmo que vincule a la población, por un lado, y a la o las variables sintomáticas por otro (CELADE, 1998:78-79). A continuación se reseñan los principales procedimientos empleados para calcular la población de poblaciones menores a partir de variables sintomáticas (Bay, 1998; Teixeira Jardim, 2001; Chaves Esquivel, 2001):

- Distribución por prorratio: Supone que la población entre áreas menores se distribuye en la misma proporción registrada para las variables sintomáticas. Es de

fácil cálculo y requiere información sobre un momento temporal, aunque es muy sensible a la calidad de los registros sintomáticos. La fórmula de cálculo es:

$$P_{i,t+n} = \frac{S_{i,t+n}}{S_{T,t+n}} \times P_{T,t+n} \quad [1]$$

Donde:

$P_{i,t+n}$ = población del área i en el momento t+n.

t = año del último censo.

n = número de años transcurridos desde el último censo.

$S_{i,t+n}$ = valor de la variable sintomática del área i en el momento t+n.

$S_{T,t+n}$ = valor de la variable sintomática del área mayor en el momento t+n.

$P_{T,t+n}$ = población proyectada del área mayor en el momento t+n.

- Distribución proporcional: Se diferencia del método anterior en que ajusta la estimación a los cambios en las variables sintomáticas y en población mayor, en dos momentos de tiempo. Tiene la ventaja de ser consistente con la población total y es menos sensible a los cambios de cobertura de las variables sintomáticas; la desventaja radica en que se requiere información para dos momentos de tiempo. La población se obtiene mediante las siguientes fórmulas:

$$P_{i,t+n} = P_{T,t+n} * \frac{S_{i,t+n}}{S_{T,t+n}} * F_a \quad [2]$$

$$F_a = \frac{P_{T,t+n}}{\sum \left(P_{i,t} * \frac{S_{i,t+n}}{S_{i,t}} \right)} \quad [3]$$

Donde:

F_a = factor de ajuste de la fórmula para que la suma de las $P_{i,t+n}$ sea igual a $P_{T,t+n}$.

- Tasas vitales: Se basa en la evolución de las estadísticas vitales y de las tasas de natalidad y mortalidad, tanto de la población mayor como de las subáreas. Su aplicación está condicionada a la disposición de registros de hechos vitales por localidades, las que no se publican ni difunden en la actualidad para el nivel municipal. Las fórmulas son las siguientes:

$$P_{i,t+n} = a * \left(\frac{B_{i,t+n}}{b_{i,t+n}} \right) + (1-a) * \left(\frac{D_{i,t+n}}{d_{i,t+n}} \right) \quad [4]$$

En donde:

$$b_{i,t+n} = \frac{\frac{B_{i,t}}{P_{i,t}}}{\frac{B_{T,t}}{P_{T,t}}} * \left(\frac{B_{T,t+n}}{P_{T,t+n}} \right) \quad [5]; \quad d_{i,t+n} = \frac{\frac{D_{i,t}}{P_{i,t}}}{\frac{D_{T,t}}{P_{T,t}}} * \left(\frac{D_{T,t+n}}{P_{T,t+n}} \right) \quad [6]$$

a = coeficiente de la ponderación (0,5 para igual ponderación entre las estimaciones derivadas de natalidad y mortalidad).

$B_{i,t+n}$ = nacimientos del área i en el momento $t+n$.

$b_{i,t+n}$ = tasa bruta de natalidad del área i en el momento $t+n$.

$D_{i,t+n}$ = defunciones del área i en el momento $t+n$.

$d_{i,t+n}$ = tasa bruta de mortalidad del área i en el momento $t+n$.

$B_{T,t}$ = nacimientos vivos del área mayor en el año del último censo.

$P_{T,t}$ = población proyectada del área mayor en el año del último censo.

$D_{T,t}$ = defunciones del área mayor en el año del último censo.

- Razón censal: Supone que una población local cambia proporcionalmente con las tasas de ocurrencia de un indicador sintomático del área mayor. Requiere información para dos fechas distintas. Las fórmulas son:

$$P_{i,t+n} = \frac{S_{i,t+n}}{r_{i,t+n}} \quad [7]$$

$$r_{i,t+n} = \frac{r_{T,t+n}}{r_{T,t}} r_{i,t} \quad [8]; \quad r_{i,t} = \frac{S_{i,t}}{P_{i,t}} \quad [9]$$

Donde:

$r_{i,t}$ = tasa de ocurrencia del indicador sintomático S para la localidad i en el año del último censo.

- Diferencia de tasa: Es una variante del método anterior, que emplea diferencias para el cálculo de las tasas de ocurrencia en vez de razones, y posee el mismo alcance. Se aplica la fórmula:

$$r_{i,t+n} = r_{i,t} + (r_{T,t+n} - r_{T,t}) \quad [10]$$

- Compuesto: Emplea distintos registros sintomáticos para estimar proporcionalmente la población de un área local por grupos de edades; la población total se obtiene por adición de la población por los distintos grupos etarios. La principal ventaja es la utilización de cada variable sintomática para estimar el grupo etario con el que está más relacionado; la desventaja es la necesidad de contar con un conjunto de registros sintomáticos de manera simultánea.

- Población de 0-4 años: Utiliza registros de nacimientos.

$$P_{i,t+n}(0-4) = \frac{P_{i,t}(0-4) \times \frac{B_{i,t+n}}{B_{i,t}}}{P_{T,t}(0-4) \times \frac{B_{T,t+n}}{B_{T,t}}} \times P_{T,t+n}(0-4) \quad [11]$$

- Población de 5-14 años: Emplea matrícula escolar primaria.

$$P_{i,t+n}(5-14) = \frac{P_{i,t}(5-14) \times \frac{M_{i,t+n}}{M_{i,t}}}{P_{T,t}(5-14) \times \frac{M_{T,t+n}}{M_{T,t}}} \times P_{T,t+n}(5-14) \quad [12]$$

Donde: $M_{i,t+n}$ = matrícula escolar primaria del área i en el momento t+n.
 $M_{T,t+n}$ = matrícula escolar primaria del área mayor en el momento t+n.

- Población de 15-59 años: Utiliza registros de electores.

$$P_{i,t+n}(15-59) = \frac{P_{i,t}(15-59) \times \frac{E_{i,t+n}}{E_{i,t}}}{P_{T,t}(15-59) \times \frac{E_{T,t+n}}{E_{T,t}}} \times P_{T,t+n}(15-59) \quad [13]$$

Donde: $E_{i,t+n}$ = número de electores del área i en el momento t+n.
 $E_{T,t+n}$ = número de electores del área mayor en el momento t+n.

- Población de 60 años y más: Emplea registros de defunciones.

$$P_{i,t+n}(60+) = \frac{P_{i,t}(60+) \times \frac{D_{i,t+n}}{D_{i,t}}}{P_{T,t}(60+) \times \frac{D_{T,t+n}}{D_{T,t}}} \times P_{T,t+n}(60+) \quad [14]$$

Donde D = defunciones registradas.

- Correlación de razón: Considera que el cambio poblacional está correlacionado con la variación de una o más variables sintomáticas, a través de un modelo de regresión lineal. Su principal fortaleza radica en que las estimaciones se obtienen a partir de modelos probabilísticos. Los inconvenientes que presenta son la necesidad de contar con un número razonable de observaciones en dos momentos de tiempo para poder construir los modelos de regresión, y en el requisito de independencia entre las variables sintomáticas para evitar la multicolinealidad. Las fórmulas son las siguientes:

$$P_{i,t+n} = Y_{i,t+n} * \frac{P_{i,t}}{P_{T,t}} * P_{T,t+n} \quad [15]$$

$$Y_{i,t} = a_0 + a_1 \cdot X_{1,i,t} + \dots + a_j \cdot X_{n,i,t} + u_i \quad [16]$$

$$Y_{i,t} = \frac{\frac{P_{i,t}}{P_{T,t}}}{\frac{P_{i,0}}{P_{T,0}}} \quad [17]; \quad X_{j,i,t} = \frac{\frac{S_{j,i,t}}{S_{j,T,t}}}{\frac{S_{j,i,0}}{S_{j,T,0}}}; \quad (j = 1, \dots, n) \quad [18]$$

Donde:

Y_t = razón entre la proporción de la población del área i sobre el total, en el momento t y el momento 0 .

$X_{j,i,t}$ = razón entre la proporción de la variable sintomática j del área i sobre el total, en el momento t y el momento 0 .

a_i = coeficientes del modelo de correlación lineal múltiple.

u = error de la estimación según el modelo.

- Correlación de tasa: Consiste en un ajuste logarítmico del método anterior, que intenta adaptar el procedimiento de estimación a períodos breves de tiempo. Posee los mismos alcances del método de correlación de razón. Las fórmulas son:

$$Y_{i,t}^* = a_0 + a_1 \cdot X_{1,t}^* + \dots + a_i \cdot X_{n,t}^* + u_i^* \quad [19]$$

Donde:

$$Y_{i,t}^* = \frac{\ln(Y_{i,t})}{k} \quad [20]; \quad X_{i,t}^* = \frac{\ln(X_{i,t})}{k} \quad [21]$$

k = Intervalo de tiempo transcurrido entre los censos.

- Correlación de diferencia: Se trata también de un modelo de regresión lineal, que se diferencia en medir la variación demográfica a través de diferencias en vez de razones. Contiene las mismas ventajas y limitaciones del método de correlación de razón. Las fórmulas son:

$$W_{i,t} = a_t + a_i \cdot Z_{i1} + \dots + a_j \cdot Z_{jn} + e \quad [22]$$

$$W_{i,t} = \frac{P_{i,t}}{P_{T,t}} - \frac{P_{i,0}}{P_{T,0}} \quad [23]; \quad Z_{j,i,t} = \frac{S_{j,i,t}}{S_{T,t}} - \frac{S_{j,i,0}}{S_{T,0}}; \quad (j = 1, \dots, n) \quad [24]$$

Donde:

$W_{i,t}$ = diferencia entre la proporción de la población del área i en el momento t y el momento 0 .

$Z_{j,i,t}$ = diferencia entre la proporción de cada variable sintomática del área i en el momento t y el momento 0 .

Otro concepto involucrado en este trabajo es el de "población menor". Esta noción se encuentra condicionada por la finalidad del análisis al que es sometido. En el caso de las proyecciones, una población puede considerarse menor cuando su magnitud es insuficiente para el cálculo directo de indicadores demográficos, especialmente de las tasas específicas de mortalidad para todos los grupos etarios requeridos. Ello puede ocurrir cuando la población total es inferior a 100.000 habitantes, como cantidad referencial.

Se puede considerar asimismo la noción de población menor desde un punto de vista jurisdiccional. En el caso de Argentina, el Instituto Nacional de Estadísticas y Censos (INDEC) realiza proyecciones demográficas luego de cada censo nacional para el total del país, la Ciudad Autónoma de Buenos Aires y cada

una de las provincias, aplicando el método de los componentes (INDEC, 2005:4). Para las áreas subprovinciales (partidos, departamentos o distritos escolares), se efectúan estimaciones matemáticas basadas en las proyecciones provinciales (INDEC, 2008:5-8). De esta manera, no se cuenta con proyecciones por componentes de las poblaciones residentes en unidades geográficas inferiores. Esto puede responder a que la población de un área subprovincial no es suficiente para la obtención de indicadores demográficos confiables, o bien no se disponen de estadísticas vitales desagregadas para la escala que se requiere. En este caso nos encontraríamos en presencia de una población menor, debido a la carencia de una proyección por componentes o bien de la información necesaria para su resolución.

3- METODOLOGÍA

Para el desarrollo del procedimiento propuesto deben seguirse los siguientes pasos:

- 1º- Recolección de información administrativa susceptible de ser empleada como variables sintomáticas de la población. Se requiere datos para los años coincidentes con los últimos tres censos de población (1980, 1991 y 2001 para el ejemplo propuesto).
- 2º- Aplicación de los procedimientos con variables sintomáticas para el año del último censo. Se efectúan estimaciones de la población en estudio para el año 2001, y se coteja con los resultados del mismo. Se calcula la diferencia relativa de las diversas estimaciones con relación al censo de población.
- 3º- Selección de variables sintomáticas y métodos de estimación. A partir del análisis de los errores de estimación calculados en el paso anterior, se seleccionan aquellos métodos y variables que muestran mayor aproximación con los resultados censales, generalmente menores a 10 puntos porcentuales.
- 4º- Proyección por el método de los componentes con distintos niveles de migración. Se define la población base y las hipótesis de fecundidad y mortalidad para todo el período comprendido en la proyección (2001-16), a partir de la extrapolación logística de los niveles de tasa global de fecundidad y de esperanza de vida. Se define un período de 15 años a partir del último censo a fin de no exceder las posibilidades de la extrapolación logística de los niveles de fecundidad y mortalidad (ARRIAGA, 2001:417). Se definen cuatro escenarios hipotéticos de migración basados en la tasa neta del período intercensal más reciente (1991-2001).
- 5º- Estimación de la población para el último año con variables sintomáticas disponibles. Con los métodos seleccionados se efectúa la estimación de la población total de las localidades del Gran Córdoba para el año 2005, que es el último año para el que se cuenta con registros sintomáticos. Para ello se utiliza la proyección de la población de la provincia de Córdoba (INDEC, 2005), como área mayor para el cálculo con variables sintomáticas.
- 6º- Comparación entre las proyecciones y la población estimada con variables sintomáticas. Se coteja la población total estimada en el paso anterior con las proyecciones definidas en el paso N° 4. De la comparación se puede confirmar alguna de las cuatro proyecciones construidas y adoptar su tasa implícita de

migración, o bien surgir la necesidad de ajustarlas. En el segundo caso se puede definir una nueva hipótesis de migración por la diferencia proporcional entre la población estimada con variables sintomáticas y las proyecciones más próximas; la tasa de migración se deriva de la aplicación de esa misma diferencia proporcional a las hipótesis de migración de las proyecciones más cercanas.

Para efectuar la estimación de la tasa neta anual de migración a partir de las proyecciones definidas y las variables sintomáticas, para un año poscensal determinado, se aplica la siguiente fórmula:

$$m = \left[\frac{PS - PC_{inf}}{PC_{sup} - PC_{inf}} * (m_{sup} - m_{inf}) \right] + m_{inf} \quad [25]$$

Donde:

m = tasa neta de migración.

PS = población estimada con variables sintomáticas.

PC_{inf} = población proyectada por componentes inferior a PS .

PC_{sup} = población proyectada por componentes superior a PS .

m_{sup} = tasa neta de migración de la proyección por componentes PC_{sup} .

m_{inf} = tasa neta de migración de la proyección por componentes PC_{inf} .

- 7º- Proyección definitiva 2005-16. En caso de no aceptar ninguna de las proyecciones hipotéticas disponibles, se construye una nueva proyección con el método de los componentes. Para ello se define la nueva población base para el año 2005 a partir de la estimación del paso N° 5, se le aplica la tasa de migración calculada en el paso anterior, y se replican las mismas hipótesis de fecundidad y mortalidad de las restantes proyecciones.

Para la aplicación de la fórmula [25] se requiere contar al menos con dos proyecciones por componentes, entre cuyos valores queden comprendidas las estimaciones con variables sintomáticas. De cada estimación de población se puede derivar una tasa neta de migración, a partir de la ubicación relativa de la estimación con variables sintomáticas respecto a las proyecciones definidas por componentes. Ello implica considerar que, bajo los supuestos de fecundidad y mortalidad definidos para todas las proyecciones por componentes e igual calidad de la información sintomática, la diferencia registrada entre las estimaciones con variables sintomáticas y dichas proyecciones responde exclusivamente a la migración.

Otra posible interpretación podría postular que, la diferencia relativa entre las estimaciones con variables sintomáticas y las proyecciones por componentes, responde al efecto conjunto de las diferencias entre los niveles reales de fecundidad y mortalidad y las hipótesis respectivas, junto a la migración vigente y los cambios en la calidad de los registros sintomáticos. De esta manera, la tasa neta de migración así calculada sería una estimación del nivel migratorio más un componente residual de las diferencias entre los niveles de fecundidad y mortalidad hipotéticos respecto a los niveles reales, y las variaciones en la calidad de los registros durante el período poscensal.

4- APLICACIÓN

La metodología propuesta se aplica a las localidades del Gran Córdoba para el período 2001-16, a fin de probar su eficacia en un caso testigo. De acuerdo al Instituto Nacional de Estadísticas y Censos (INDEC), el aglomerado Gran Córdoba estuvo conformado en el censo 2001 por la ciudad de Córdoba (1.267.521 habitantes) y 12 localidades del Departamento Colón (100.780 habitantes). Estas últimas localidades contabilizaron 70.879 habitantes, por lo que el crecimiento medio anual intercensal fue de 33,5 por mil (INDEC, 2003). La ciudad capital, por su parte, ha experimentado una tasa anual de crecimiento de 9,9 por mil en el mismo período, por lo que se destaca el rápido incremento demográfico de las localidades periféricas, impulsadas especialmente por la migración desde la ciudad de Córdoba hacia los municipios vecinos.

Se describe en primer lugar la selección de variables sintomáticas. En estas localidades no se dispone de información sintomática suficiente para el período 1980-1991, se toma como referencia los procedimientos y variables sintomáticas que mostraron para el departamento Colón un error menor a los 10 puntos porcentuales respecto al censo 2001. De esta manera se asume el supuesto que las estimaciones que se practican sobre las localidades del Gran Córdoba tienen la misma aproximación que el departamento Colón. Se adopta este supuesto considerando que estas localidades representan 59 por ciento de la población departamental, por lo que las posibles diferencias entre ambas áreas pueden suponerse no significativas y que no invalidan su aplicación.

En el Cuadro 1 se presenta el grado de precisión de las estimaciones de población con variables sintomáticas para el departamento Colón, respecto a los resultados del Censo 2001. Se puede observar que las estimaciones más ajustadas se derivan de los electores, con los primeros cuatro procedimientos las diferencias son menores a la unidad. En orden decreciente de exactitud, las estimaciones más aproximadas se obtienen a partir de los registros de nacimientos con el método de distribución por prorratio (-1,3%), seguido de la media de nacimientos, hechos vitales, alumnos y electores con el mismo procedimiento (1,6%).

Seguidamente se ubican las estimaciones obtenidas con el método compuesto, que emplea de manera conjunta a nacimientos, alumnos primarios, electores y defunciones (-2,7%). A continuación sigue el método de correlación de diferencias, a partir de los registros de defunciones y alumnos (2,8%)⁴. Sigue luego en orden de exactitud el método de correlación de razón, a partir de los registros de defunciones y alumnos (-3,5%)⁵. En quinto lugar se observa la estimación construida con el procedimiento de diferencia de tasas, con la matrícula escolar primaria (3,9%).

⁴ El modelo tiene un coeficiente de determinación $R^2 = 0,8$. La selección de variables independientes se efectuó aplicando diversos métodos.

⁵ El modelo tiene un coeficiente de determinación $R^2 = 0,66$. La selección de variables independientes se efectuó aplicando diversos métodos.

Cuadro 1: Diferencias porcentuales de la población estimada con variables sintomáticas respecto al Censo 2001. Departamento Colón.

MÉTODO DE ESTIMACIÓN	VARIABLES SINTOMÁTICAS							
	Naci- mientos	Defun- ciones	Hechos Vitales	Alum- nos	Elec- tores	Vivien- das	Super- ficie	Media (*)
Distribución por prorrato	-1,3	-23,0	-8,2	16,4	-0,7	57,7	87,4	1,6
Distribución proporcional	-13,4	-18,5	-15,4	6,1	0,1	16,4	47,6	-5,7
Tasas vitales			-15,6					
Razón censal	-13,4	-17,8	-15,2	5,8	0,2	25,6	48,8	-5,6
Diferencia de tasas	-15,4	-17,8	-16,0	3,9	0,2	25,1	56,4	-6,8
Método compuesto	-2,7	-2,7		-2,7	-2,7			
Correlación de razón: 1- Alumnos 2- Defunciones y alumnos				-7,1				
		-3,5		-3,5				
Correlación de tasas: 1- Alumnos 2- Defunciones y alumnos				-7,5				
		-5,5		-5,5				
Correlación de diferencias: - Defunciones y alumnos		2,8		2,8				

(*) Media de las estimaciones efectuadas con nacimientos, hechos vitales, alumnos y electores.

Fuente: Elaboración propia a partir de registros oficiales de estadísticas vitales, matrícula escolar, electores y construcción en la provincia de Córdoba, 1980, 1991 y 2001.

Se observan finalmente una serie de modelos cuyas estimaciones registran diferencias entre los 5 y 10 puntos porcentuales. Se destacan la correlación de tasas a partir de defunciones y alumnos (-5,5%)⁶, la media de estimaciones hechas con distribución proporcional y razón censal, y la razón censal aplicada a los alumnos (5,8%).

En segundo lugar se procede a definir las proyecciones por componentes de esta población. La población base (101.441 habitantes) se construye a partir de la población censada en 2001 (100.780 habitantes), ajustada por omisión censal (2,75%) y con corrección de la población menor de 5 años a partir de la población definida por INDEC para la proyección provincia de Córdoba (Gerencia de Estadísticas y Censos, 2003). La estructura etaria de la población de 5 años y más se mantiene proporcional a los resultados censales, dado que los test de calidad de declaración de la edad son aceptables.

⁶ El modelo tiene un coeficiente de determinación $R^2 = 0,7$. La selección de variables independientes se efectuó aplicando diversos métodos.

La fecundidad se define a partir de las tasas globales calculadas en los años 1991 (2,95 hijos por mujer) y 2001 (2,28). Se formula la hipótesis que las tasas globales irán descendiendo gradualmente desde el valor observado en 2001 hasta los 2,1 hijos por mujer para el año 2010, y 2,05 para el año 2016. No se emplean funciones matemáticas de extrapolación porque producen valores muy bajos de fecundidad para el período considerado.

Para la mortalidad se calcula la esperanza de vida del área conformada por los departamentos Colón y Capital en conjunto, dado que la población en estudio se encuentra en su mayoría en el límite de ambas jurisdicciones. Esta cercanía geográfica permite a la población residente el aprovechamiento de la oferta sanitaria de la ciudad de Córdoba. Partiendo de las esperanzas de vida calculadas para ambos departamentos en los años 1980, 1991 y 2001, se efectúa una extrapolación logística hasta el año 2016.

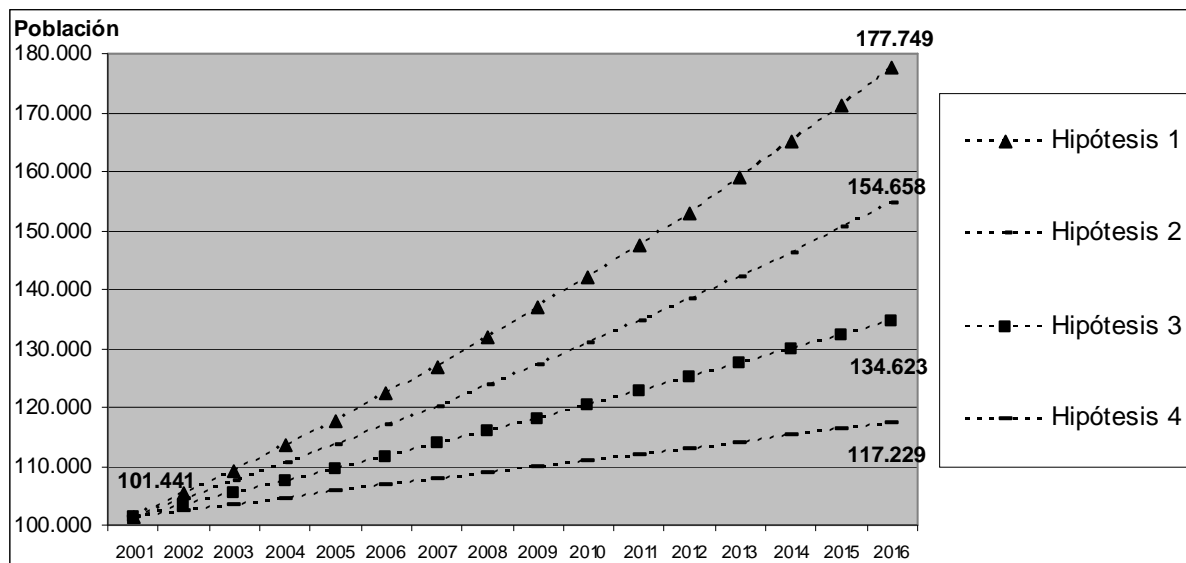
En el caso de la migración, se calcula indirectamente la tasa neta de migración 1991-2001 por el método de la ecuación compensadora, ya que no se dispone de información censal sobre los emigrantes de estas localidades. La estructura etaria de los inmigrantes se deriva de la población censada que cinco años antes residía en otra localidad. De esta manera se asume el supuesto que los emigrantes no afectan la estructura etaria, ya que se trata de un área donde predominan los inmigrantes.

Se propone la proyección de la población de referencia con cuatro hipótesis distintas de migración, a partir de la tasa neta de migración intercensal (18,4‰):

- Hipótesis 1: supone que la tasa de migración equivale a 1,5 veces la tasa observada en el período 1991-2001 (27,6 ‰ anual);
- Hipótesis 2: supone que la migración es igual al período 1991-2001 (18,4 ‰ anual);
- Hipótesis 3: supone que la tasa de migración se reduce a la mitad (9,2 ‰ anual);
- Hipótesis 4: supone migración nula.

Debido a la definición de cuatro hipótesis distintas de migración, se obtienen cuatro proyecciones distintas de la población en referencia. Como las proyecciones mantienen los mismos supuestos de fecundidad y mortalidad, las diferencias entre las mismas responden exclusivamente a los diferentes niveles de migración definidos. En el Gráfico 1 se ilustran las proyecciones resultantes.

Gráfico 1: Población total proyectada de las localidades del Gran Córdoba 2001-16 por hipótesis de migración según año.



Fuente: Elaboración propia.

En tercer lugar, se procede a la estimación de la población poscensal con variables sintomáticas para los años en que se dispone de registros. La información recopilada para las localidades del Gran Córdoba consta de nacimientos, defunciones y matrícula escolar primaria de los años 2001-05, y los electores empadronados en las elecciones realizadas en 2003 y 2005. En el Cuadro 2 se presenta la información para el total de las localidades.

Cuadro 2: Variables sintomáticas seleccionadas de las localidades del Gran Córdoba años 2001-05.

Variable sintomática	2001	2002	2003	2004	2005
Nacimientos	1.723*	2.015	2.114	2.205	2.508
Defunciones	566	606	743	715	628
Alumnos primarios	13.658	13.917	14.249	14.393	14.304
Electores	71.522**	-	76.581	-	81.726

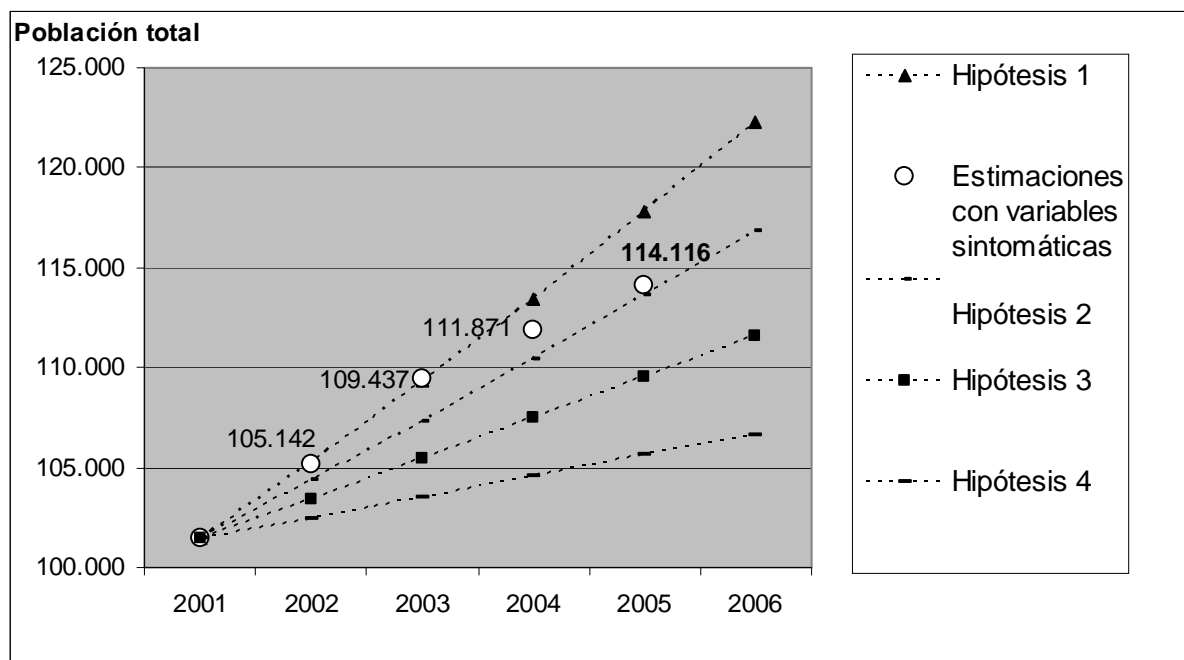
(*) Promedio de los nacimientos registrados en el período 2000-02; los nacimientos registrados en el año 2001 fueron 1394.

(**) Promedio de las elecciones realizadas el 2 de septiembre y el 14 de octubre de 2001.

Fuente: Elaboración propia a partir de datos de los Ministerios de Salud y de Educación de la provincia de Córdoba, y Tribunales Federales de la ciudad de Córdoba.

Con la información anterior se procede a calcular la población de las localidades del Gran Córdoba entre los años 2002 y 2005 (Gráfico 2). En el Cuadro 3 se detallan las estimaciones obtenidas aplicando los procedimientos seleccionados para el año 2005.

Gráfico N° 2: Población estimada con variables sintomáticas y proyecciones de las localidades del Gran Córdoba 2001-06.



Fuente: Elaboración propia.

Cuadro 3: Población estimada con variables sintomáticas seleccionadas. Localidades del Gran Córdoba 2005.

Método	2005						
	Naci- mientos	Hechos vitales	Alumnos	Defunciones y alumnos	Electores	Media*	Todas las variables
Distribución por prorrateo	120.321	107.519			116.290	118.212	113.945
Distribución proporcional					120.691		
Razón censal			113.370		115.173	113.616	
Diferencia de tasas			112.470		114.210	113.201	
Método compuesto							
Correlación de razón			111.720				
Correlación de tasas			110.507				
Correlación de diferencias				110.489			
Estimación media	114.116						

(*) Media de las estimaciones efectuadas con nacimientos, hechos vitales, alumnos y electores.

Fuente: Elaboración propia a partir de datos de los Ministerios de Salud y de Educación de la provincia de Córdoba, y Tribunales Federales de la ciudad de Córdoba.

Debido a que se obtienen numerosas estimaciones para cada año, una primera opción que se presenta es calcular una media aritmética excluyendo los valores extremos. Para los años 2002 a 2004 se excluyen de los promedios las estimaciones calculadas con los registros de nacimientos, debido a que resultan excesivamente elevados en relación con las restantes estimaciones⁷. Para el año 2005 se adopta como población definitiva a la media de todas las estimaciones, 114.116 habitantes, dado que no se observan valores extremos.

⁷ Es posible que la serie de registros de nacimientos posterior al año 2001 haya experimentado variaciones que alteren la relación observada con el crecimiento de población en los años previos.

Otra posibilidad es la elección de un método de estimación en particular, que se considere recomendable por sus supuestos teóricos. Podríamos seleccionar el método compuesto, por ejemplo, teniendo en cuenta que utiliza diferentes registros sintomáticos (nacimientos, alumnos, electores y defunciones) para estimar la población de distintos grupos etarios, con los que guarda mayor correlación. De esta manera la estimación de la población para el año 2005, por el método compuesto, arroja un total de 113.945 habitantes.

Si bien el método compuesto puede contar con ventajas teóricas en relación con los restantes, sólo resulta un 0,15% inferior a la media de las todas las estimaciones calculada en el procedimiento anterior. El cálculo de un promedio, basado en los métodos que mostraron un grado de error inferior al 10 por ciento respecto al último censo disponible (Cuadro 1), es un criterio más bien pragmático. Este recurso puede disminuir el efecto de los supuestos implícitos en cada procedimiento y las limitaciones propias de cada variable, aunque también podría arrastrar sus características menos deseables y equipara procedimientos con diversos grados de solidez estadística. La elección depende del criterio del investigador.

En cuarto lugar, se cotejan las estimaciones seleccionadas con las proyecciones por componentes, como se ilustra en el Gráfico 2. Se puede observar que en los años 2002 y 2003 las variables sintomáticas indican que la población habría seguido una trayectoria muy cercana a la hipótesis 1 de proyección por componentes (migración 50% superior a la calculada entre 1991 y 2001). En el año 2003 la población estimada (109.437 habitantes) alcanza a superar la población proyectada por componentes.

La estimación adoptada para el año 2004 muestra que la población de las localidades del Gran Córdoba se habría ubicado entre las proyecciones N° 1 y 2. Por su parte, la estimación para el año 2005 se acerca a la proyección por componentes N° 2. Esto estaría indicando una desaceleración del crecimiento poblacional desde el año 2002, que lo acercaría a la tasa de migración de la década anterior.

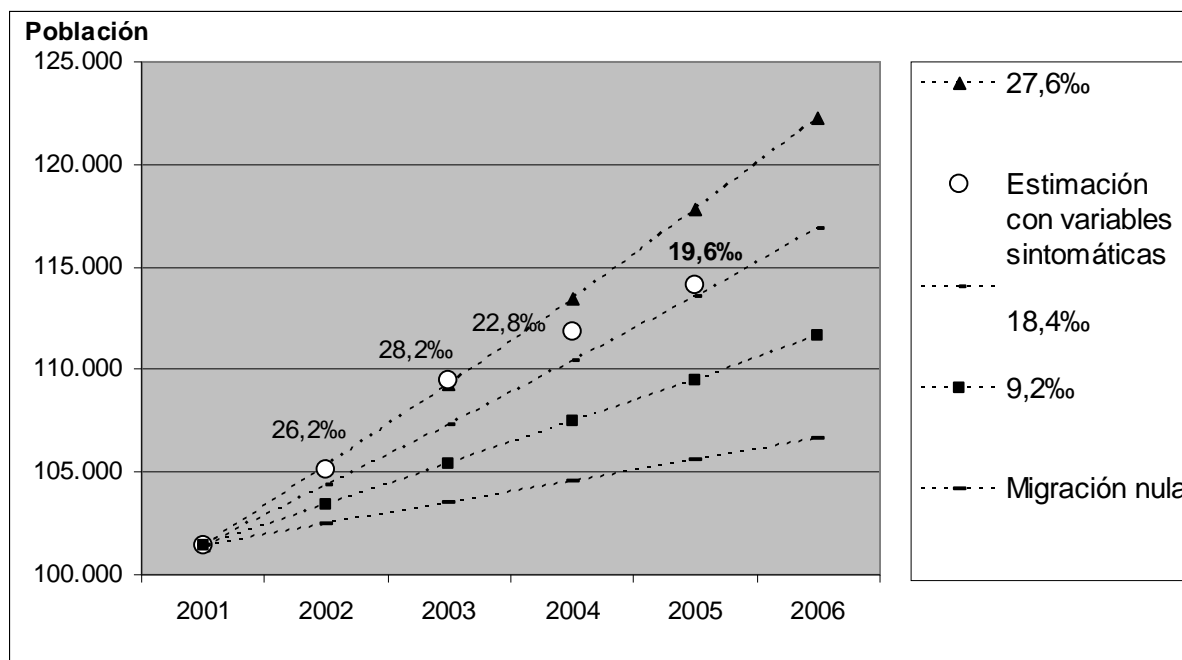
Como la estimación al año 2005 no coincide exactamente con ninguna hipótesis se presentan dos opciones: se adopta la proyección más cercana (hipótesis N° 2) o se construye una nueva proyección. Si se opta por una nueva proyección, corresponde estimar el nivel correspondiente de migración. En el Cuadro 4 y el Gráfico 3 se presenta el cálculo de las tasas netas de migración para cada año posterior al censo, de acuerdo a la fórmula propuesta (punto 6° de la Metodología).

Cuadro 4: Cálculo de la tasa anual de migración de las localidades del Gran Córdoba con proyecciones y variables sintomáticas, años 2002-05.

	2002	2003	2004	2005
Estimación de población con variables sintomáticas	105.142	109.437	111.871	114.116
Población proyectada hipótesis 1	105.290	109.299	113.462	117.785
Población proyectada hipótesis 2	104.343	107.335	110.412	113.573
Tasa migración hipótesis 1 (‰)	27,6	27,6	27,6	27,6
Tasa migración hipótesis 2 (‰)	18,4	18,4	18,4	18,4
Tasa de migración estimada (‰)	26,2	28,2	22,8	19,6

Fuente: Elaboración propia.

Gráfico 3: Tasas netas de migración derivadas de las estimaciones con variables sintomáticas y proyecciones de las localidades del Gran Córdoba 2002-06.



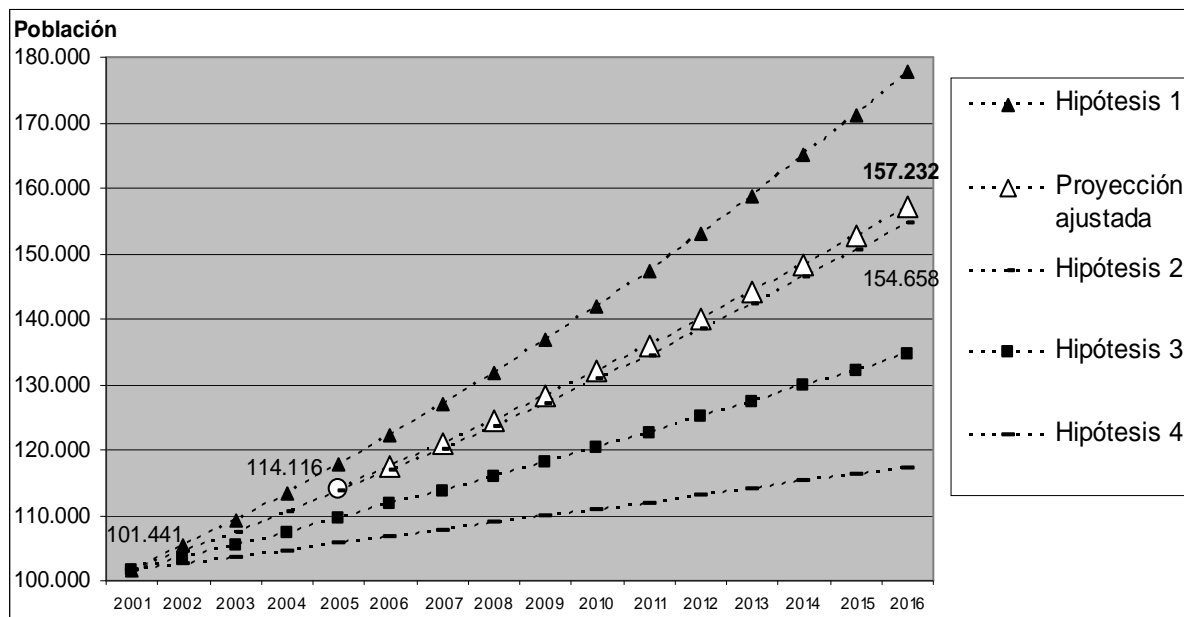
Fuente: Elaboración propia.

De los años considerados solamente en el 2003 se supera el mayor nivel de migración previsto, por lo que su valor debería ser recalculado a partir de una nueva proyección por componentes que posea una tasa de migración mayor a la hipótesis 1 (27,6‰).

Para el año 2005 se deduce una tasa de 19,6 por mil. Esto equivale a suponer que en ese momento la población de las localidades del Gran Córdoba estaba evolucionando bajo las condiciones de fecundidad y mortalidad definidas para todas las proyecciones, y una tasa neta anual de migración de 19,6 por mil habitantes.

Una vez que se ha calculado la nueva tasa de migración, se está en condiciones de correr una nueva proyección. La población total estimada con variables sintomáticas (114.116 habitantes) se adopta como nueva población base. Para definir la estructura por sexo y edad, se puede aplicar proporcionalmente la estructura etaria de la población proyectada más cercana (la hipótesis N° 2) para el mismo año. Luego se incorpora el nivel estimado de migración (19,6 ‰) abierto por tasas por sexo y edad. Si se mantienen las mismas hipótesis de fecundidad y mortalidad, para el año 2016 las localidades alcanzarían una población total de 157.232 habitantes, lo que representa un incremento demográfico del 55 por ciento en sólo 15 años (Gráfico N° 4).

Gráfico 4: Población total proyectada de las localidades del Gran Córdoba 2001-16 por hipótesis de migración según año y ajuste con variables sintomáticas.



Fuente: Elaboración propia.

Esta nueva proyección supone, por su parte, que la tasa de migración estimada en el año 2005 se mantiene constante para el resto del período. Este supuesto puede ser revisado regularmente, a medida que se dispone de nuevos registros de información sintomática que permitan modificar la población total y la tasa de migración. De esta manera se puede realizar una nueva proyección cada vez que se cuente con una nueva estimación confiable de la población total, o bien seleccionar la hipótesis de proyección que a juicio del usuario sea más conveniente.

5- COMENTARIOS FINALES

Se considera, en primer término, que el procedimiento propuesto para el ajuste de una proyección demográfica con variables sintomáticas no presenta inconvenientes teóricos ni metodológicos que invaliden su utilización. Esto se debe a la aplicación independiente de cada método, ya que el ajuste se realiza con los resultados obtenidos en cada procedimiento y se basa en el juicio del investigador. Su originalidad reside en la definición de la población base de una proyección a partir de variables sintomáticas en años poscensales, y en la estimación de la tasa neta de migración contemporánea.

El ajuste propuesto permite el aprovechamiento de las ventajas de cada procedimiento y relativiza sus limitaciones. El método de los componentes ofrece proyecciones robustas, un detalle anual de la estructura de la población y una amplia gama de indicadores demográficos. Brinda también la posibilidad de ensayar proyecciones bajo diversas hipótesis demográficas, e incorporar nuevas estadísticas vitales a medida que se disponen. Su limitación radica en que depende del último censo de población, no pudiendo evaluarse su exactitud hasta la realización de otro censo, por lo general cada diez años en América Latina.

Las variables sintomáticas ofrecen en cambio la posibilidad de estimar la población para cada año que se dispone de información, pudiendo también conocer el grado de aproximación de cada procedimiento respecto al último censo. De esta manera se puede contar con estimaciones para los años posteriores al censo, con un atraso aproximado de uno a dos años respecto al tiempo presente. Su principal desventaja consiste en que calcula únicamente la población total, sin detallar la estructura por edad y sexo.

Es preciso considerar también que los registros susceptibles de ser empleados como variables sintomáticas pueden presentar problemas de calidad. Los registros administrativos suelen tener diversos grados de cobertura territorial, omisiones diferenciales por áreas o sesgos propios de la naturaleza de los datos (por ejemplo la declaración del lugar de residencia en los registros electorales o el subregistro de estadísticas vitales). Estos inconvenientes son de difícil solución en el tratamiento de la información, si no se cuenta con estudios de cobertura de las fuentes que originan los registros. Aún así los datos pueden ser utilizados, en la medida que los sesgos se mantengan relativamente constantes durante el período considerado y que los procedimientos ensayados muestren una aceptable aproximación a los resultados censales.

Con respecto a la aplicación de la metodología propuesta a las localidades del Gran Córdoba, los resultados muestran la persistencia de un sostenido crecimiento demográfico. La aplicación de las variables sintomáticas permite estimar la tasa anual de migración en 19,6 por mil para el año 2005, bajo los supuestos de fecundidad y mortalidad planteadas para todas las proyecciones. Esto implica que el ritmo migratorio en 2005 fue levemente superior al observado en el período intercensal 1991-2001 (18,4%). Si se mantuvieran constantes estos parámetros, la población alcanzaría a 157.232 habitantes en 2016, lo que significaría un crecimiento del 55 por ciento en sólo 15 años.

En síntesis, el empalme de ambas metodologías permite rehacer un conjunto de proyecciones disponibles para cada año posterior al censo. Aunque los censos son los instrumentos adecuados para evaluar las proyecciones, las variables sintomáticas permiten ajustar anualmente una proyección mientras no se cuente con un nuevo censo. Esto redundaría en una importante economía de tiempo para los usuarios de información demográfica, quienes pueden utilizar proyecciones ajustadas periódicamente hasta la disposición de los próximos resultados censales.

Finalmente se considera que esta propuesta es particularmente válida para aquellas poblaciones menores que registran cambios relevantes en su magnitud. Como se ha apreciado en el caso de las localidades del Gran Córdoba, la migración puede constituirse en el factor dinámico de mayor relevancia en una población menor. Las variables sintomáticas permiten actualizar los últimos datos censales, registrar indirectamente el nivel de los movimientos migratorios cada año, y ofrecer una perspectiva a mediano o largo plazo cuando se combina con proyecciones por componentes. De esta manera, la información proporcionada puede constituir una valiosa herramienta para la planificación de la gestión pública y privada.

BIBLIOGRAFÍA

- ARRIAGA, Eduardo (2001): *El análisis de la población con microcomputadoras*. Córdoba, Doctorado en Demografía - Universidad Nacional de Córdoba.
- ARRIAGADA LUCO, C. - RODRÍGUEZ VIGNOLI, J. (2003): Segregación residencial en áreas metropolitanas de América Latina: magnitud, características, evolución e implicancias de política. Santiago, CEPAL-CELADE, Serie Población y Desarrollo N° 47.
- BAY, Guiomar (1998): “El uso de variables sintomáticas en la estimación de la población de áreas menores”, en Notas de Población N° 67/68. Santiago, CELADE, 1998:181-208.
- CELADE (1998): *Demografía II*. México, Programa Latinoamericano de Actividades en Población, Carlos Welti Editor.
- CHAVES ESQUIVEL, Edwin (2001): “Variables sintomáticas en las estimaciones poblacionales a nivel cantonal en Costa Rica”, en Notas de Población N° 71. Santiago, CELADE, 2001:51-72.
- GONZÁLEZ, Leandro (2007): *Ajuste de proyecciones de poblaciones menores con variables sintomáticas. Localidades del Gran Córdoba 2001-16*. Córdoba, Tesis de Doctorado en Demografía, Universidad Nacional de Córdoba, inédito.
- INDEC (2003): Censo Nacional de Población, Hogares y Viviendas 2001. Buenos Aires, INDEC.
- INDEC (2005): Proyecciones provinciales de población por sexo y grupos de edad 2001-2015. Buenos Aires, INDEC, Serie Análisis Demográfico N° 31.
- INDEC (2008): Estimaciones de población total por departamento y año calendario. Período 2001-2010. Buenos Aires, INDEC, Serie Análisis Demográfico N° 34.
- IUSPP (1985): *Diccionario Demográfico Multilingüe*. Liege, CELADE.
- HOWE, Andrew (2004): “Assessing the accuracy of Australia’s small area population estimates, 2001”. Canberra, Australian Population Association.
- PINTO DA CUNHA, José M. (2002): Urbanización, redistribución espacial de la población y transformaciones socioeconómicas en América Latina. Santiago, CEPAL-CELADE, Serie Población y Desarrollo N° 30.
- RODRÍGUEZ VIGNOLI, Jorge (2004): Migración interna en América Latina y el Caribe: estudio regional del período 1980-2000. Santiago, CEPAL-CELADE, Serie Población y Desarrollo N° 50.
- TEIXEIRA JARDIM, Maria de L. (2001): “Uso de variables sintomáticas para estimar la distribución espacial de población. Aplicación a los municipios de Rio Grande do Sul, Brasil”, en Notas de Población N° 71. Santiago, CELADE, 2001:21-50.