En *Estimaciones y proyecciones de población en América Latina. Desafíos de una agenda pendiente*. Río de Janeiro (Brasil): ALAP Editor.

Estimaciones de población en áreas menores en América Latina: revisión de métodos utilizados.

González, Leandro M. - Torres, Eduardo.

Cita:

González, Leandro M. - Torres, Eduardo (2012). *Estimaciones de población en áreas menores en América Latina: revisión de métodos utilizados. En Estimaciones y proyecciones de población en América Latina. Desafíos de una agenda pendiente. Río de Janeiro (Brasil): ALAP Editor.*

Dirección estable: https://www.aacademica.org/leandro.m.gonzalez/58

ARK: https://n2t.net/ark:/13683/ptoh/gFV

Esta obra está bajo una licencia de Creative Commons.
Para ver una copia de esta licencia, visite
https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es.

Acta Académica es un proyecto académico sin fines de lucro enmarcado en la iniciativa de acceso abierto*. Acta Académica fue creado para facilitar a investigadores de todo el mundo el compartir su producción académica. Para crear un perfil gratuitamente o acceder a otros trabajos visite: https://www.aacademica.org*.

Estimaciones de Población en Áreas Menores en América Latina:

revisión de métodos utilizados

Leandro M. González[1]
Eduardo Torres[2]

Resumen

Este trabajo presenta sintéticamente los métodos de estimaciones de población para áreas menores más utilizados en América Latina. Para cada uno de estos métodos se presenta brevemente su forma de cálculo, brindando referencias bibliográficas del método propuesto originalmente para consultas más detalladas. En primer lugar se consignan siete procedimientos de estimación basados en funciones matemáticas. A continuación se detallan nueve metodologías que emplean variables sintomáticas. Finalmente, se realiza un recorrido por las aplicaciones recientes realizadas en América Latina, tanto por estudiosos del ámbito académico, las organizaciones regionales (en especial CELADE) como también por las Oficinas Nacionales de Estadísticas. Con ello no sólo se pretende ofrecer una compilación de los métodos de estimaciones disponibles, sino también una guía para facilitar su estudio, promover su utilización y alentar la investigación en la materia. La gran variedad de informes y aplicaciones disponibles en formato electrónico provenientes de otras geografías, sin duda, permitirán completar el contenido de este capítulo. El uso de estas técnicas en los diferentes países y regiones del continente favorecerá el perfeccionamiento de los procedimientos conocidos, a la vez que aportará nuevas variantes que enriquezcan el "portafolio de herramientas" del demógrafo latinoamericano.

[1] Investigador de CIECS-CONICET y Centro de Estudios Avanzados-U.N. Córdoba (Argentina), leandrogonzalez@yahoo.com.ar
[2] Investigador de CIECS-CONICET y Facultad de Ciencias Económicas-U.N. Córdoba (Argentina), torresedu@gmail.com

Introducción

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).

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 se va modificando por efecto de las defunciones, los nacimientos y las migraciones que anualmente se producen en la población. 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).

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 (2001) advierte que estos métodos no producen *proyecciones* propiamente dichas, sino *estimaciones* de población. Según el autor, 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.

Una variante de los métodos matemáticos de estimación demográfica son las llamadas "variables sintomáticas". A partir de series de registros administrativos disponibles para áreas menores, se construyen modelos matemáticos que relacionan la evolución de dichos registros con el volumen y cambio de una población. De esta manera la información "sintomática" que se obtiene de diversas fuentes posibilita actualizar la población censada en un área, hasta la disponibilidad de un nuevo censo de población.

Tradicionalmente los institutos nacionales de estadística de América Latina utilizan el método por componentes para realizar sus proyecciones oficiales, en conjunto con el Centro Latinoamericano y Caribeño de Demografía (CELADE). En cambio, las estimaciones que emplean variables sintomáticas no están generalizadas en ámbitos oficiales, ya sea por desconocimiento de las técnicas, falta de capacitación del personal y también por la baja frecuencia con que se efectúan estimaciones para áreas pequeñas.

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. Naturalmente la expresión debe ser evaluada para cada país, de acuerdo a la magnitud de su población y a las características de su división territorial interna. Varios países de América Latina cuentan

con provincias o estados cuyas poblaciones permiten la proyección por el método de los componentes, por lo que es conveniente aplicar este procedimiento para el área mayor y luego estimar las poblaciones de las subáreas con métodos matemáticos.

El objetivo de este capítulo es sistematizar los métodos de estimaciones de población para áreas menores más utilizados en América Latina. Para cada uno de estos métodos se presentará brevemente su forma de cálculo, brindando referencias importantes del método propuesto originalmente para consultas más detalladas. Antes de presentar los diferentes procedimientos de estimación, es preciso tener en cuenta algunos criterios que permiten evaluar la precisión de los diversos métodos, que es resumidamente presentado en la siguiente sección.

Evaluación de métodos

Una de las tareas más difíciles en las estimaciones es evaluar la precisión de estas, dado que el valor verdadero de la población es desconocido[3]. Así, en lo general, se analiza el error de las estimaciones producidas con los diferentes procedimientos respecto a la población empadronada en los censos, que contiene errores de cobertura y que pueden ser distintos regionalmente. El criterio más ampliamente empleado para evaluar precisión es el porcentaje de error[4] absoluto (PEA) (Smith *et al.*, 2001), calculado de la siguiente forma:

$$PEA_t = \left[\frac{PE_t - PC_t}{PC_t} \right] * 100 \qquad (1)$$

Donde:

PE_t = Población estimada al momento t.

PC_t = Población censada al momento t.

Entre la mayoría de los trabajos publicados en torno a estimaciones de áreas menores, se utiliza como medida de precisión de los procedimientos al promedio de los PEA, ya sea para el total de las áreas o bien por aglomerados de áreas de acuerdo a algún criterio. Entre los criterios de clasificación de las poblaciones menores se puede utilizar la magnitud demográfica, tasas de crecimiento intercensal o región geográfica, entre otros criterios que pueden ser de interés para el investigador. Otras medidas que pueden aplicarse a los PEA son la mediana, coeficiente de variación, porcentajes de poblaciones entre distintos porcentajes de error, o también valores de la PEA dentro de los cuales se encuentra un alto porcentaje de las áreas menores (80 o 90%).

[3] Buena revisión de los criterios para evaluar las estimaciones y proyecciones de población, con respecto a precisión, nivel de detalle geográfico y demográfico, alcance temporal, plausibilidad y validad entre otros, están disponibles en Smith y colegas (2001) y Howe (2004).

[4] Es importante siempre acordarse de que no es error de estimación, pero de hecho es variación entre estimación y población empadronada.

Con relación al margen de error aceptable para una estimación, un primer criterio aceptado es que los PEA sean inferiores a 10 por ciento, es decir en un rango de +10 y -10 puntos porcentuales respecto a la población censada. Como advierte Howe (2004) este límite debe ser flexible de acuerdo al tamaño de la población, ya que un error del 5 por ciento puede ser aceptable para una población de 2000 habitantes pero inaceptable para una población de 200.000. Smith y colegas (2001) considera que para poblaciones municipales, y para un horizonte de diez años, un error promedio aceptable puede ubicarse entre 8 y 14 por ciento, sin embargo, la variabilidad en el tamaño de la población de municipios en América Latina es muy variable, con municipios muy chicos y muy grandes.

A continuación se presentarán los principales métodos de estimaciones de poblaciones menores más utilizados en América Latina, siguiendo el esquema presentado en el Cuadro 1. Los métodos para realizar proyecciones para poblaciones menores se presentan aquí agrupados en dos grandes grupos: los basados en funciones matemáticas y los más recientes procedimientos basados en variables sintomáticas. Si bien los métodos con variables sintomáticas constituyen una extensión de los procedimientos matemáticos, se los presentará en un apartado especial a fin de cotejar sus semejanzas y diferencias.

Cuadro 1
Lista de Métodos presentados

Métodos basados en funciones matemáticas	Métodos basados en variables sintomáticas
a) Función lineal	a) Distribución por prorrateo
b) Función exponencial	b) Distribución proporcional
c) Función logística	c) Método de tasas vitales
d) Estimación logística de proporciones	d) Método de razón censal
e) Método de los incrementos relativos	e) Método de diferencia de tasa
f) Método de relación de cohortes	f) Método compuesto
g) Método de diferencial de crecimiento	g) Método de correlación de razón
	h) Método de correlación de tasa
	i) Método de correlación de diferencia

Métodos basados en funciones matemáticas

Como mencionado, Arriaga (2001) sugiere que estos métodos no producen *proyecciones* propiamente dichas, sino *estimaciones* de población. Estas técnicas son útiles para estimar la población de pequeñas áreas que no cuentan con estadísticas vitales en períodos intercensales. Los resultados obtenidos pueden ser cotejados y ajustados con proyecciones confiables del total agregado de todas las áreas estimadas, obtenidas con métodos más precisos.

Asimismo, Arriaga (2001) recomienda que las estimaciones basadas en estas funciones se efectúen para períodos cortos de tiempo, que no excedan los 15 años. Durante períodos breves de tiempo, cualquier procedimiento matemático proporcionará resultados similares si se dispone de una proyección de la población total de todas las

subáreas, que sirva como control de ajuste. A continuación se presentarán los métodos matemáticos de estimación más utilizados en América Latina.

Función lineal

La función lineal es el procedimiento más sencillo, ya que supone que el incremento anual promedio de una población de un período reciente se repetirá en el futuro. Las fórmulas son las siguientes (Arriaga, 2001:413-17):

$$P_{t+n} = P_t + n.Y \quad (2)$$

$$Y = \frac{P_t - P_0}{u} \quad (3)$$

Donde:

P_{t+n} = población al momento t+n

P_t = población al momento t

n = período de tiempo cubierto por la estimación

Y = incremento medio anual (en términos absolutos)

P_0 = población en un momento previo 0

u = período de tiempo entre los momentos 0 y t

La principal limitación de la función lineal se puede presentar en aquellas áreas donde se ha verificado un descenso de la población, durante el período previo tomado como referencia. En caso de estimar la población por un período demasiado extenso el resultado podría arrojar un resultado negativo, lo cual no es aceptable en términos demográficos.

Función exponencial

La función exponencial se basa en el hecho que las poblaciones humanas tienden a crecer exponencialmente, cuando están sometidas a condiciones de fecundidad y mortalidad constantes, y en ausencia de migración. La fórmula utilizada es (Arriaga, 2001):

$$P_{t+n} = P_t \cdot e^{r\,n} \quad (4)$$

Donde:

P_t = población al momento t

P_{t+n} = población al momento t+n

e = exponencial

r = tasa anual de crecimiento exponencial

n = período de tiempo cubierto por la estimación

Si la tasa de crecimiento no está disponible, se puede calcular a partir de un período de tiempo previo. La tasa media anual de crecimiento exponencial se calcula de la manera siguiente:

$$r = \frac{h\left(\frac{P_{t+n}}{P_t}\right)}{u} \quad (5)$$

Donde:

ln = representa el logaritmo natural

u = período de tiempo entre los momentos t y t+n

En el Manual VIII (Naciones Unidas, 1978) se cita también el uso de la fórmula del interés compuesto o tasa anual[5] para el cálculo del crecimiento poblacional. En este caso se asimila el concepto de población al de capital, y supone que al final de cada año se añade un incremento que guarda una relación proporcional fija a la población existente al principio del año. Se señala que la aplicación de esta noción de incremento en relación a la población a comienzos de un año no se corresponde con otras tasas demográficas (como la natalidad y la mortalidad), que tienen como referencia la población a mitad de año o período de tiempo. También se advierte que la fórmula del interés compuesto carece de propiedades matemáticas convenientes que pueden producir resultados incongruentes para cálculos poblacionales, que se evitan cuando se emplea la tasa de crecimiento exponencial (Naciones Unidas, 1975:26-27). De todas maneras, si se emplea poblaciones a mitad de año como base de estimaciones de poblaciones futuras y se computa correctamente los períodos de tiempo entre las poblaciones de referencia, los resultados obtenidos con ambas fórmulas de crecimiento son prácticamente iguales.

Función logística

La función logística ha sido ampliamente usada en Demografía para suavizar funciones de las tablas de mortalidad, estimar distribuciones de población por edades correspondientes a subáreas, y proyectar poblaciones de pequeñas áreas entre otras utilidades (Arriaga, 1986). Esta función supone que la población (o un índice) crece al principio aceleradamente y, luego de un valor máximo, su ritmo disminuye con la misma intensidad con que creció (INDEC, 1996). Una de las fórmulas más generales de la función logística es la siguiente:

$$Y = \frac{1}{1+e^{a+wn}} \quad (6)$$

En donde:

a = constante (e^a representa la razón entre las diferencias del índice considerado respecto a las asíntotas superior e inferior)

w = tasa anual de crecimiento del índice considerado

n = período de tiempo

[5] Partiendo de la relación $P_{t+n}=P_t*(1+r)^n$, la tasa de crecimiento es $r = \sqrt[u]{\frac{P^t}{P^0}} - 1$.

Si se hace el supuesto fundamental que la tasa de crecimiento de la función (w) se mantiene constante durante todo el período de tiempo considerado, desde un valor mínimo a otro máximo (asíntotas inferior y superior), se tiene entonces una logística. Entonces:

$$e^a = \frac{U - I_j}{I_j - L} \tag{7}$$

En donde:

L y U = son las asíntotas inferior y superior respectivamente

I_j = valor de un índice determinado en un momento de tiempo

Las fórmulas anteriores permiten estimar la proporción del índice en relación con las asíntotas, en un momento determinado. Si se desea obtener una estimación directa del índice para cualquier fecha, y se conoce (o se estima) el valor de las asíntotas inferior y superior, se emplea la siguiente fórmula (Arriaga, 2001:368):

$$I_{j+n} = L + \frac{V}{1 + \frac{U - I_j}{I_j - L}\, e^{wn}} \tag{8}$$

Donde

$$V = U\text{-}L; \tag{9}$$

$$w = \frac{1}{n}\ 1n \left[\frac{\frac{U - I_{j+n}}{I_{j+n} - L}}{\frac{U - I_j}{I_j - L}} \right] \tag{10}$$

Por otra parte, Murphy (S/D) (*apud* Granados, 1989:135) propuso una modificación a la función logística a fin de ser calculada a partir de la población en dos momentos. Sostiene que los resultados de las estimaciones logísticas de la población no están muy afectados por el valor de la asíntota superior, por lo que este valor puede ser calculado en una forma relativamente burda. Propone estimar la asíntota superior como el valor de la población obtenida al mantener constante durante 60 años la tasa de crecimiento del último período intercensal. La fórmula de cálculo es la siguiente:

$$P = \frac{P_0 * R^{\frac{60}{T}} * \left(R^{\frac{60}{T}} - 1\right)^{\frac{t-t_0}{T-1}}}{\left(R^{\frac{60}{T}} - 1\right)^{\frac{t-t_0}{T-1}} + \left(R^{\frac{60}{T-1}} - 1\right)^{\frac{t-t_0}{T}}} \tag{11}$$

Donde:

P_t = población proyectada al momento t

P_0 y P_1 = poblaciones en cada uno de los censos

t_0 y t_1 = fechas de los censos

$R = P_1/P_0$

$T = t_1 - t_0$

Es preciso aclarar que la fórmula anterior puede ser utilizada cuando la población es creciente en el período considerado; en caso de ser decreciente se debe remplazar el término ($R^{60/T}$ – 1) por (1 – $R^{60/T}$). Asimismo cabe remarcar que el comportamiento de la población y de su tasa de crecimiento es exactamente igual a la curva logística planteada inicialmente.

Estimación logística de proporciones

La función logística puede ser empleada también para estimar poblaciones totales, para luego ser desagregadas por grupos etarios a través de tablas de contingencia. Sobre esta posibilidad Arriaga (2001) afirma:

> La logística es recomendada para estimar poblaciones pequeñas de subáreas para las cuales no se puede aplicar el método de proyección por las componentes. Las estimaciones con la función logística tampoco deben hacerse por períodos largos de tiempo, y pueden mejorarse si existe una población total de control a la cual pueda ajustarse la suma de las poblaciones de las subáreas. (…)
>
> La función logística puede ser usada también para proyectar las poblaciones pequeñas de subáreas si se cuenta con un total de control para la suma de la población por subáreas. En este caso, en lugar de estimar directamente la población para cada subárea, se estiman primero las proporciones de la población de cada subárea con relación al total-control de la población. Puesto que las proporciones varían de 0 a 1, se supone que las asíntotas inferiores y superiores de la logística pueden tener estos valores como límite (Arriaga, 2001:417-18).

Este procedimiento puede ser aplicado a través de las planillas de cálculo que componen el P.A.S. (*Population Analysis Spreadsheets*)[6], desarrollado por Eduardo Arriaga y Peter Johnson en el Bureau of the Census de Estados Unidos (Arriaga, 2001). La planilla "PROYLOG" requiere el ingreso de la siguiente información:

- Población total o por sexos de todas las subáreas de un área mayor en dos censos.
- Fechas exactas de ambos censos.
- Población del área mayor proyectada para los años que se desea estimar.

La planilla estima con la función logística la evolución de la proporción de la población de cada subárea para los años deseados, basada en la evolución de las poblaciones relativas de cada subárea en los dos censos informados. A continuación ajusta las proporciones estimadas para que la suma de las mismas sea igual a 1, y luego calcula la población de cada subárea a partir de la población proyectada del área mayor.

[6] El sistema y su documentación están disponibles en http://www.census.gov/population/international/software/pas/, consulta en 11 de junio de 2012.

Las asíntotas que la planilla usa por defecto son 0 y 1, pero pueden ser modificadas en caso de que las poblaciones menores muestren cambios muy fuertes en el período intercensal. De esta manera se evita la obtención de resultados inconsistentes o exagerados para el período de la estimación.

Una vez estimada la población total (o por sexo) de cada subárea puede utilizarse otra planilla del mismo Sistema para estimar la estructura etaria. La planilla "CTBL32" necesita la siguiente información:

- Población por edad del área mayor, proyectada para un año calendario deseado.
- Población por edad de cada área menor del último censo disponible.
- Población total de cada subárea estimada para el año deseado por procedimientos matemáticos (como por ejemplo, la función logística de las proporciones).

Esta planilla emplea tablas de contingencia que, a través de procesos iterativos, ajusta la distribución etaria censal de las poblaciones menores a los totales estimados para un año de interés y, a su vez, a la estructura etaria de la población mayor proyectada para el mismo momento (Arriaga, 2001:66-68). Este procedimiento está descrito en el Manual VIII (Naciones Unidas, 1975), orientado a proyecciones de población urbana y rural, y es denominado "método de eliminación de la diferencia".

Método de los incrementos relativos

El método de los incrementos relativos (también llamado de participación en el crecimiento, *Apportionment Method o AiBi*) fue introducido en Brasil por Madeira y Simões (1972)[7]. Se fundamenta en la determinación de la proporción del crecimiento absoluto de la población de un área mayor que ha correspondido a la población de cada área menor que la compone, en un determinado período de tiempo. Cuando el crecimiento demográfico de las áreas mayores no presenta cambios significativos, el procedimiento supone para el largo plazo una disminución de las diferencias en los ritmos de crecimiento poblacional de las áreas menores (INDEC, 1996).

La aplicación de este método requiere disponer de la proyección de la población del área mayor para el período que se va a estimar, y la población de cada una de las áreas menores en los dos últimos censos. Para la estimación de la población total de cada área se acepta que (INDEC, 1996):

$$P_i^t = a_i \cdot P_T^t + b_i \qquad (12)$$

Donde:

P_i^t = población del área menor i en el año t

P_T^t = población del área mayor T en el año t

[7] Este método es utilizado para las estimativas oficiales de la población total anual de los Estados y municipios en Brasil (Brito et al., 2010) hace muchos años.

El coeficiente de proporcionalidad del incremento de la población del área menor en relación al incremento de la población del área mayor es igual a:

$$a_i = \frac{P_i^1 - P_i^0}{P_T^1 - P_{Ti}^0} = \frac{P_i}{P_T} \tag{13}$$

$$b_i = \frac{P_i^1 + P_i^0 - \frac{P_i}{P_T}(P_i + P_0)}{2} \tag{14}$$

Cuando se estima la totalidad de las poblaciones de las áreas menores, se tiene que: $\Sigma a_i=1$ y $\Sigma b_i=0$.

Método de relación de cohortes

Este procedimiento fue propuesto por Duchesne (1987) como una adaptación del método de los componentes para áreas menores. Partiendo de una proyección por componentes de la población del área mayor, se calculan las tasas de crecimiento de sus cohortes y se aplican a las cohortes de las poblaciones menores, diferenciando entre cada área menor. Los pasos de acuerdo a Brito y colegas (2010:39-40) son los siguientes:

1. Cálculo de las tasas de crecimiento (CR) entre dos censos de las cohortes de la población mayor (PT):

$${}_{10}CR_{Xi}^{t,t+10} = \frac{{}_{10}PT_{x+10}^{t+10}}{{}_{10}PT_x^t} \tag{15}$$

2. A partir de dos censos anteriores, se calcula el índice de diferencial de crecimiento por cohortes (K) de la población de cada área menor (Pi) en relación con la población mayor:

$${}_{10}K_x = \frac{\left({}_{10}Pi_{x+10}^t \Big/ {}_{10}Pi_x^{t-10}\right)}{\left({}_{10}Pt_{x+10}^t \Big/ {}_{10}Pt_x^{t-10}\right)} \tag{16}$$

3. A cada cohorte de la población menor se le aplica la tasa de crecimiento de la cohorte anterior, multiplicada por el factor K correspondiente a cada área menor:

$${}_{10}Pi_{x+10}^{t+10} = {}_{10}K_x * {}_{10}CR_{Xi}^{t,t+10} * {}_{10}Pi_x^t \tag{17}$$

Método de diferencial de crecimiento

Este procedimiento es semejante al anterior, ya que estima la población por sexo y edad de un conjunto de subáreas, y las ajusta a la proyección de la población mayor

que las contiene. El método fue adaptado por CELADE (1991:6.29) a partir del procedimiento desarrollado por Naciones Unidas (1975). La población de un área menor es estimada a partir de la diferencia del crecimiento de cada grupo etario, que se obtiene como la diferencia entre las tasas exponenciales de crecimiento intercensal de cada población menor y del complemento del área menor respecto al área mayor. Las fórmulas son las siguientes:

$$ {}_5Pi_x^t = \frac{100}{1 + \left(R_x^t \Big/ Pi_x^t \right) * \exp^{-DCx * n}} \qquad (18) $$

$$ {}_5DC_x = {}_5rc_x - {}_5ri_x \qquad (19) $$

Donde:

$R_{t,x}^t$: diferencia entra las poblaciones del área mayor y del área menor en el momento t, del grupo etario x,5.

DC_x: diferencial de crecimiento del grupo de edades a estimar, entre la población del área menor y el complemento de la población del área mayor.

${}_5rc_x$: tasa exponencial de crecimiento intercensal de la población del complemento del área menor respecto al área total.

${}_5ri_x$: tasa exponencial de crecimiento intercensal de la población del área menor.

Métodos basados en variables sintomáticas

Howe (2004:3) define a las variables sintomáticas como cualquier conjunto de datos disponibles que se encuentran relacionados de alguna manera a los cambios en el tamaño de la población. Se trata, por lo general, de registros estadísticos que se encuentran asociados al volumen y cambio de una población.

Los especialistas del CELADE afirman que "*uno de los grandes problemas de las estimaciones y proyecciones de población es el monitoreo de sus resultados*". Debido a que los censos se realizan cada diez años aproximadamente, "*sólo después de transcurrido un decenio es posible verificar la exactitud de las estimaciones y proyecciones efectuadas y rectificar las proyecciones a futuro*" (CELADE, 1998:78).

De acuerdo a CELADE (1988),

> [..], 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.

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).

Bay (1998), por su parte, identifica a las variables sintomáticas con la información estadística relacionada con los cambios en el tamaño de la población. Estos datos permiten la aplicación de métodos basados en relaciones de proporción con poblaciones mayores o modelos de regresión, a fin de detectar cambios en la tendencia del movimiento y crecimiento de la población en períodos intercensales.

En cuanto a su utilidad, Álvarez (2001) afirma que las variables sintomáticas permiten actualizar cifras de población en áreas localizadas, ya que se apoyan en registros administrativos poscensales que entregan visiones recientes de las tendencias demográficas. De esta manera, resultan de gran utilidad para examinar la validez de aquellas proyecciones y señalar correcciones focalizadas.

No es posible definir un listado exhaustivo de los registros que pueden servir como variables sintomáticas. La mayoría de los trabajos coinciden en emplear datos sobre nacimientos, defunciones, matrículas escolares y número de electores son los registros más utilizados, aunque las posibilidades pueden ser más amplias de acuerdo a la disponibilidad y periodicidad de la información en las áreas de interés (Brito *et al.*, 2010). En algunos países desarrollados se emplean también los cambios de residencia en las declaraciones de impuestos a los ingresos, la afiliación al sistema público de seguridad social y los registros periódicos de la población residente en instituciones (fuerzas armadas, congregaciones religiosas, hospitales, hogares geriátricos y de menores, etc.).

Muchos otros registros podrían ser empleados como variables sintomáticas de la población, en la medida que se pudiera acceder a la información. Se necesita construir series de tiempos que incluyan al menos tres censos de población (alrededor de 20 años para el caso de censos decenales), para contar con una medida de la exactitud de las estimaciones respecto a la población del último censo. En el caso de los métodos que emplean regresiones, se requiere además un número adecuado de jurisdicciones menores para poder construir los modelos.

Howe (2000:28) señala los criterios básicos que debe reunir la información para ser empleada como estimadora de la población:

- Habilidad indicativa: debe tener la capacidad de describir cómo cambia la población en el tiempo.
- Serie histórica: la información debe estar disponible para un adecuado número de años, por lo menos 10 años y preferentemente 15.
- Consistencia: las variables deben tener una consistencia razonable durante el tiempo en que es relacionada con la población.

- Disponibilidad geográfica: la información debe estar disponible para el nivel de desagregación geográfica para el cual se requiere estimar la población, teniendo en cuenta los cambios de límites entre las jurisdicciones.
- Disponibilidad temporal: las variables deben estar disponibles poco tiempo después del momento para el cual se estima la población.

Los métodos que utilizan estos indicadores indirectos del tamaño de la población tienen la ventaja de ser sencillos de aplicar y pueden ser actualizados frecuentemente. Por otro lado evidencian limitaciones como la omisión diferencial de registros por áreas, interrupción de registros, alteraciones en la división político-administrativa, cambios en los programas sociales de vivienda, educación, salud y servicios públicos (Bay, 2001).

Chaves Esquivel (2001) advierte que la exactitud de un método no puede ser generalizado, en el sentido que no asegura su eficacia cuando es aplicado a situaciones diferentes. Para ello se deben considerar las hipótesis sobre los métodos disponibles y la calidad de la información que se utiliza, evitando el uso mecánico de los métodos desarrollados. Observa también que la distribución de errores de estimación frente a los resultados censales guarda relación con el tamaño poblacional; mientras más reducida es una población, menor es la precisión de las estimaciones.

La aplicación de estos procedimientos requiere una revisión continua. Métodos que pueden ser útiles en una década determinada pueden resultar imprecisos en las siguientes. Es preciso considerar también que los supuestos adoptados por cada procedimiento y la información empleada sean confiables (Long, 1993). El autor clasifica las variables sintomáticas en dos grupos: información de flujo (*flow data*) o de cantidad (*stock data*). Los primeros representan medidas de los componentes del cambio poblacional (nacimientos, defunciones y migrantes); el método que emplea estas variables es el de los componentes con sus diversas variantes de estimación de los migrantes (principalmente con registros administrativos o matrícula escolar). El segundo tipo de información (de cantidad) es aquella que está correlacionada con el tamaño total de la población, y cuyos cambios son adoptados como indicadores del cambio poblacional. La mayoría de los procedimientos con variables sintomáticas utiliza este último criterio metodológico.

Los métodos que utilizan datos de flujo son recomendables para grandes poblaciones, en las cuales el crecimiento vegetativo explica gran parte del cambio demográfico. Los registros administrativos son empleados para estimar el aporte migratorio, que generalmente representa una porción reducida de la dinámica poblacional. En las poblaciones más reducidas, en cambio, no se puede contar con estas ventajas y por lo tanto son más apropiados los procedimientos que emplean variables de cantidad (Long, 1993).

A continuación se presentan los supuestos metodológicos y las fórmulas que emplean cada uno de los métodos con variables sintomáticas.

Distribución por prorrateo

Este procedimiento supone que la razón entre la población de cada área local y la población total es igual a la razón correspondiente a una variable sintomática (Jardim, 2001). Para su aplicación se requiere disponer de la distribución de la variable sintomática en dos momentos –el período base y en el año deseado-, tanto para las áreas menores como para el total, la población en el período inicial (por lo general, el último censo disponible) de cada área menor y una proyección de la población del área total para el año deseado. La fórmula de cálculo es:

$$P_{i,t+n} = \frac{S_{i,t+n}}{S_{T,t+n}} \times P_{T,t+n} \qquad (20)$$

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.

Las principales ventajas de este procedimiento radican en que puede ser aplicado con una sola variable sintomática, y puede utilizarse información para un solo año calendario. La principal desventaja radica en que las estimaciones están afectadas por los cambios en la calidad de la información sintomática, en caso de no ser constante en el tiempo. Bay (1998) recomienda calcular una media aritmética de un conjunto de estimaciones de población, que se consideren más próximas a la realidad.

Distribución proporcional

Este método parte del supuesto que la población varía en la misma proporción que la variable sintomática. La diferencia con el método de distribución por prorrateo radica en que se efectúa un ajuste a partir de los cambios ocurridos en la variable sintomática y en la población del área mayor, entre los dos momentos considerados. Jardim (2001) afirma que este ajuste permite que haya consistencia de las estimaciones con la población total, e impide que se vean influenciadas por cambios en la cobertura de las variables sintomáticas, siempre que éstas presenten la misma tendencia en todas las áreas menores. La información requerida son los datos de la variable sintomática para dos momentos (por lo general, un censo y el año deseado), la población por área en el momento inicial y una proyección de la población del total de áreas para el año de interés.

La población de un área determinada se obtiene mediante las siguientes fórmulas:

$$P_{i,t+n} = P_{i,t} * \frac{S_{i,t+n}}{S_{T,t+n}} * F_a \qquad (21)$$

$$F_a = \frac{P_{T,t+n}}{\sum\left(P_{i,t} * \frac{S_{i,t+n}}{S_{i,t}}\right)} \quad (22)$$

Donde:

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

Los alcances de este método son semejantes que el de distribución por prorrateo, con la diferencia que se necesita información para dos momentos de tiempo y que las jurisdicciones sean compatibles en los mismos. Bay (1998) afirma que este método tiene la ventaja de permitir la utilización de variables cuya distribución no se asemeje a la de la población, siempre que su variación sea un buen indicador de los cambios de tamaño de la población. Asimismo, se puede calcular la media de un conjunto de estimaciones, que probablemente redunde en una estimación más adecuada.

Método de tasas vitales

Este método fue propuesto por Donald Bogue en 1950 (*apud* Jardim, 2001:34), para la estimación de la población de áreas menores utilizando solamente información proveniente de estadísticas vitales. Requiere información sobre el número de nacimientos y defunciones de cada área menor en el año base y en el año considerado para la estimación, como también la población del área mayor y sus tasas brutas de natalidad y de mortalidad.

Bogue parte del supuesto de que hay una relación inversa entre las variaciones anuales de las tasas de natalidad y mortalidad de las áreas menores, y que la desviación proporcional de esas tasas con respecto a las tasas del área mayor se mantiene constante a lo largo del tiempo. El autor sostiene que es posible hacer una estimación fidedigna de la población de áreas menores usando la razón entre las tasas locales y las del área mayor (Bogue, 1950 *apud* Jardim, 2001:34). Las fórmulas son las siguientes:

$$P_{i,t+n} = a * \left(\frac{B_{i,t+n}}{b_{i,t+n}}\right) + (1 - a) * \left(\frac{D_{i,t+n}}{d_{i,t+n}}\right) \quad (23)$$

$$b_{i,t+n} = \frac{B_{i,t} \,/\, P_{i,t}}{B_{T,t} \,/\, P_{T,t}} * \left(\frac{B_{T,t+n}}{P_{T,t+n}}\right) \quad (24)$$

$$d_{i,t+n} = \frac{D_{i,t} \,/\, P_{i,t}}{D_{T,t} \,/\, P_{T,t}} * \left(\frac{D_{T,t+n}}{P_{T,t+n}}\right) \quad (25)$$

En donde:

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

B= Hijos nacidos vivos

b= Tasa bruta de natalidad

D= Defunciones

d= Tasa bruta de mortalidad

Este método permite realizar distintas ponderaciones entre las estimaciones calculadas a partir de las tasas de natalidad y mortalidad, según el criterio del usuario (Jardim, 2001:34-35).

Método de razón censal

Chaves Esquivel (2001:54-56) presenta este procedimiento con variables sintomáticas, que aplicó en Costa Rica. En este método se consideran las tasas de ocurrencia de un determinado indicador sintomático, partiendo del supuesto que la población local presenta una razón de cambio proporcional a la del área mayor, en el período comprendido entre el último censo y la fecha de estimación.

Primeramente se determina la tasa de ocurrencia de cada variable sintomática para el área menor durante el último período censal. La fórmula es:

$$r_{i,t} = \frac{S_{i,t}}{P_{i,t}} \qquad (26)$$

Donde:

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

$S_{i,t}$ = información de la variable sintomática S observada en la localidad i en el año del último censo

$P_{i,t}$ = población de la localidad i en el último censo

En segundo lugar, se aproxima la tasa de la variable sintomática a la fecha t+n por medio de un parámetro ϕ, que representa la razón de cambio en la variable sintomática entre el año censal t y la fecha t+n, de modo que:

$$r_{i,t+n} = \phi\, r_{i,t} \qquad (27)$$

Para estimar el valor de ϕ es necesario tener información del área mayor (T) que abarque a las localidades menores, sobre la cual se pueden hacer estimaciones independientes de la población en la fecha t+n. De este modo, se calculan las tasas de ocurrencia de S en el área T, en la fecha del censo (t) y la fecha t+n, utilizando la primera fórmula. Si la razón de cambio en la variable sintomática para el área mayor puede aproximarse efectivamente al valor de ϕ, se tiene:

$$\Phi = \frac{r_{T,t+n}}{r_{T,t}} \qquad (28)$$

Entonces, la estimación de $r_{i,t+n}$ es:

$$r_{i,t+n} = \Phi\ r_{i,t} = \frac{r_{T,t+n}}{r_{T,t}}\, r_{i,t} \qquad (29)$$

Finalmente, la estimación poblacional para la localidad i, en el momento t, se obtiene por:

$$P_{i,t+n} = \frac{S_{i,t+n}}{r_{i,t+n}} \quad (30)$$

Por lo general se obtiene más de una estimación demográfica al considerar diferentes variables sintomáticas. La estimación final es generalmente el promedio simple de las estimaciones individuales, pudiéndose también ponderarlas.

La principal ventaja que destaca el autor es la simplicidad del cálculo. La desventaja se señala en la posibilidad que las series sintomáticas podrían verse afectadas por factores distintos a los que justifican los cambios poblacionales.

Método de diferencia de tasa

También propuesto por Chaves Esquivel (2001:56) como una variante del método de razón censal. La principal diferencia entre ambos métodos radica en la forma de estimar la tasa de ocurrencia de la variable sintomática para cada área menor. En este método se supone que el crecimiento de esta tasa para un área menor es igual al crecimiento registrado en el área mayor. De esta manera, la aproximación para la tasa de ocurrencia de este indicador en la fecha t, para el área menor i, se obtiene por la siguiente fórmula:

$$r_{i,t+n} = r_{i,t} + (r_{T,t+n} - r_{T,t}) \quad (31)$$

Método compuesto

Este método es una modificación al método compuesto de Bogue y Duncan (1950, (*apud* Jardim, 2001:35-37), que consiste en la realización de estimaciones de población desagregadas por grupos de edad y sexo. Consiste en la utilización de información sintomática como estimadora de la población por tramos de edad, que luego son adicionadas para la obtención de la población total. La autora propone utilizar los datos de nacidos vivos para estimar la población de 0 a 6 años; la matrícula escolar primaria para estimar la población de 7 a 14 años; el número de electores para estimar la población de 15 a 49 años; y el número de defunciones para estimar la población de 50 años y más Estos grupos etarios pueden ser modificados de acuerdo al criterio del investigador y teniendo en cuenta las características de la población en estudio[8].

[8] González y colegas (2010, 2009:3-4) emplearon este método para la provincia de Córdoba (Argentina) con los siguientes grupos etarios: nacimientos para el grupo 0-4 años, matrícula escolar primaria para el grupo 5-14, electores para el grupo 15-69 años y defunciones para el grupo de 70 años y más.

Las fórmulas empleadas son las siguientes:

- Población de 0-6 años: registro de nacimientos.

$$P_{i,t+n}(0-6) = \frac{P_{i,t}(0-6) \times \frac{B_{i,t+n}}{B_{i,t}}}{P_{T,t}(0-6) \times \frac{B_{T,t+n}}{B_{T,t}}} \times P_{T,t+n}(0-6) \quad (32)$$

Donde: P(0-6) = población de 0-6 años

B = nacimientos registrados

- Población de 7-14 años: matrícula escolar primaria.

$$P_{i,t+n}(7-14) = \frac{P_{i,t}(7-14) \times \frac{M_{i,t+n}}{M_{i,t}}}{P_{T,t}(7-14) \times \frac{M_{T,t+n}}{M_{T,t}}} \times P_{T,t+n}(7-14) \quad (33)$$

Donde M= matrícula escolar primaria.

- Población de 15-49 años: electores.

$$P_{i,t+n}(15-49) = \frac{P_{i,t}(15-49) \times \frac{E_{i,t+n}}{E_{i,t}}}{P_{T,t}(15-49) \times \frac{E_{T,t+n}}{E_{T,t}}} \times P_{T,t+n}(15-49) \quad (34)$$

Donde E= número de electores.

- Población de 50 años y más: registros de defunciones.

$$P_{i,t+n}(50+) = \frac{P_{i,t}(50+) \times \frac{D_{i,t+n}}{D_{i,t}}}{P_{T,t}(50+) \times \frac{D_{T,t+n}}{D_{T,t}}} \times P_{T,t+n}(50+) \quad (35)$$

Donde D = defunciones registradas.

Método de correlación de razón

Es un método desarrollado originalmente por Crosetti y Schmitt (1954 *apud* Bay 1998:187) que se basa en el supuesto de que la evolución de la población está correlacionada con la variación de un conjunto de variables sintomáticas, obteniéndose la población por medio de un modelo de regresión lineal

Este método se basa en las siguientes fórmulas:

$$Y_{i,t} = a_0 + a_1 \cdot X1_{i,t} + \ldots + a_j \cdot Xn_{i,t} + u_i \quad (36)$$

$$Y_{i,t} = \frac{P_{i,t} / P_{T,t}}{P_{i,0} / P_{T,0}} \quad (37)$$

$$Xj_{i,t} = \frac{Sj_{i,t} / Sj_{T,t}}{Sj_{i,0} / Sj_{T,0}}; \quad (j = 1, \ldots, n) \quad (38)$$

Donde:

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

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

u = error de la estimación del modelo.

La población del área i en el momento t+n se encuentra mediante la estimación del modelo de regresión múltiple, basado en el período 0,t (generalmente los dos últimos censos disponibles), la posterior estimación de $Y_{i,\ t+n}$ con base en el período t,t+n, y una proyección del área total T en el momento t+n, o sea:

$$P_{i,t+n} = Y_{i,t+n} * \frac{P_{i,t}}{P_{T,t}} * P_{T,t+n} \quad (38)$$

Documento del Bureau of the Census (1997:2) menciona que Michael Batutis (1991) afirma que la denominación de este método (correlación de razón) es algo oscura, ya que consiste en un modelo de regresión lineal a partir del procedimiento de mínimos cuadrados. El mismo documento se apunta que, posiblemente esto surge de la elección que realizan Crosetti y Schmitt (1954) de los indicadores sintomáticos, a partir del análisis de una matriz de correlación de orden cero entre variables independientes y la población. A pesar de que este procedimiento sería deseable, no es intrínseco al modelo y por lo tanto un mejor nombre sería "regresión de razones" o simplemente "regresión".

Snow (1911:586-88) –autor del que se tiene la referencia más antigua sobre el tema hasta el momento– sostiene que la hipótesis fundamental de este método es que, si una población actual es obtenida con una fórmula definida con la información correspondiente a un momento anterior, la fórmula que se pudiera construir para el momento presente no diferiría de manera significativa de la que se dispone. Por lo tanto, los valores de la población derivados con este modelo son los más probables para cualquier momento posterior al censo. La fórmula no ofrece el valor exacto de la población sino el más probable, basado en la experiencia de una serie de áreas durante el último censo disponible.

Batutis (1991 *apud* Bureau of the Census, 1997:4-5) advierte que la efectividad del método depende de dos factores principales. El primero es la disposición de una

proyección confiable de la población del área mayor; en caso de presentar sesgos, la estimación de las áreas menores será distorsionada. El segundo factor es el supuesto de estabilidad de la relación entre la variable dependiente y las independientes, a lo largo del período de la estimación. Si las relaciones entre las variables cambian para algunas áreas menores el procedimiento efectuará estimaciones defectuosas. Asimismo el modelo puede verse afectado por la multicolinealidad entre las variables independientes, que puede reducir la eficacia de las estimaciones.

Bay (1998) sostiene que la principal ventaja de este procedimiento radica en que la estimación se construye con más de una variable sintomática. Las desventajas se encuentran en la necesidad de contar con mayor información, la que debe ser compatibilizada para los momentos elegidos para estimar el modelo. La autora recomienda la utilización preferentemente de este método, ya que ofrece las mejores aproximaciones con los resultados censales.

Jardim (2001) aplica también este método para estimar la población por grupos de edades. Para ello define los cuatro grupos de mayor interés para planificadores y demógrafos: población infantil (0-4 años), niños en edad de cursar estudios básicos (7-14 años), población en edad reproductiva y productiva (15-49 años) y tercera edad.

Es conveniente considerar también que, de acuerdo a este modelo, las variables sintomáticas pueden ser comprendidas de dos maneras diferentes: como variables correlacionadas con la población o como variables explicativas. En el primer sentido, los registros sintomáticos manifiestan una asociación lineal significativa con la población. Es decir que se observa una variación conjunta en los valores de nacimientos, defunciones, matrícula escolar, electores y la población total de las jurisdicciones consideradas. Esta asociación es medida por los coeficientes de correlación general (R y R^2). Desde esta óptica resulta lógico esperar que los cambios en la dinámica demográfica sean coincidentes con las variaciones en los registros de hechos vitales, demanda educacional y electores, por citar algunas de las principales variables correlacionadas.

La segunda interpretación de las variables sintomáticas se refiere a la naturaleza de la relación entre las variables involucradas. El modelo de regresión lineal múltiple considera a la población (o su cambio) como variable dependiente o explicada por los registros sintomáticos (o sus cambios), los que son considerados variables independientes o explicativas. De esta manera, nuevos registros sintomáticos pueden predecir valores contemporáneos de población. Esta capacidad predictiva es descripta por los coeficientes a_i de la ecuación de regresión.

Al considerar a los registros sintomáticos como variables independientes, se podría pensar que, teóricamente y de acuerdo al Teorema Central del Límite[9], un mayor

[9] El teorema central del límite postula que la suma de n variables independientes, provenientes de poblaciones con distribución cualesquiera, tiende a distribuirse normalmente, a medida que n crece, con media igual a la suma de las medias y varianza igual a la suma de las varianzas (Giuliodori, 1996:183).

número de variables sintomáticas asegurarían una mejor estimación de la población. Es decir que cuanto mayor sea el número de registros sintomáticos independientes que se incorporen al modelo de correlación múltiple, la población tiene más probabilidades de distribuirse normalmente y por lo tanto de ser mejor estimada. Para ello se requiere también de un número grande de observaciones, que en este caso estaría dado por una gran cantidad de mediciones para las variables sintomáticas (áreas geográficas o jurisdicciones).

De acuerdo a esta última interpretación, el cambio poblacional es función del cambio en las variables sintomáticas. Pero si se analiza la naturaleza intrínseca de la información considerada, esta relación es aceptable desde un punto de vista funcional o empírico. Es decir, se asume la relación de dependencia de la población respecto de las variables sintomáticas a fin de aplicar el modelo para la estimación de la población, para aquellos momentos de tiempo en los que sólo se dispone de información sintomática actualizada.

La índole propia de los registros sintomáticos (hechos vitales, alumnos, electores, viviendas construidas, vehículos registrados, etc.) lleva a considerar que éstos existen a partir de la existencia de una población. Es decir que, por el hecho de haber una población determinada, es de esperar la existencia de fenómenos derivados de la misma como demandas de educación, salud, vivienda, transporte y servicios públicos, entre otras. Por lo tanto la relación explicativa sería inversa: la dinámica demográfica explicaría la dinámica de las variables sintomáticas.

Ahora bien, se podría plantear también que la migración tiende a seguir incentivos para la fijación del lugar de residencia de las personas. Suponiendo que los individuos que deciden cambiar de residencia buscan destinos que resulten atractivos para la satisfacción de sus necesidades o expectativas, las variables sintomáticas podrían dar indicios de las áreas con mayores y más variadas ofertas educativa, edilicia y de infraestructura urbana. Por ende, esta oferta diferencial podría considerarse como polos de atracción migratoria y la información sintomática podría operar así como variable explicativa de la población, o al menos de su crecimiento migratorio.

Las dos interpretaciones presentadas pueden servir para una mejor evaluación de los alcances y limitaciones del modelo de regresión. Queda a criterio del investigador la elección del criterio que mejor interprete la aplicación del modelo.

Método de correlación de tasa

Este método es criticado por Chaves Esquivel (2001), que señala que el método de correlación de razón presenta una inconsistencia interna:

> Los coeficientes del modelo se calculan tomando como referencia el tiempo transcurrido entre los censos. Sin embargo, las estimaciones no coinciden con el período base, que por lo general, es menor que el tiempo transcurrido entre los censos. Esto puede afectar la estabilidad natural de los coeficientes y las estimaciones producidas.

> Se ha propuesto una modificación del método de correlación de razón llamada "tasa de correlación", que se basa en una aproximación exponencial de cambio y se realiza tomando el logaritmo natural de cada razón y dividiéndolo por el tiempo transcurrido entre los censos utilizados como referencia (Chaves Esquivel, 2001: 57-58).

De toda forma, el modelo se puede expresar de la siguiente forma:

$$Y^*_{i,t} = a_0 + a_1 \cdot X^*1_{i,t} + \ldots + a_i \cdot X^*n_{i,t} + u_i \quad (40)$$

Donde:

$$Y^*_{i,t} = \frac{\ln(Y_{i,t})}{k} \quad (41)$$

$$X^*_{i,t} = \frac{\ln(X_{i,t})}{k} \quad (41)$$

k = Intervalo de tiempo transcurrido entre los censos.

Swanson y Tedrow (1984 *apud* Chaves Esquivel, 2001) señalan tres razones para suponer que la técnica de correlación de tasa da resultados más precisos, que la de correlación de razón:

i. Como la primera proviene de una transformación de la segunda, se deduce que ambas tienen la misma tendencia. Si ésta se determina durante varios años de estimaciones, el método de correlación de razón presentará mayores inconsistencias, porque estas extrapolaciones están basadas en una relación estructural de cambio a lo largo de un período mayor de años.

ii. La transformación de las tasas reduce las variancias, por lo que se podría lograr una mayor exactitud con la correlación de tasa.

iii. Puesto que se ha reducido el efecto de la autocorrelación espacial, es de esperar que las estimaciones sean más precisas (apud Chaves Esquivel, 2001:58).

Método de correlación de diferencia

Este método fue desarrollado por O'Hare (1976) y consiste en una modificación del procedimiento de correlación de razón, manteniendo los mismos supuestos. La principal diferencia radica en que las variaciones de la población son calculadas por diferencias en lugar de razones. Las fórmulas segundo Bay (1998:189) son las siguientes:

$$W_{i,t} = a_t + a_i * Zi_n + \ldots + a_j * Zj_n + e \quad (43)$$

$$W_{i,t} = \frac{P_{i,t}}{P_{T,t}} - \frac{P_{i,0}}{P_{T,0}} \quad (44)$$

$$Zj_{i,t} = \frac{S_{i,t}}{S_{T,t}} - \frac{S_{i,0}}{S_{T,0}} \;\; ; \;\; (j = 1, \ldots, n) \quad (45)$$

Donde:

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

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

La población del área i en el momento t+n se encuentra mediante la estimación del modelo de regresión múltiple, basado en el período 0,t (generalmente los dos últimos censos), la posterior estimación de $W_{i,t+n}$ con base en el período t,t+n , y una proyección del área total T en el momento t+n, o sea:

$$P_{i,t+n} = \left[W_{i,t+n} + \frac{P_{i,t}}{P_{T,t}} \right] * P_{T,t+n} \qquad (46)$$

O'Hare (1976) advierte que, cuando se aplican los métodos de correlación de razón y de diferencia a las mismas poblaciones, se puede observar que el coeficiente de correlación múltiple R^2 es mayor en el modelo de diferencias que en el de razones; asimismo son mayores también las correlaciones parciales entre las variables independientes y la población.

Bay (1998) afirma que este método presenta las mismas ventajas y desventajas que el método de correlación de razón. Aún así destaca que las variaciones de la población calculadas por diferencias permiten lograr una mayor correlación entre las variables (cambio poblacional y de variables sintomáticas), y que las estimaciones estarían menos afectadas por cambios temporales de las variables.

Experiencias de aplicación en América Latina

En este apartado se realiza una breve reseña de los trabajos publicados sobre estimaciones de poblaciones menores en América Latina, a fin de considerar qué procedimientos fueron los más utilizados, los alcances de las metodologías e información utilizadas, y rescatar las recomendaciones realizadas por los autores para intentar mejorar los futuros trabajos en la región. A continuación se efectuará una revisión de la producción de estimaciones para áreas subnacionales que ofrecen las oficinas nacionales de estadísticas en la región.

En primer lugar se cita el artículo de Bay (1998), ya citado en los apartados anteriores, publicado por CELADE, donde se presentan cuatro métodos de estimación de áreas menores que utilizan variables sintomáticas, y realiza una aplicación en Chile y Costa Rica. Respecto a las estimaciones para áreas pequeñas la autora afirma que básicamente se realizan extrapolaciones matemáticas de niveles y tendencias registradas en el pasado cercano. Esto provoca que los resultados se vuelvan imprecisos cuanto más se alejan las estimaciones del período base, fundamentalmente a causa de las migraciones.

Como alternativa la autora propone la aplicación de métodos basados en variables sintomáticas, utilizando modelos de regresión. Los métodos que aplica para pro-

yectar áreas menores son: distribución por prorrateo, distribución proporcional, de correlación de razón, y el de correlación de diferencia. En el caso de Chile, estima la población de 295 comunas en 1992 utilizando como variables sintomáticas los nacimientos y defunciones, la matrícula de educación básica y la inscripción de automóviles; posteriormente compara las proyecciones obtenidas con los resultados del censo de población de 1992. Para Costa Rica utiliza casi la misma información –con excepción de inscripción de automóviles- para estimar la población de los 85 cantones existentes en 1984, fecha del último censo disponible (Bay, 1998).

De la comparación de las proyecciones obtenidas y los resultados censales concluye que, en el caso de Chile, los modelos de correlación de razón son los más adecuados para realizar estimaciones de poblaciones menores; asimismo observa que en los modelos de prorrateo y distribución proporcional resultan mejores los promedios de las estimaciones obtenidas. Para Costa Rica se obtuvieron estimaciones satisfactorias con los métodos de distribución proporcional y de correlación de razón.

La autora sostiene finalmente que a pesar de los problemas de disponibilidad y calidad de la información, se pueden realizar estimaciones demográficas bastante razonables a escala local. Recomienda la utilización preferentemente del método de correlación de razón, que ofrece las mejores aproximaciones con los resultados censales.

Otra investigación en este campo fue realizada por Jardim (2001), en la que se aplican métodos de actualización de la población de los municipios de Rio Grande do Sul (Brasil) durante la década de 1990, empleando variables sintomáticas. Aplica los métodos expuestos por Bay (1998), una variante del método de correlación de razón para estimar la población por grupos de edad, la técnica de tasas vitales y el método compuesto propuesto por Bogue y Duncan (1950 *apud* Jardim, 2001). En relación con las localidades reducidas, la autora encontró que la variable número de electores suele conducir a una sobreestimación de la población, mientras que las demás variables tienden a provocar una subestimación.

Las principales conclusiones a la que llega la autora son que el método de correlación de razón ofrece resultados muy eficaces, por lo que se considera como el más adecuado para estimar la población de los municipios de Rio Grande do Sul. En cuanto a los errores que observa en las estimaciones de las poblaciones de menor tamaño, concluye que son inherentes al proceso de estimación de universos menores. En general observa que los resultados son satisfactorios y que existen muchas posibilidades para perfeccionar la metodología, sugiriendo la posibilidad de que la eficacia aumentaría si los municipios se estratificaran por tamaño.

Un trabajo semejante realizó Chaves Esquivel (2001) para la estimación de la población para los cantones de Costa Rica en los años 1990. Utiliza la información del censo 1984, el último realizado, y las variables disponibles por cantones (matrícula escolar, declaraciones de impuestos, número de abonados a servicios básicos y electores). Los métodos aplicados son el de razón censal, de diferencia de tasas, de correlación de razón y de correlación de tasa. El autor evalúa el grado de precisión de las estimaciones de cada método a la luz de los resultados del censo 1984. Afirma

que son pocas las variables sintomáticas confiables que se pueden disponer en Costa Rica. Aun así es posible realizar estimaciones poblacionales con una calidad aceptable. Las mejores estimaciones las ofrecieron las técnicas que utilizan regresiones lineales. Si bien los cuatro métodos aplicados presentan diferencias no muy marcadas en los resultados, las técnicas de correlación de razón y de correlación de tasa son evidentemente más precisas.

En Jannuzzi (2005) se propone un método integrado, que emplea proyecciones por componentes para un nivel regional y un sistema de ecuaciones diferenciales para áreas municipales, a partir de un modelo de especies competitivas provenientes de la Ecología. Este enfoque identifica a las poblaciones locales con especies que conviven en un hábitat cerrado, con una limitada capacidad de soporte. Dentro de dicho hábitat, cada especie depende de su crecimiento natural (nacimientos menos defunciones) y de la forma en que interactúan con las otras especies existentes (competencia, parasitismo o predación). Cada población local crece según su grado de atracción dentro del área mayor, que depende de numerosos factores económicos, precio de la tierra y de las viviendas, contaminación ambiental, costo del transporte, etc.

Este procedimiento fue aplicado en Brasil para distintas escalas demográficas. El autor empleó el modelo para estimar la población de 96 distritos de la ciudad de San Pablo entre los años 2005-10. Asimismo fue aplicado para proyectar las 5661 ciudades y localidades brasileñas en el período 2005-20, para el gobierno federal. Advierte el autor que la metodología desarrollada no es una simple extrapolación matemática sino un verdadero modelo demográfico, ya que se basa en una proyección por cohortes y en un sistema de ecuaciones, que discrimina para cada población menor la contribución de la tasa de crecimiento natural y el balance migratorio. Afirma que el modelo puede ser más útil para plazos de cinco a diez años de plazo, en situaciones de escasez de recursos e infraestructura urbana. Señala también que la precisión del método depende de la validez de las hipótesis de la proyección por componentes, como de las especificaciones de los valores de crecimiento natural y migración de las poblaciones menores (Jannuzzi, 2005).

Brito y colegas (2010) analizan la precisión de las estimaciones de población de 60 municipios del Estado de Río de Janeiro. Los autores comparan los resultados del censo 2000 y del conteo 2007 con las estimaciones que arrojan tres métodos: incrementos relativos (o AiBi), relación de cohortes de Duchesne y correlación de razones. Para este último procedimiento se utilizaron información de estadísticas vitales, matrícula escolar y electores. Los resultados obtenidos muestran que los tres procedimientos aplicados ofrecen buenas estimaciones respecto al censo y al conteo de población, siendo el método de correlación de razones el que registra menores errores (4,49% en promedio). La principal explicación que ofrecen los autores se refiere a que las variables sintomáticas representan datos de la dinámica demográfica reciente de las poblaciones, a diferencia de los otros métodos que se basan en el crecimiento demográfico intercensal previo. Observan que las estimaciones son más ajustadas cuando están referidas a municipios mayores (donde el crecimiento vegetativo es más influyente), aunque en el segmento de municipios entre los 150.000 y 1,5 millón

de habitantes los errores medios de estimación son semejantes a las jurisdicciones entre 15.00 y 150.000 habitantes. Asimismo se registran menores errores de estimación en aquellos municipios con tasas de crecimiento intercensal moderadas, en los tres métodos aplicados.

Los autores recomiendan profundizar los trabajos a fin de incorporar nuevas herramientas estadísticas e información correlacionada con la población, incluyendo variables económicas. Esto requiere de un importante esfuerzo de producción de series de datos regulares y confiables por parte de las oficinas públicas. Proponen también estratificar los municipios de acuerdo a sus características económicas y sociales, tratando de identificar el componente migratorio de estas áreas (Brito *et al.*, 2010:44-54).

Álvarez (2001) realizó una estimación de la población de los 509 departamentos existentes en Argentina entre los años 1991 y 1996, utilizando los registros administrativos existentes como variables sintomáticas. Afirma que mientras las proyecciones realizadas por INDEC-CELADE para todo el país (1995) se efectuaron mediante el método de los componentes, la proyección por departamentos (1996) fue elaborada con métodos matemáticos ya que no se contaba con información confiable para dichas jurisdicciones. Por tal motivo se encaró la búsqueda de información administrativa por departamentos que pudieran ser indirectamente asociadas al crecimiento poblacional, se efectuaron estimaciones a 1996 y se compararon los resultados obtenidos con las proyecciones efectuadas por INDEC para el mismo año.

En primer lugar se efectuó la recopilación de información básica, tratando de obtener la mayor cantidad de variables registradas al nivel de departamento para las fechas de los últimos censos (1980 y 1991) y los posteriores al último empadronamiento (1992-1999). Se utilizaron los registros vitales (nacimientos y defunciones) y los registros escolares (matrícula en educación primaria común). Se tomó información inédita de nacimientos registrados por departamento de residencia de la madre y defunciones registradas por departamento de residencia del fallecido. Para que las cifras de áreas menores no estuvieran expuestas a oscilaciones aleatorias, se efectuó un suavizamento mediante los datos del año de referencia, del anterior y del posterior.

La primera conclusión que Álvarez (2001) obtiene es que, mediante el empleo de variables sintomáticas, se pueden obtener modelos capaces de predecir el crecimiento poblacional de áreas menores con un nivel de precisión bastante elevado. En este sentido observa que:

- los métodos de estimación con variables sintomáticas basados en modelos de correlación lineal son más eficaces que los generados a partir de distribuciones o relaciones (como los métodos de distribución por prorrateo, distribución proporcional y método de las tasas vitales);
- entre los métodos de estimación los más eficaces son los que relacionan los valores de dos fechas mediante cocientes (métodos de correlación de razones y de tasas);
- el método de correlación de tasas fue el que evidenció la mayor precisión;

- entre las variables sintomáticas utilizadas en este estudio la más predictiva fue la matrícula escolar, más por su evolución durante el período intercensal que como resultado de su peso relativo en una sola fecha;
- los registros vitales constituyen un elemento importante para formular modelos de estimación y su valor se incrementa cuando se los trata como un todo (suma de nacimientos y defunciones);
- los modelos de correlación mejoran su ajuste en la medida en que se integran distintas variables sintomáticas en un modelo multivariado;
- la estratificación de unidades espaciales debe ser analizada con detenimiento ya que sólo en algunos casos incrementa la precisión de los modelos correlacionales.

La segunda conclusión es que la aplicación del método de variables sintomáticas es una opción factible para examinar proyecciones poblacionales basadas en modelos matemáticos. En este sentido observa que:

- permiten actualizar cifras de población en áreas localizadas, ya que se apoyan en registros administrativos poscensales que entregan visiones recientes de las tendencias demográficas;
- son herramientas apropiadas para efectuar estimaciones de población en departamentos con una dinámica demográfica cambiante;
- existiendo otras proyecciones de población de áreas espaciales menores, comúnmente basadas en métodos matemáticos, constituye un elemento de gran utilidad para examinar la validez de aquellas proyecciones y señalar correcciones focalizadas.

Finalmente, el Álvarez (2011) señala que las estimaciones por variables sintomáticas pueden contribuir al fortalecimiento de las administraciones locales, tanto ofreciendo cifras más adaptadas a las situaciones específicas cuanto en su exigencia de recopilación y sistematización de datos a niveles geográficos desagregados. Asimismo el uso de variables sintomáticas para estimar el crecimiento poblacional es una estrategia para suplir la falta de registros poblacionales continuos que den cuenta de los movimientos migratorios internos, mas no debe verse como un sustituto sino como una demostración de la importancia de contar con tales registros (Álvarez, 2001).

Otras aplicaciones de variables sintomáticas en Argentina fueron ensayadas por González y colegas (2010) para los 26 departamentos de la provincia de Córdoba. A partir de los censos 1991 y 2001 y la información sintomática disponible (estadísticas vitales, electores y matrícula escolar primaria), se aplicaron diversos procedimientos de estimación de la población departamental para ser comparada con los resultados del censo provincial 2008 y nacional 2010.

Estos autores encuentran que la información de electores es la que ofrece las mejores estimaciones de la población, con los métodos de correlación de tasa y de diferencias (1,8 y 2,4% de error medio), distribución proporcional (2,6%), diferencia de tasas (2,6%) y razón censal (2,8%). En orden de precisión sigue la información de alumnos primarios, con diferentes procedimientos que registran errores medios entre

2,2 y 2,9 por ciento. Los registros de defunciones pueden ser utilizados cuando se combinan con los electores en los métodos de regresión lineal, con un error promedio mayor que los obtenidos por la variable electores de manera individual pero con mejores estimaciones para algunas áreas en particular. El método compuesto, por su parte, registra un error medio de 3 por ciento.

En general se observa que los mayores errores de estimación se registran en los departamentos menores a 50.000 habitantes y en aquellos de rápido crecimiento intercensal. Aún así cada departamento muestra diferentes errores según sea el procedimiento y las variables utilizadas, por lo que el investigador debe analizar la conveniencia de los mismos según cuál sea el área geográfica de su interés. Las posibilidades de aplicación son numerosas y la mayor complejidad de cálculo de los procedimientos no necesariamente conducen a estimaciones más precisas (González y otros, 2010:5-14).

En otro trabajo González (2010) 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 a la población estimada para un año poscensal, o bien definir una nueva proyección. La tasa de migración para el período poscensal se puede derivar 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, ya que no hay certidumbre sobre el nivel del flujo migratorio para los años posteriores al último censo y su variación puede afectar seriamente la validez de una proyección. Para ilustrar los alcances de la metodología, el autor la aplica a la población las localidades del Gran Córdoba (Argentina), que registró una tasa anual de migración de 18,4 por mil 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 levemente superior a la década anterior (19,6 ‰). De mantenerse esa tendencia, la población crecería un 55 por ciento hacia el año 2016 (González, 2010:117-27). Una aplicación posterior de la misma técnica a la población del departamento Colón (que contiene a las localidades anteriores), con información sintomática del año 2007, permitió ajustar la proyección de la población y verificó un error de -0,7 por ciento respecto al censo 2010.

Para finalizar, se ofrece una reseña de la labor de estimaciones de población para áreas subnacionales que publican regularmente las Oficinas Nacionales de Estadísticas en la región. Los países para los que se dispone de información son los siguientes:

- Argentina: El Instituto Nacional de Estadísticas y Censos (INDEC) ha publicado proyecciones de la población para las provincias argentinas, y luego estimaciones de por departamentos para el período 2001-2010, basados en el Censo 2001. Para el nivel de desagregación más reducido (departamento o partido),

INDEC calculó la población total por año calendario, a partir de las proyecciones provinciales (obtenidas por el método de los componentes) y el ritmo de crecimiento de cada área en el período intercensal 1991-2001. Los procedimientos de cálculo utilizados fueron la función logística de proyección de proporciones y el método de los incrementos relativos. El informe no aclara si las estimaciones son una media de los resultados de ambos procedimientos, o si bien se aplicaron de manera diferenciada (INDEC, 2008).

- Bolivia: El Instituto Nacional de Estadística ha realizado proyecciones de la población por provincias y municipios, edad y sexo, para los años calendarios del período 2000-10 en los departamentos de Cochabamba y Oruro. El programa utilizado es el "Proyecciones de Población por Edades Simples y Años Calendario" (PPESYAC, CELADE), que fundamentalmente se basa en técnicas de interpolación y ajustes sucesivos, por orden jerárquico descendente, desde el nivel nacional hacia el municipal. El procedimiento metodológico considera dos fases de cálculo articulados: 1) apertura de la población por grupos quinquenales de edad en edades simples, aplicando los Multiplicadores de Sprague; 2) Apertura de la población quinquenal en años calendario, aplicando los multiplicadores de Karup-King (INE Bolivia 2005:17).
- Brasil: El Instituto Brasileño de Geografía y Estadísticas (IBGE) publicó en 2008 la proyección de la población nacional para el período 1980-2050 por el método de los componentes. Las estimaciones de las poblaciones municipales fueron definidas a partir del análisis del comportamiento demográfico observado en los censos de 2000 y 2010, más el conteo poblacional de 2007. Se definieron dos grupos: el primero comprendió a los municipios para los que se adoptó la tendencia de crecimiento en el período 2000-2010, y el segundo con aquellos a los que no se pudo explicar el comportamiento demográfico en los últimos tres relevamientos. Para este último grupo se adoptó como parámetro de cálculo la participación relativa de la población municipal en la población nacional del censo 2010. Para el primer grupo se aplicó el método de incrementos relativos, basado en la tendencia de crecimiento de la última década, tomando como área mayor de referencia la población nacional. Este grupo de municipios se subdividió a su vez en dos conjuntos: por un lado las jurisdicciones mayores de 100.000 habitantes, que se estimaron directamente con este método; por el otro lado los municipios menores fueron reagrupados de acuerdo al tamaño de las poblaciones y las tasas de crecimiento intercensal, se estimó la población conjunta y luego se desagregó por municipios de acuerdo a su población relativa (IBGE 2011).
- Chile: El Instituto Nacional de Estadísticas (INE) publicó en 2008 las proyecciones de población por edad y sexo de las 15 Regiones que componen el país, basados en el Censo 2002. Las Regiones fueron proyectadas por el método de los componentes, y luego se obtuvo la población por comunas a través del método de relación de cohortes de Duchesne (INE Chile, 2008). La información por comunas se ofrece para el período 2000-2010 en los años terminados en 0 y 5, por grandes grupos etarios.

- Ecuador: El Instituto Nacional de Estadísticas y Censos ha definido proyecciones de poblaciones por provincias, cantones y áreas menores por sexo y grupos de edad para el período 2001-2010. Se aplicó el método de los componentes por provincias y se utilizó el programa PRODEM de CELADE, a fin de compatibilizarla con la proyección nacional. Las áreas menores fueron estimadas a través de procedimientos matemáticos, empleando el mismo programa (INEC, 2004).
- Perú: El Instituto Nacional de Estadística e Informática ha desarrollado estimaciones de la población por departamentos, provincias y distritos, por sexo y grupos quinquenales, para el período 2005-2015. El procedimiento utilizado fue el de la "tabla cuadrada" o tabla de contingencia (planilla CTBL32 elaborado por el *Bureau of the Census*), basada en la proyección por componentes de la población nacional. Esta técnica permite ajustar las poblaciones de las subáreas provenientes del último censo de población a la proyección por componentes de la población total (INEI, 2010).
- Uruguay: El Instituto Nacional de Estadística ha actualizado en 2005 la proyección de la población nacional desagregada por departamentos y áreas urbana-rural, para el período 1996-2025 por subáreas y 2025-2050 para el total del país. La proyección del total nacional se obtuvo por suma de las proyecciones por componentes de la población urbana y rural nacional, obtenidas a su vez por agregación de estas áreas correspondientes a cada departamento. Posteriormente se proyecta, sobre la base de los resultados obtenidos para el año 2025, en forma independiente el total del país hasta el año 2050 por el método de componentes. La desagregación de los resultados de población residente en localidades mayores a 5000 habitantes se realizó mediante el procedimiento de estimación logística de proporciones (INE Uruguay, 2005).

La enumeración realizada hasta aquí se limita a los países que cuentan con información metodológica disponible para áreas subnacionales. Se aclara que la mayoría de los países de la región cuentan con estimaciones de áreas menores disponibles en los portales electrónicos de las Direcciones Nacionales de Estadísticas, con diverso grado de desagregación y actualización de resultados.

Referencias bibliográficas

Álvarez, Gustavo (2001), "Estimación de población en áreas menores mediante variables sintomáticas: una aplicación para los departamentos de la República Argentina (1991 y 1996)". Santiago, CELADE, *Serie Población y Desarrollo* N° 13, 36 págs., http://www.eclac.org/publicaciones/xml/3/6843/lcl1481e.pdf.

Arriaga, Eduardo (2001), *El análisis de la población con microcomputadoras*. Córdoba, Doctorado en Demografía - Universidad Nacional de Córdoba.

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, págs.181-208, http://www.cepal.cl/publicaciones/xml/1/5431/LCG2048_p7.pdf .

____ (2001), "Estimaciones indirectas de indicadores demográficos para áreas menores. Situación en América Latina" en *Notas de Población* No 71. Santiago, CELADE, http://www.cepal.cl/publicaciones/xml/3/7223/LCG2114_p1.pdf.

Brito, L., Cavenaghi, S., Jannuzzi, P. (2010), "Estimativas e projeções populacionais para pequenos domínios: uma avaliação da precisão para municípios do Rio de Janeiro em 2000 e 2007", en *Revista Brasileira de Estudos Populacionais*, Río de Janeiro, vol. 27, Nº 1, págs. 35-57.

BUREAU OF THE CENSUS (1997), *Documentation for Puerto Rico Population Estimates*. Washington, Bureau of the Census, http://www.census.gov/population/methods/prmethod.txt .

CELADE (1998), *Demografía II*. México, Programa Latinoamericano de Actividades en Población, Carlos Welti Editor.

____ (1991), *PRODEM. Manual del usuario*. Santiago, CELADE.

Chaves Esquivel, Edwin (2001), "Variables sintomáticas en las estimaciones poblacionales a nivel cantonal en Costa Rica", en *Notas de Población* Nº 71. Santiago, CELADE, págs. 51-72, http://www.cepal.cl/publicaciones/xml/3/7223/LCG2114_p3.pdf .

Duchesne, Louis (1987*), Método de proyecciones de población por sexo y edad para áreas intermedias y menores, Método de relación de cohortes*. Santiago, CELADE.

Giuliodori, Roberto (1996), *Estadística descriptiva y probabilidad*. Córdoba, Ediciones Eudecor.

González, Leandro M. (2010), "Ajuste de proyecciones de poblaciones menores con variables sintomáticas. El caso del Gran Córdoba (Argentina) 2001-16."; Santiago, en *Notas de Población* Nº 91, Santiago, CELADE, págs. 105-28, http://www.cepal.cl/publicaciones/xml/3/42923/lcg2484-P_4.pdf.

González, L., Torres, E., Celton, D. (2010), "Estimación de la población de áreas subprovinciales con variables sintomáticas. Córdoba (Argentina)."; *IV Congreso de la Asociación Latinoamericana de Población*, La Habana, http://www.alapop.org/Congreso2010/DOCSFINAIS_PDF/ALAP_2010_FINAL368.pdf.

Granados, María P. (1989), "Técnicas de proyecciones de población de áreas menores. Aplicación y evaluación", en GRANADOS, María P. (comp.), *Métodos para proyecciones subnacionales de población*. Bogotá, CELADE.

Howe, Andrew (2004), *Assessing the accuracy or Australia's small area population estimates, 2001*. Canberra, Australian Population Association, http://www.apa.org.au/upload/2004-5C_Howe.pdf.

____ (2000), "Methods and Procedures for Estimating Small Area Populations in Australia". Adelaide, Australian Bureau of Statistics - Small Area Population Unit, *Demography Working Paper* 2000/3, http://www.publicpractice.com.au/PDF%20Files/Small%20Area%20Popn%20forecasting.pdf.

IBGE (2011), "Metodologia das estimativas da população residente nos municípios brasileiros para 1° de julho de 2011"; Río de Janeiro, IBGE, Segunda edição,

http://www.ibge.gov.br/home/estatistica/populacao/estimativa2011/metodologia_08112011.pdf .

INDEC (1996), *Estimaciones de población por Departamento 1990-2005*. Buenos Aires, 1996, Serie Análisis Demográfico N° 8, pág. 73.

____ (2008): "Estimaciones de población total por departamento y año calendario. Período 2001-2010"; Buenos Aires, INDEC, http://www.indec.mecon.ar/nuevaweb/cuadros/2/estimaciones-serie34.pdf.

INE BOLIVIA (2005), *Proyecciones de población por Provincias y Municipios, según Sexo, Edades Simples y Años Calendario, Período 2000 - 2010. Cochabamba*, La Paz, INE, http://www.ine.gob.bo/PDF/DIFD/ProyeccionesPoblacionProvinciasMunicipiosCochabamba.pdf .

INE CHILE (2008), *Proyecciones y Estimaciones de Población. 1990-2020. País y Regiones*, INE, Santiago, http://palma.ine.cl/demografia/menu/EstadisticasDemograficas/DEMOGRAFIA.pdf.

INE URUGUAY (2005), *Estimaciones y proyecciones de la población de Uruguay (Revisión 2005);* INE, Montevideo, http://www.ine.gub.uy/biblioteca/metodologias/informe%20metodol%F3gico.pdf .

INEC (2004), *Proyecciones de población por provincias, cantones, áreas, sexo y grupos de edad. Período: 2001 - 2010,* Quito, INEC, http://www.inec.gob.ec/inec/index.php?option=com.

INEI (2010), *Estimaciones y Proyecciones de Población total y edades quinquenales, según Departamento, Provincia y Distrito, 2005-2015*, Lima, INEI, http://www.inei.gob.pe/biblioineipub/bancopub/Est/Lib1010/Libro.pdf.

IUSPP (1985) *Diccionario Demográfico Multilingüe*. Liege, CELADE.

Jardim, Maria de L Teixeira. (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* Nº 71. Santiago, CELADE, págs. 21-50, http://www.cepal.cl/publicaciones/xml/3/7223/LCG2114_p2.pdf.

Jannuzzi, Paulo de M. (2005), "Population Projections for Small Areas: Method and Applications for Districts and Local Population Projections in Brazil" en *Congreso Internacional de la Unión Internacional para el Estudio Científico de la Población*. IUSSP, http://iussp2005.princeton.edu/download.aspx?submissionId=51422.

Long, John F. (1993), "Postcensal Population Estimates: States, Counties, and Places". Washington, US Bureau of the Census, *Population Division working Paper* Nº 3, http://www.census.gov/population/www/documentation/twps0003.html.

Madeira, J.L. – Simões, C.C.S. (1972), "Estimativas preliminares da população urbana e rural, Segundo as unidades da Federação, 1960/1980: por uma nova metodologia", en *Revista Brasileira de Estatística*. Río de Janeiro, ABE-IBGE, V. 33, Nº 129, págs. 3-11.

NACIONES UNIDAS (1975), "Manual VIII. Métodos para hacer proyecciones de la población urbana y rural". New York, Departamento de Asuntos Económicos y Sociales, *Estudios de Población*, Nº 55.

O'Hare, William (1976), "Report on a Multiple regresión Method for Making Population Estimates" en *Demography*. Maryland, Vol. 13, Nº 3, págs. 369-79.

Snow, E. C. (1911), "The Application of the Method of Multiple Correlation to the Estimation of Post-Censal populations", en *Journal of the Royal Statistical Society* V. 74, Nº 6. Londres, págs. 575-629. www.jstor.org.