
población y desarrollo

Estimación de población en áreas menores mediante variables sintomáticas: una aplicación para los departamentos de la República Argentina (1991 y 1996)

Gustavo Álvarez



NACIONES UNIDAS



Centro Latinoamericano y Caribeño de Demografía
(CELADE) – División de Población.

Santiago de Chile, enero de 2001

Este documento fue preparado por el Licenciado Gustavo Álvarez, en el marco de la cooperación del CELADE con el Instituto Nacional de Estadísticas y Censos (INDEC), dentro del programa BID/CELADE (ATN/SF-5827-RG). Se contó con la asesoría técnica de Guiomar Bay, consultora del CELADE, y con la colaboración de Soledad Triano y Daniel Fridman de la Dirección de Estadísticas Poblacionales del INDEC.

Las opiniones expresadas en este documento, que no ha sido sometido a revisión editorial, son de exclusiva responsabilidad del autor y pueden no coincidir con las de la Organización.

Publicación de las Naciones Unidas

LC/L.1481-P

ISBN: 92-1-321774-9

Copyright © Naciones Unidas, enero de 2001. Todos los derechos reservados

N° de venta: S.01.II.G.14

Impreso en Naciones Unidas, Santiago de Chile

La autorización para reproducir total o parcialmente esta obra debe solicitarse al Secretario de la Junta de Publicaciones, Sede de las Naciones Unidas, Nueva York, N.Y. 10017, Estados Unidos. Los Estados miembros y sus instituciones gubernamentales pueden reproducir esta obra sin autorización previa. Sólo se les solicita que mencionen la fuente e informen a las Naciones Unidas de tal reproducción.

Índice

Resumen	5
Introducción	7
I. Consideraciones acerca de la información básica	9
II. Presentación de distintos modelos de aplicación de variables sintomáticas	13
III. Resultados de aplicación de variables sintomáticas ...	19
A. Examen de la precisión de los modelos frente a datos censales de 1991	20
B. Comparación entre estimación por variables sintomáticas y proyecciones poblacionales por métodos matemáticos hacia 1996.....	25
IV. Conclusiones y recomendaciones	29
Bibliografía	33
Serie Población y desarrollo: números publicados	35

Índice de cuadros

Cuadro 1	Análisis del error de las estimaciones de población por departamento según los métodos de prorrateo, de distribución proporcional y de tasas vitales. Argentina, 1991	21
Cuadro 2	Análisis del error de las estimaciones de población por departamento según los métodos de correlación de razones, de diferencias y de tasas. Argentina, 1991.....	23
Cuadro 3	Análisis del error de las estimaciones de población por departamento del modelo MND según los métodos de correlación de razones, de diferencias y de tasas con unidades estratificadas por crecimiento de población, densidad poblacional y urbanización. Argentina, 1991	24
Cuadro 4	Comparación entre estimación de población por departamento por métodos matemáticos (INDEC 1996) y por método de correlación de tasas mediante la aplicación de variables sintomáticas MND y estratificando las unidades por nivel de crecimiento entre 1980 y 1991. Argentina, 1996	25
Cuadro 5	Comparación entre valores de estimación de población por departamento por métodos matemáticos (matemat) y por método de variables sintomáticas (sintom). Departamentos seleccionados. Argentina, 1996.....	27

Resumen

Los métodos de estimación indirecta de indicadores sociodemográficos para áreas geográficas menores han tenido un importante desarrollo en América Latina y el Caribe en los últimos años. Algunos de estos métodos tienen referentes en trabajos de hacia varias décadas atrás y durante la década de 1990 confluyeron tres principales factores para que se produjera un significativo impulso a esta línea de investigación: a) el fuerte aumento de la demanda por este tipo de información, originado en requerimientos legales vinculados a la representación política, ciertas asignaciones presupuestarias en el marco de procesos de descentralización, o de la focalización de las políticas sociales en general; b) una mayor frecuencia, periodicidad y sistematización del levantamiento de encuestas de hogares y demográficas con representación nacional, urbana y rural; y c) la más amplia disponibilidad de medios computacionales de bajo costo y alta velocidad de procesamiento de grandes bases de datos, como lo son las encuestas y, sobre todo, los censos de población y vivienda.

En este marco, resulta fundamental conocer el tamaño de la población para todas las unidades geográficas del país. Dentro de este panorama, este trabajo utiliza los registros administrativos existentes (variables sintomáticas) junto con modelos estadísticos que permitan actualizar las estimaciones de población de áreas menores en períodos intercensales.

El programa BID/CELADE (ATN/SF-5827-RG) ha venido trabajando en este área, la que ha tenido una instancia de síntesis en el curso taller internacional sobre el tema¹. Como parte del seguimiento de esta actividad se ha considerado oportuno realizar un estudio ilustrativo de un país con información adecuada de censos y registros. Para ello se ha seleccionado el caso de Argentina, que podría servir, junto a los estudios anteriormente elaborados en Chile, Costa Rica y Brasil, de base para aplicaciones de estos métodos en otros países de la región.

¹ “Taller sobre estimaciones e indicadores demográficos y económicos para áreas menores”, llevado a cabo en la ciudad de Lima, Perú, entre el 27 y el 31 de marzo de 2000.

Introducción

Durante los últimos años se ha acentuado, tanto en la Argentina como en el resto de la región latinoamericana, la demanda de información con un alto nivel de desagregación espacial. Esta tendencia ha sido propulsada tanto por la estrategia de focalización espacial adoptada por los organismos públicos nacionales como por la creciente iniciativa de los municipios en la gestión de políticas y programas sociales. Frente a estas nuevas exigencias, el repertorio tradicional de información ofrecida por los organismos oficiales de estadística requiere ser actualizado tanto por el fortalecimiento de los datos localizados como por la implementación de metodologías específicas para las áreas geográficas menores.

Las proyecciones de población para el total del país (INDEC-CELADE, 1995) han sido realizadas por el método de las componentes, contemplando el crecimiento poblacional intercensal y definiendo hipótesis acerca de la evolución de la fecundidad, la mortalidad y las migraciones internacionales. Mas esta metodología no ha podido ser replicada al nivel de las jurisdicciones más elementales, los departamentos, por cuanto la información no es suficientemente confiable y la inestabilidad de la migración interna no admite la formulación de hipótesis a mediano plazo (INDEC, 1996). Por tal motivo, las estimaciones de población por departamento se llevaron a cabo mediante métodos matemáticos que tomaron como insumo la población enumerada en los censos de población y vivienda de 1980 y de 1991.

El Centro Latinoamericano y Caribeño de Demografía (CELADE) ha promovido la investigación de alternativas metodológicas para mejorar las estimaciones poblacionales derivadas de extrapolaciones matemáticas. Al respecto, se hicieron experiencias con la metodología de variables sintomáticas. La misma consiste en actualizar las cifras de población de áreas pequeñas a través de información derivada de registros administrativos continuos como los vitales (nacimientos y defunciones), escolares (matrícula) y otros (permisos de construcción, registro de automóviles, consumo de energía eléctrica). Mediante esta metodología se pueden establecer correcciones a las tendencias poblacionales observadas contando con información reciente de las variables sintomáticas del crecimiento, esto es indirectamente asociadas al fenómeno en cuestión.

En este contexto, este informe expone los resultados de un ejercicio de estimación de la cantidad de habitantes por departamento de la República Argentina,² mediante diversos indicadores indirectos del tamaño de la población, poniendo en juego distintas opciones de la metodología de variables sintomáticas.

La sección siguiente constituye un repaso del proceso de recopilación de la información básica que se utilizó como insumo. En tal sentido, se plantean las limitaciones de los datos disponibles, los ajustes que se aplicaron a los mismos y las aclaraciones pertinentes sobre ciertas decisiones metodológicas adoptadas para solventar el desarrollo del ejercicio de estimación.

A continuación, se establece una breve descripción de los modelos utilizados. Se plantean sus respectivas fórmulas y se señalan sintéticamente las virtudes y limitaciones que portan. Acerca de los desarrollos matemáticos y las evaluaciones empíricas que los sustentan, se remite a literatura especializada preexistente.

Los resultados que se presentan en el informe son de dos tipos: evaluación entre alternativas metodológicas para estimar población por variables sintomáticas y análisis comparativo de las estimaciones por variables sintomáticas frente a las proyecciones oficiales formuladas por métodos matemáticos. En el primer caso, se examina la precisión relativa de cada alternativa metodológica frente a la población enumerada en el censo nacional de población y vivienda de 1991 (INDEC, 1992) a fin de detectar la opción más confiable. En tanto que la segunda parte, compara la estimación de la población por departamento de 1996 lograda con la alternativa más fiable frente a las “Estimaciones de la población por departamento” (INDEC, 1996).

Finalmente se sintetizan las principales conclusiones obtenidas mediante este ejercicio y se enuncian recomendaciones a fin de aprovechar esta experiencia para desarrollar un programa sistemático de aplicación de esta nueva metodología.

² La República Argentina hacia 1996 estaba constituida por 24 provincias (jurisdicciones principales) que contenían en su interior 509 departamentos (unidades espaciales menores).

I. Consideraciones acerca de la información básica

La recopilación de información básica se llevó a cabo 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). Sin embargo fueron escasas las fuentes que cumplieran con todos estos requisitos simultáneamente. El examen de los anuarios estadísticos argentinos de la década del '90 ofrece varios registros tomados a nivel departamental mas frecuentemente no se publican para todas las provincias y cuando se requiere componer los mismos antes de 1993 casi ninguno está disponible. La extensión en los últimos años de los registros tomados a nivel departamental permite especular con una mejora de las condiciones de aplicación de los métodos ensayados en el futuro, pero sus falencias en la década del '80 impidieron utilizarlos para este ejercicio.

A título de ejemplo, se encontró que una larga nómina de indicadores presentados en los anuarios estadísticos oficiales de la década del '90 estaban en el ámbito de departamentos pero no para todas las provincias. Entre ellos se cuentan: energía eléctrica facturada por categoría de usuario, líneas telefónicas instaladas por cada 100 habitantes, establecimientos hoteleros y viviendas construidas mediante programas provinciales.

Dadas las restricciones comentadas, se decidió utilizar los registros vitales (nacimientos y defunciones) y los registros escolares

(matrícula en educación primaria común). Entre ellos se encontraron algunos faltantes pero se consideró que con ciertos ajustes y estimaciones complementarias podrían ser utilizados para la metodología de variables sintomáticas.

Los ajustes que se establecieron en la información proveniente de las estadísticas vitales fueron ciertamente modestos, con supuestos de cumplimiento bastante seguro. En primer lugar, se tomó información inédita de nacimientos registrados por departamento de residencia de la madre y defunciones registradas por departamento de residencia del fallecido.³ Para que las cifras de áreas menores no estuvieran expuestas a oscilaciones aleatorias, se efectuó un suavizamiento mediante los datos del año necesario, del anterior y del posterior. Esta operación se llevó a cabo con datos del trienio 1990-92 y 1995-97.⁴

En cambio, para el año 1980 no se contaba con información de los años precedentes. Por tal motivo, se asumieron los totales provinciales publicados (PNES, 1984) y se los distribuyó respetando las proporciones de los departamentos según los datos disponibles.

Asimismo se estableció, para todos los años, una distribución proporcional de dos tipos de información espacial inespecífica: los eventos vitales de provincia conocida pero departamento ignorado (frecuentemente registrados en una provincia distinta a la de residencia) y los eventos vitales de provincia desconocida.

Las correcciones y estimaciones adicionales que se efectuaron con los datos de la matrícula escolar fueron mucho más severas. Dada la imposibilidad de rescatar información departamental para las fechas de interés se tomó la decisión de asumir más riesgos, ponderando la importancia de utilizar otra fuente adicional que complementara los resultados brindados por los registros vitales.

La información básica disponible era: total de alumnos matriculados en ciclo primario común por provincia para el período 1970-1988 y los años 1994, 1996 y 1997 y cantidad de alumnos matriculados por departamento para los años 1996 en adelante. La primera operación consistió en identificar el dato más aproximado para dar cuenta del año 1991. En tal sentido, se examinó la relación existente en 1980 entre la matrícula escolar del ciclo primario y la población que asistía al mismo nivel según el censo poblacional de la misma fecha. Dado que se apreció una relación regular,⁵ con pequeñas variaciones entre provincias, se aplicó la inversa de dicha relación como multiplicador para derivar la población matriculada a partir de la información censal de 1991.

Ya que se aplicó la misma relación observada en 1980 a los datos censales de 1991 se efectuó un examen de esa estimación. Al respecto, se comprobó que la matrícula escolar estimada de 1991 a partir de datos censales presentaba un valor consistente con los datos disponibles para la serie anual que se interrumpía en 1988 y se renovaba a partir de 1994.

³ La información de estadísticas vitales fue proporcionada por la Dirección de Estadística e Información de Salud (DEIS) del Ministerio de Salud de la Nación. En los casos en que no había información por departamento se recurrió a información disponible en las provincias que había sido recopilada por el Sistema de Estadísticas Sociodemográficas (SESD) del INDEC. La recopilación de registros administrativos fue posible gracias a Mercedes Fernández de la Dirección de Estadística e Información de Salud, y Elma Manino de la Dirección de Estadísticas Sectoriales del INDEC.

⁴ Cuando la información de los años disponibles presentaba un coeficiente de variabilidad menor al 20 % se adoptó el valor medio; en cambio, cuando las oscilaciones eran más importantes se tomó como representativo el valor mediano.

⁵ Sólo en el caso de una provincia se apreció una desviación importante en esta relación pero la misma pudo ser identificada como una irregularidad en la serie de los registros escolares. En tal caso, se sustituyó el valor de la matrícula de 1980 por una interpolación hecha con los datos de matrícula de 1978, 1979, 1981 y 1982.

Finalmente la información departamental se estimó para el año 1980 y el año 1991, a partir de aplicar las relaciones observadas entre matrícula escolar y población enumerada que asistía al nivel primario respectivas de cada provincia.

Otra cuestión a resolver fue el tratamiento de los cambios en las divisiones político administrativas durante el período del análisis. En principio, cabe aclarar que las modificaciones producidas en la división departamental entre 1980 y 1991 fueron escasas y muy localizadas. Después de 1991 se produjeron otros cambios, principalmente en la provincia de Buenos Aires, que afectaron a zonas densamente pobladas pero que tuvieron un impacto muy reducido en cantidad de unidades administrativas. A los efectos de trabajar con las mismas unidades espaciales para las dos estimaciones del ejercicio (la de 1991 y la de 1996) se decidió agrupar todas aquellas unidades que hubiesen sido creadas en el período o afectadas por desmembramientos, labor que concluyó en la reunión de veintiséis departamentos en nueve grupos.

Por otra parte, dado que muchos departamentos tienen escasa población y pocos eventos vitales se hallaban dificultades para darles un tratamiento similar a los restantes. En algunos de ellos se registraron menos de cinco nacimientos o defunciones (y aún ninguna) lo que fue motivo para agruparlos. El criterio que se siguió para estos agrupamientos fue respetar la contigüidad espacial y la homogeneidad demográfica. Estos agrupamientos para departamentos muy pequeños siempre se hicieron reuniendo departamentos de la misma provincia y contemplando afinidades en cuanto a tamaño poblacional, densidad (habitantes por kilómetro cuadrado), por ciento de población urbana y ritmo de crecimiento entre 1980 y 1991. En esta etapa, se reunieron diez departamentos en cuatro grupos.

A fin de tener una dimensión de los agrupamientos realizados, cabe puntualizar que las estimaciones se harán para cuatrocientas ochenta y seis áreas menores formadas a partir de quinientos nueve departamentos. De las áreas menores mencionadas, sólo en trece casos se trata de agrupamientos; esto representa una baja proporción de las áreas analizadas (menos de un tres por ciento).

La jurisdicción de ciudad de Buenos Aires (denominada Capital Federal hasta 1994) amerita un comentario aparte. Por su historia legal particular, no posee subdivisiones administrativas consolidadas en su interior. A razón de ello, los registros disponibles se encuentran presentados según distintas divisiones territoriales (básicamente distritos escolares y circunscripciones electorales) que no tienen afinidad entre sí. Teniendo en cuenta estas limitaciones, la ciudad de Buenos Aires ha sido tratada como una sola unidad espacial.⁶

Debido a las limitaciones encontradas y considerando las correcciones aplicadas, finalmente se trabajó con la siguiente información:

- población departamental enumerada en los censos poblacionales de 1980 y de 1991;
- población departamental estimada para el año 1996 por las técnicas de la Función Logística Modificada de Murphy y los Incrementos Relativos de Madeira y Simoes;⁷
- población total de la ciudad de Buenos Aires estimada para el año 1996;

⁶ Dado su tamaño y sus características demográficas singulares, esta jurisdicción no fue utilizada para calcular el modelo. En cambio, las estimaciones que se hicieron de esta área provinieron de aplicar la misma metodología que para las restantes áreas geográficas del estudio (departamentos y grupos).

⁷ La descripción de la metodología utilizada por INDEC para la estimación de la población por departamentos en el período 1990-2005 se puede consultar en Anexo I Notas Metodológicas de INDEC (1996).

- población de 5 años y más que asistía a nivel primario por departamento en censo poblacional de 1980 con las correcciones derivadas de las relaciones —en el ámbito total de cada provincia— obtenidas con matrícula escolar del ciclo primario común;
- población de 5 años y más que asistía a nivel primario por departamento en censo poblacional de 1991⁸ con las correcciones mencionadas para datos de 1980;
- matrícula escolar del ciclo primario común por departamento del año 1996;
- nacimientos por departamento de residencia de la madre y defunciones por departamento de residencia registrados en 1980;
- nacimientos por departamento de residencia de la madre y defunciones por departamento de residencia registrados en 1990, 1991, 1992, 1995, 1996 y 1997;
- tasa de crecimiento geométrico anual intercensal 1980-1991 por departamento (por mil personas);
- densidad de población por departamento en 1991 (en habitantes por kilómetro cuadrado);
- por ciento de población urbana por departamento en 1991.

⁸ Si bien la tabulación publicada indica en su título que se trata de población de 3 años y más, se comprobó que no había niños de menos de 5 años asistiendo al nivel primario.

II. Presentación de distintos modelos de aplicación de variables sintomáticas

En esta sección se hará una breve presentación de distintas propuestas metodológicas basadas en variables sintomáticas. El carácter de la misma es meramente descriptivo ya que explicar los supuestos y el origen de ellos excedería los alcances de este informe. De todos modos, estas alternativas ya han sido evaluadas durante décadas y existe una considerable bibliografía técnica que permite satisfacer estos intereses.⁹ La presentación general se apoyará en un artículo de Guiomar Bay (1998) donde se efectúa un examen comparativo de estas propuestas.

El método más sencillo es la *distribución por prorrateo*, que se basa en el supuesto de que la distribución de la población en un área mayor se distribuya entre sus áreas menores en forma idéntica a la distribución de una variable sintomática. La población del área menor i en el momento $t+n$ está expresada por:

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

⁹ Al respecto se podría consultar Crosetti y Schmitt (1954); Grier y Schmitt (1966); O'Hare (1976); Basavarajappa, Bender y Verma (1982); Swanson y Tedrow (1984); Teixeira Jardim (1992); Long (1993) y Bay (1998).

Donde:

- $t+n$ es el momento n años posteriores al último censo (momento t);
- $P_{(i,t+n)}$ población del área menor i en el momento $t+n$;
- $P_{(\bullet,t+n)}$ estimación disponible de la población del área mayor en el momento $t+n$;
- $S_{(i,t+n)}$ valor de la variable sintomática del área menor i en el momento $t+n$;
- $S_{(\bullet,t+n)}$ valor de la variable sintomática del área mayor en el momento $t+n$.

El método de *distribución proporcional* supone que la población varía durante un período en la misma proporción que lo hace la variable sintomática en el mismo período. La población del área menor i en el momento $t+n$ se deduce de esta relación contemplando un factor de ajuste (F_a) para que la suma de las áreas menores coincida con el área mayor:

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

Donde:

- t es el momento del último censo;
- $t+n$ es el momento n años posterior al último censo;
- $P_{(i,t+n)}$ población del área menor i en el momento $t+n$;
- $P_{(i,t)}$ población del área menor i en el momento t ;
- $S_{(i,t)}$ valor de la variable sintomática del área menor i en el momento t ;
- $S_{(i,t+n)}$ valor de la variable sintomática del área menor i en el momento $t+n$;
- F_a factor de ajuste entre suma de áreas menores estimadas y área mayor.

El método de las *tasas vitales de Bogue* consiste en estimar la población de áreas menores utilizando solamente datos de los registros vitales.¹⁰ Bogue (1950) tomó como supuesto la existencia de una relación inversa entre las variaciones anuales de las tasas de natalidad y de mortalidad de las áreas menores y que el desvío de las tasas propias en relación a las tasas del área mayor permanece constante a través del tiempo. La expresión del modelo es:

$$P_{(i,t+n)} = a \times \frac{B_{(i,t+n)}}{b_{(i,t+n)}} + (1-a) \times \frac{D_{(i,t+n)}}{d_{(i,t+n)}}$$

Donde:

- a es el coeficiente de ponderación entre las dos estimaciones;

¹⁰ La descripción de este método se basa en Teixeira Jardim (1992).

$$b_{(i,t+n)} = [(B_{(i,t)} / P_{(i,t)}) / (B_{(\bullet,t)} / P_{(\bullet,t)})] * (B_{(\bullet,t+n)} / P_{(\bullet,t+n)})$$

$$d_{(i,t+n)} = [(D_{(i,t)} / P_{(i,t)}) / (D_{(\bullet,t)} / P_{(\bullet,t)})] * (D_{(\bullet,t+n)} / P_{(\bullet,t+n)})$$

B es la cantidad de nacidos vivos;

b es la tasa bruta de natalidad;

D es la cantidad de defunciones;

d es la tasa bruta de mortalidad.

El método de *correlación de razones* se aplica asumiendo que la evolución de la población está correlacionada con la variación de una o más variables sintomáticas. La correlación se estima por medio de un modelo de regresión, en tanto que la variación se establece por la razón entre los valores de una variable correspondientes a dos fechas. El modelo asume la siguiente forma:

$$Y_{(i,t)} = a_{(0)} + a_{(1)} * X1_{(i,t)} + \dots + a_{(n)} * Xn_{(i,t)} + e_{(i)}$$

Donde:

$$Y_{(i,t)} = \frac{P_{(i,t)} / P_{(\bullet,t)}}{P_{(i,0)} / P_{(\bullet,0)}}$$

$$Xj_{(i,t)} = \frac{Sj_{(i,t)} / Sj_{(\bullet,t)}}{Sj_{(i,0)} / Sj_{(\bullet,0)}}$$

0 es el momento del penúltimo censo;

t es el momento del último censo;

$Y_{(i,t)}$ razón entre la proporción del área menor i en los momentos 0 y t ;

$Xj_{(i,t)}$ razón entre la proporción de la variable sintomática j del área menor i en los momentos 0 y t ;

$e_{(i)}$ error de la estimación según el modelo.

Finalmente la población del área menor i puede ser estimada en el momento $t+n$, a partir de la aplicación del modelo de regresión a las razones de variación de las variables sintomáticas entre el momento t y $t+n$, es decir que:

$$P_{(i,t+n)} = Y_{(i,t+n)} \times \frac{P_{(i,t)}}{P_{(\bullet,t)}} \times P_{(\bullet,t+n)}$$

Donde:

$Y_{(i,t+n)}$ razón de variación de la población del área menor i estimada para el período $t, t+n$ por el modelo de regresión.

El método de *correlación de diferencias* es semejante al anterior. La diferencia estriba en que la variación se establece por la diferencia entre los valores de una variable correspondientes a dos fechas. El modelo asume la siguiente forma:

$$W_{(i,t)} = a_{(0)} + a_{(1)} * Z1_{(i,t)} + \dots + a_{(n)} * Zn_{(i,t)} + e_{(i)}$$

Donde:

$$W_{(i,t)} = [P_{(i,t)} / P_{(\bullet,t)}] - [P_{(i,0)} / P_{(\bullet,0)}]$$

$$Zj_{(i,t)} = [Sj_{(i,t)} / Sj_{(\bullet,t)}] - [Sj_{(i,0)} / Sj_{(\bullet,0)}]$$

$W_{(i,t)}$ diferencia entre la proporción del área menor i en los momentos 0 y t ;

$Zj_{(i,t)}$ diferencia entre la proporción de la variable sintomática j del área menor i en los momentos 0 y t ;

$e_{(i)}$ error de la estimación según el modelo.

Finalmente la población del área menor i en el momento $t+n$ será:

$$P_{(i,t+n)} = [W_{(i,t+n)} + (P_{(i,t)} / P_{(\bullet,t)})] * P_{(\bullet,t+n)}$$

Donde:

$W_{(i,t+n)}$ diferencia de proporciones del área menor i estimada para el período $t, t+n$ por el modelo de regresión

El método de *correlación de tasas* es otra variante del modelo de correlación de razones.¹¹ En este caso, se establece una aproximación exponencial del cambio y, en tal sentido, se toma el logaritmo natural de cada razón y se lo divide por el tiempo transcurrido entre los censos tomados como referencia. En estos casos, el modelo asume la siguiente forma:

$$Y^*_{(i,t)} = a_{(0)} + a_{(1)} * X^*1_{(i,t)} + \dots + a_{(n)} * X^*n_{(i,t)} + e_{(i)}$$

Donde:

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

$$X^*j_{(i,t)} = \frac{\ln Xj_{(i,t)}}{k}$$

k es el tiempo transcurrido entre las dos fechas de referencia.

¹¹ La descripción del método de correlación de tasas se basa en Chaves Esquivel (s/f).

Finalmente la población del área menor i puede ser estimada en el momento $t+n$, a partir de la aplicación del modelo de regresión a las tasas de variación de las variables sintomáticas entre el momento t y $t+n$, es decir que:

$$P_{(i,t+n)} = e^{[Y^{*(i,t+n)*k}]} \times \frac{P_{(i,t)}}{P_{(\bullet,t)}} \times P_{(\bullet,t+n)}$$

Donde:

$e^{[Y^{*(i,t+n)*k}]}$ tasa de variación exponencial de la proporción del área menor i estimada para el período $t, t+n$ por el modelo de regresión si hubiese transcurrido k años entre t y $t+n$.

Finalmente resta comentar que los tres últimos métodos, basados en correlaciones, pueden ser refinados en la medida en que se utilicen variables complementarias para distinguir estratos de áreas. De esta manera, se mejoran las estimaciones obtenidas por un modelo único para todas las divisiones geográficas en cuestión.

III. Resultados de aplicación de variables sintomáticas

El examen del método de variables sintomáticas se desarrollará en dos aspectos. La primera parte de esta sección, ilustra acerca de la precisión relativa de varios modelos basados en diversas variables sintomáticas y bajo supuestos disímiles sobre la relación entre ellas y el crecimiento poblacional. En tal sentido, se tomó como referencia para sopesar la confiabilidad de cada modelo la población enumerada por departamento en el censo nacional de población y vivienda de 1991.

En tanto que la segunda parte, contiene los hallazgos más destacados de una comparación entre los resultados obtenidos por el método de variables sintomáticas y las estimaciones de población oficiales, producidas sobre la base de métodos matemáticos. A tal efecto, se comparan las proyecciones departamentales del año 1996 con las estimaciones generadas mediante el seguimiento de variables sintomáticas con posterioridad al último censo.

A. Examen de la precisión de los modelos frente a datos censales de 1991

Según se había anunciado anteriormente, las variables sintomáticas que se utilizaron para las estimaciones provenían de dos fuentes: registros vitales y registros educativos. En consecuencia, las distintas estimaciones de población ensayadas se derivaron de:

- matrícula en ciclo primario común;
- nacimientos;
- defunciones generales;
- eventos vitales (nacimientos más defunciones generales).

En tanto que los métodos que se aplicaron fueron los seis descriptos en la sección previa. Adicionalmente se establecieron algunas modificaciones a ellos como promedios de estimaciones o alternativas a las ponderaciones de cada componente según se aclara oportunamente. Los métodos que se aplicaron fueron:

- distribución por prorrateo;
- distribución proporcional;
- método de tasas vitales;
- correlación de razones;
- correlación de diferencias;
- correlación de tasas.

Por otra parte, con el propósito de refinar los modelos de correlación se emplearon tres variables complementarias para estratificar los departamentos. En todos los casos, se construyeron tres niveles y para definir los umbrales se buscó que cada categoría reuniera la misma cantidad de unidades. A continuación se presentan los criterios de estratificación asociados a cada variable:

Variable de estratificación	Definición operacional	Estrato	Definición de estrato
Crecimiento 1980-1991	Tasa de crecimiento geométrico anual (por mil)	Crecimiento bajo	Menos de 6.4
		Crecimiento medio	De 6.4 a 16.3
		Crecimiento alto	Más de 16.3
Densidad poblacional 1991	Cantidad de habitantes por kilómetro cuadrado	Densidad baja	Menos de 3
		Densidad media	De 3 a 12.9
		Densidad alta	Más de 12.9
Urbanización en 1991	Por ciento de población en aglomerados de 2000 habitantes o más	Urbanización baja	Menos de 51.3
		Urbanización media	De 51.3 a 76.5
		Urbanización alta	Más de 76.5

Para examinar la justeza de los modelos se tuvo en cuenta el error absoluto porcentual (EAP), definido como:

$$\text{EAP} = \text{Absoluto de } [(\text{Pob. estimada} - \text{Pob. censada}) / \text{Pob. censada}] * 100$$

Obtenido este elemento de comparación en el ámbito de cada unidad espacial, se formularon indicadores agregados para disponer de medidas de resumen. Los que se utilizan en el documento fueron derivadas de los EAP:

- media aritmética de los EAP;
- desviación estándar de los EAP;
- porcentaje de departamentos con EAP inferior a 10%;
- valor de EAP correspondiente al percentil 80;
- valor mediano de EAP (percentil 50);
- valor de EAP correspondiente al percentil 5;
- valor de EAP correspondiente al percentil 95.

En primer lugar, podemos observar la precisión relativa de los métodos más sencillos, vale decir aquellos que no se basan en un modelo de correlación lineal. En el cuadro 1 se aprecian los indicadores pertinentes para expresar la justeza de cada estimación.

Cuadro 1
ANÁLISIS DEL ERROR DE LAS ESTIMACIONES DE POBLACIÓN POR DEPARTAMENTO
SEGÚN LOS MÉTODOS DE PRORRATEO, DE DISTRIBUCIÓN PROPORCIONAL Y DE TASAS VITALES
ARGENTINA, 1991

Método	Modelo	EAP promedio	Desviac. estándar	% de EAP >10	EAP perc .80	EAP mediano	EAP perc .05	EAP perc .95
Prorrateo	M. Matrícula	21.4	15.5	72.9	35.5	18.0	2.2	47.9
	N. Nacimientos	25.6	19.8	74.9	41.3	21.6	2.3	61.5
	D. Defunciones	22.0	15.6	72.9	35.0	20.1	1.8	49.9
	E. Eventos Vitales	15.9	14.2	55.9	26.3	11.9	1.0	41.5
	M-N. Promedio de M y N	22.6	16.3	73.9	36.7	19.0	3.0	52.8
	M-D. Promedio de M y D	9.2	8.5	32.6	14.2	7.1	0.8	25.2
	M-E. Promedio de M y E	17.1	13.7	61.4	29.0	13.5	1.5	41.4
	N-D. Promedio de N y D	11.7	10.8	45.6	17.6	9.1	0.9	31.4
Distribución proporcional	M-N-D. Prom. de M, N y D	11.7	10.7	45.6	19.5	8.6	0.6	31.2
	M. Matrícula	7.4	7.0	27.3	11.4	6.1	0.5	19.0
	N. Nacimientos	15.7	16.2	53.6	23.8	10.9	1.2	44.7
	D. Defunciones	20.7	18.2	69.2	32.4	16.2	1.5	51.1
	E. Eventos Vitales	14.3	13.4	55.4	21.6	10.9	1.3	40.1
	M-N. Promedio de M y N	9.0	8.3	32.4	14.1	7.0	0.8	25.3
	M-D. Promedio de M y D	12.0	9.4	50.7	18.0	10.3	1.3	28.7
	M-E. Promedio de M y E	8.6	7.2	31.6	12.7	7.1	0.5	23.1
Método de tasas vitales	N-D. Promedio de N y D	14.7	13.8	55.0	21.7	11.4	1.4	38.4
	M-N-D. Prom. de M, N y D	10.6	9.3	43.5	15.7	9.0	0.6	27.8
	Tb. Estimación por Natalidad	15.7	16.2	53.6	23.8	10.9	1.2	44.7
	Td. Estimación por Mortalidad	20.7	18.2	69.2	32.4	16.2	1.5	51.1
	Tbde Tb y Td equiponderadas ¹²	14.7	13.9	55.2	21.8	11.4	1.4	38.9
Tbdv Tb y Td con distinto peso ¹²	14.3	13.7	54.8	21.4	10.9	1.3	38.2	

¹² Para asignarle peso a la estimación proveniente de las TBN y de las TBM se tuvo en cuenta para cada caso la inversa del coeficiente de variabilidad de la respectiva tasa. Sobre el fundamento de esta alternativa ver Teixeira Jardim (1992).

En líneas generales, se advierte que las aproximaciones más adecuadas se obtienen por el método de distribución proporcional y las menos precisas por el de prorrateo. Con todo, este último se caracteriza por presentar situaciones marcadamente distintas según las variables sintomáticas puestas en juego. En efecto, el modelo M-D por prorrateo presenta uno de los menores valores de EAP promedio, siendo el modelo N de este mismo método el portador de las mayores inexactitudes.

El método de las tasas vitales representa una situación intermedia frente a los otros métodos examinados. En sus distintas opciones, se comprueba que las estimaciones por la tasa bruta de natalidad o por la tasa de mortalidad son de inferior precisión a una combinación de ambas. En tanto que a la hora de conformar una estimación combinada, la alternativa de ponderar diferencialmente las estimaciones originadas por una u otra tasa mejora la aproximación obtenida con una ponderación equivalente, aunque la precisión ganada no es tan significativa.

En cuanto al desempeño de las variables sintomáticas, se advierte en general que son más precisas las estimaciones derivadas de promedios entre modelos de variables utilizadas individualmente. Esta superioridad de los modelos combinados es más destacada entre las variantes del método de prorrateo.

Analizadas las variables sintomáticas en forma singular, no se observan regularidades sostenidas a través de los diversos métodos. En tal sentido, las estimaciones M (por matrícula escolar) tanto resulta ser la más adecuada con el método de distribución proporcional como una de las menos precisas con el de prorrateo. Las mejores estimaciones derivadas de los registros escolares suelen presentarse cuando se combina la matrícula escolar con la cantidad de algún registro vital.

En cuanto al uso de los registros vitales, los resultados justificarían tratarlos como una suma de nacimientos y defunciones (eventos vitales) antes que como alguno de sus componentes por separado. Mas el método de tasas vitales propuesto por Bogue constituye una opción bastante razonable si sólo se cuenta con este tipo de registros.

Los modelos de correlación, basados en información de un período suelen presentar mayores niveles de precisión para estimar el tamaño de la población. A efectos de incrementar la calidad de los modelos, se combinaron distintas variables sintomáticas en modelos de regresión lineal multivariados.

La lectura del cuadro 2 nos confirma que sólo algunos modelos superan un EAP promedio de diez por ciento, mientras que los más efectivos tienen un promedio cercano a la mitad de dicho valor. El contraste entre los métodos correlacionales y los presentados previamente puede dimensionarse cuando se analiza cuál es la proporción de departamentos con una valor de EAP superior al diez por ciento; mientras en los modelos basados en correlaciones estos errores sólo afectan a un quinto o un cuarto de las unidades, en los anteriores estas imprecisiones afectaban a más de la mitad y en ocasiones, a tres cuartas partes de ellas.

Entre los distintos métodos de correlación que se ensayaron el de diferencias fue el que presentó la menor precisión. Con todo, en muchos de los modelos diseñados con esta técnica se dieron valores promedio de EAP inferiores al diez por ciento. Esto ocurrió particularmente con los modelos que incluían como insumo la matrícula escolar, esto es M, M-N, M-D, M-E y M-N-D.

Los métodos de correlación de razones y de tasas tuvieron un desempeño bastante similar. Si bien el segundo de ellos fue concebido para refinar el primero, no se apreciaron diferencias sustantivas en la precisión de ambos. Con valores promedio de EAP semejantes al método de razones, el método de correlación de tasas tiene una menor desviación típica de los EAP aunque cabría puntualizar que esta diferencia tampoco es de gran magnitud.

Cuadro 2

**ANÁLISIS DEL ERROR DE LAS ESTIMACIONES DE POBLACIÓN POR DEPARTAMENTO
SEGÚN LOS MÉTODOS DE CORRELACIÓN DE RAZONES, DE DIFERENCIAS Y DE TASAS.
ARGENTINA, 1991**

Método	Modelo	EAP promedio	Desviac. estándar	% de EAP >10	EAP perc .80	EAP mediano	EAP perc .05	EAP perc .95
Correlación de razones	M. Matrícula	5.8	5.1	15.0	9.0	4.9	0.4	13.7
	N. Nacimientos	9.5	9.9	31.8	14.6	6.8	0.7	27.6
	D. Defunciones	10.1	9.4	38.4	15.2	7.2	0.9	28.5
	E. Eventos Vitales	9.4	9.1	35.5	14.4	7.0	0.7	25.5
	M-N. Combinac. lineal de M y N	5.7	4.9	14.6	8.8	4.4	0.5	13.5
	M-D. Combinac. lineal de M y D	5.8	5.1	14.6	9.1	4.8	0.4	13.4
	M-E. Combinac. lineal de M y E	5.8	4.8	15.6	8.9	4.7	0.6	14.0
	N-D. Combinac. lineal de N y D	9.5	9.8	32.0	14.5	6.9	0.7	27.7
	M-N-D. Comb. lineal de M, N y D	5.6	5.0	14.6	8.7	4.6	0.5	13.9
Correlación de diferencias	M. Matrícula	7.8	7.8	24.6	11.8	5.8	0.4	21.2
	N. Nacimientos	13.4	13.1	50.5	19.2	10.2	0.9	41.3
	D. Defunciones	14.1	14.0	52.2	20.2	10.5	1.0	42.1
	E. Eventos Vitales	13.6	13.7	51.3	19.1	10.2	1.0	41.8
	M-N. Combinac. lineal de M y N	7.7	7.9	24.4	11.6	5.6	0.5	21.5
	M-D. Combinac. lineal de M y D	7.7	7.7	26.5	11.6	5.8	0.6	21.4
	M-E. Combinac. lineal de M y E	7.7	7.8	24.6	11.5	5.7	0.5	21.5
	N-D. Combinac. lineal de N y D	13.5	13.3	51.3	19.3	10.4	0.9	41.7
	M-N-D. Comb. lineal de M, N y D	7.9	8.0	26.3	11.7	5.7	0.5	21.5
Correlación de tasas	M. Matrícula	5.8	5.0	15.0	9.1	4.8	0.4	13.8
	N. Nacimientos	9.5	9.0	33.3	14.4	7.0	0.7	25.7
	D. Defunciones	10.2	9.2	38.2	16.0	7.5	0.8	28.6
	E. Eventos Vitales	9.3	8.6	35.3	13.6	7.0	0.8	24.8
	M-N. Combinac. lineal de M y N	5.6	4.7	13.3	8.7	4.7	0.4	13.7
	M-D. Combinac. lineal de M y D	5.8	5.0	15.0	9.1	5.0	0.5	13.9
	M-E. Combinac. lineal de M y E	5.7	4.7	15.0	8.9	4.8	0.4	14.0
	N-D. Combinac. lineal de N y D	9.3	8.9	33.1	13.5	7.1	0.5	24.4
	M-N-D. Comb. lineal de M, N y D	5.6	4.7	13.3	8.7	4.8	0.4	13.6

La comparación entre variables sintomáticas reafirma una cuestión señalada anteriormente: las combinaciones entre variables sintomáticas refuerzan la solidez de los modelos. Al combinar las variables sintomáticas en regresiones múltiples se redujo en forma importante el nivel de error en las estimaciones. La expresión más acabada de este señalamiento se encuentra en el modelo que combina los tres datos básicos (el MND) ya que resulta ser la alternativa más precisa cuando se aplican los métodos de correlación de razones y de tasas. La precisión de este modelo combinado lleva a que el noventa y cinco por ciento de las unidades presentan un EAP inferior al quince por ciento.

El funcionamiento individual de las variables sintomáticas es bastante regular entre los distintos métodos. En dicho contexto, las estimaciones derivadas de la matrícula escolar (modelos M) son más precisas que las constituidas por los registros vitales. Justamente la combinación que adolece del dato de matrícula escolar (modelo N-D) resulta ser el modelo multivariado menos ajustado.

Entre los registros vitales, no se aprecian diferencias importantes en la precisión obtenida mediante el uso de registros de nacimientos o de defunciones; al mismo tiempo, la agregación de ambos datos en un total de eventos vitales no genera cambios significativos en la precisión de las estimaciones. Otro tanto puede afirmarse al comparar los modelos simples basados en la suma de registros vitales y los modelos multivariados que combinan cantidad de nacimientos y defunciones

registrados; en efecto, el rendimiento varía según el método adoptado pero a través de ellos se confirma que la precisión obtenida por la simple suma es equivalente a la de un modelo de combinación lineal.

En la búsqueda de formular modelos de correlación aún más refinados para estimar los cambios en el tamaño de la población, se apeló a la estratificación de los departamentos. Con esta operación se intentó incrementar la homogeneidad del comportamiento de las unidades espaciales identificando distintos contextos. Según se definió previamente se emplearon tres criterios para estratificar: crecimiento poblacional entre los últimos censos, densidad poblacional y nivel de urbanización.

Los tres métodos de correlación examinados fueron aplicados con el modelo MND, que resultó ser la combinación más precisa, constituyendo ecuaciones específicas para cada estrato. Vale decir que se conformaron tres ecuaciones, uno por estrato, para cada método y esta operación se reiteró con cada variable estratificadora. En el cuadro 3 se presentan los resultados.

Cuadro 3

ANÁLISIS DEL ERROR DE LAS ESTIMACIONES DE POBLACIÓN POR DEPARTAMENTO DEL MODELO MND SEGÚN LOS MÉTODOS DE CORRELACIÓN DE RAZONES, DE DIFERENCIAS Y DE TASAS, SEGÚN DISTINTOS CRITERIOS DE ESTRATIFICACIÓN. ARGENTINA, 1991

Criterio de estratificación	Método	EAP promedio	Desviac. estándar	% de EAP >10	EAP perc .80	EAP mediano	EAP perc .05	EAP perc .95
Crecimiento de población 1980-1991	Corr. de razones	4.6	5.0	9.0	6.4	3.3	0.2	13.3
	Corr. de diferencias	15.4	22.0	41.1	23.0	7.5	0.4	58.6
	Correlación de tasas	3.8	3.9	5.7	5.3	2.8	0.3	10.4
Densidad poblacional en 1991	Corr. de razones	8.1	7.2	30.2	12.3	6.6	0.6	19.7
	Corr. de diferencias	6.5	5.5	19.7	9.9	5.3	0.4	15.7
	Correlación de tasas	5.8	4.6	14.6	8.9	4.8	0.6	13.3
Nivel de urbanización en 1991	Corr. de razones	8.4	7.4	32.4	13.0	6.6	0.7	20.0
	Corr. de diferencias	7.6	7.9	25.7	11.8	5.5	0.5	20.8
	Correlación de tasas	5.7	4.6	14.4	8.5	4.4	0.7	13.6

La estratificación de los departamentos para formular nuevos modelos de correlación no resultó sistemáticamente más precisa que las correlaciones analizadas en el paso anterior. Se produjeron resultados diferenciados según la variable empleada para estratificar.

Los modelos de correlación estratificada más precisos se dieron cuando se tomó el ritmo de crecimiento intercensal como elemento de corte. Particularmente ajustadas fueron las estimaciones generadas por el método de correlación de razones y el de tasas. Este último se expresó, según lo previsto teóricamente, como una alternativa más precisa al primero en tanto tuvo menor valor medio de EAP (inferior al cuatro por ciento) y una menor desviación con respecto a esta cifra. La prueba más interesante del ajuste de este modelo es que el noventa y cinco por ciento de los departamentos presentaron un valor de EAP cercano al diez por ciento o inferior.

Sin embargo, la misma variable de estratificación derivó en un modelo ostensiblemente impreciso cuando se aplicó el método de correlación de diferencias. El nivel de imprecisión de esta combinación es aún superior a todos los modelos de correlación aplicados sin estratificar.

Las otras variables de estratificación, el nivel de urbanización y la densidad espacial, dieron lugar a estimaciones aceptables pero no mejoraron la precisión que ya se había obtenido trabajando con todas las unidades espaciales sin diferenciar. En estos casos, se confirmó la mayor exactitud

del método de correlación de tasas y, contrariamente a lo observado antes de estratificar, la correlación de diferencias tuvo mayor precisión que la de razones.

B. Comparación entre estimación por variables sintomáticas y proyecciones poblacionales por métodos matemáticos hacia 1996

Los resultados de la primera parte indicaron que los modelos correlacionales fueron más precisos que los más simples. Asimismo, se encontró que de las alternativas correlacionales la más ajustada fue la que se basa en correlaciones de tasas y que el modelo más afinado (MND) integraba la información básica de tres variables sintomáticas: la cantidad de alumnos en nivel primario según la matrícula escolar, la cantidad de nacimientos y la cantidad de defunciones obtenidos de los registros vitales. Por último, se advirtió que la precisión del modelo MND mediante correlación de tasas era incrementada si se estratificaba a los departamentos según la tasa de crecimiento poblacional entre los dos últimos censos. En consecuencia, se asumió que ésta era la alternativa metodológica más confiable para estimar población por departamento con las variables sintomáticas disponibles.

Dados estos antecedentes, se desarrolló esta segunda parte del ejercicio. En este caso, se comparó la estimación por variables sintomáticas más ajustada con las proyecciones de población disponibles para un año reciente, 1996. En este caso, no se cuenta con un censo de población que permita dirimir la mayor o menor precisión de ambas estimaciones, por lo tanto se hará una comparación para identificar el grado de similitud entre ellas y reconocer las unidades espaciales donde se producen mayores discrepancias.

En este caso, la comparación se llevará a cabo por medio de la diferencia absoluta porcentual (DAP), que se define como:

$$\text{DAP} = \text{Absoluto de } [(\text{Estimación V. Sintomáticas} - \text{Estimación Matem.}) / \text{Estimación Matem.}] * 100$$

Nuevamente se generarán medidas resumen, de contenido similar a las que se utilizaron en la sección anterior: media aritmética, desviación estándar, porcentaje de departamentos con DAP inferior a 10%, valor de DAP correspondiente al percentil 80, valor mediano de DAP (percentil 50), valor de DAP correspondiente al percentil 5 y valor de DAP correspondiente al percentil 95.

Cuadro 4

COMPARACIÓN ENTRE ESTIMACIÓN DE POBLACIÓN POR DEPARTAMENTO POR MÉTODOS MATEMÁTICOS (INDEC, 1996) Y POR MÉTODO DE CORRELACIÓN DE TASAS. ARGENTINA, 1996

DAP promedio	4.3
Desviación Estándar	4.9
% de DAP >10	9.2
Valor de EAP percentil .80	6.2
Valor mediano de EAP	2.6
Valor de EAP percentil. 05	0.2
Valor de EAP percentil .95	14.0

Nota: Mediante el uso de las variables sintomáticas matrícula escolar, nacimientos y defunciones generales (MND), estratificando las unidades por nivel de crecimiento entre 1980 y 1991.

La observación del cuadro 4 demuestra que a grandes rasgos las dos estimaciones tienen valores equivalentes en cuanto al tamaño de la población por departamento. La diferencia promedio es algo inferior al cinco por ciento. Asimismo, la gran mayoría de las unidades, el noventa y cinco por ciento de los Departamentos, presentan diferencias entre las estimaciones que no superan el quince por ciento.

Empero la comparación entre ambas metodologías puede resultar más fructífera si se apunta a las áreas en las que se encontraron las discrepancias más importantes. Al respecto se presenta el cuadro 5 donde aparece la descripción de los departamentos en donde se hallaron las DAP más importantes, contemplando en este caso el signo de las diferencias.

Cabe aclarar que este análisis ameritaría una revisión minuciosa de cada situación local más la entrevista a informantes claves vinculados a la recolección de información en puestos descentralizados. Con las restricciones del caso, se puede corroborar que en los casos presentados la simple extrapolación de los datos poblacionales de los últimos censos (1980 y 1991) puede llevar a resultados diferentes a los que se tendrían con información poscensal proveniente de registros administrativos.

En los primeros departamentos, el signo positivo de la diferencia porcentual indica que la estimación matemática es superior a la proveniente de variables sintomáticas. Se trata en varios casos de departamentos con importante crecimiento poblacional en el último período intercensal, pero donde la evolución de la matrícula escolar y los registros vitales en el quinquenio posterior al censo ha descendido o bien tuvo un ligero aumento.

El segundo grupo de departamentos ha tenido una estimación poblacional más grande con las variables sintomáticas que con el método matemático. En varios de ellos, dada la tendencia del último período intercensal se esperaba una reducción poblacional que los registros administrativos (matrícula escolar y eventos vitales) parecieran contradecir con su incremento.

Cuadro 5

**COMPARACIÓN ENTRE VALORES DE ESTIMACIÓN DE POBLACIÓN POR DEPARTAMENTO
POR MÉTODOS MATEMÁTICOS (MATEMAT) Y POR MÉTODO DE VARIABLES SINTOMÁTICAS (SINTOM).
DEPARTAMENTOS SELECCIONADOS. ARGENTINA, 1996**

Departamento	Censo 1991	Estimac. Matemat. 1996	Estim. Sintom. 1996	Matrícula escolar		Nacimientos		Defunciones		Eventos vitales		DAP
				M91	M96	N91	N96	D91	D96	E91	E96	
Collón Curá	7865	11658	6638	1247	855	184	124	18	21	202	145	43,1
General Lavalle	3046	4030	2636	485	417	37	17	21	17	59	35	34,6
Catan Lil	2408	2689	2255	469	400	69	58	14	13	83	70	16,1
Capital	218685	258932	219119	41381	33523	6078	6525	1311	1525	7389	8049	15,4
Iguazú	57623	72647	61505	12754	12043	1993	1885	262	245	2255	2130	15,3
Lib. General San Martín	37722	41804	35663	8521	6449	1095	1012	164	191	1258	1203	14,7
San Antonio	24216	26550	22888	4707	3614	545	509	112	108	657	617	13,8
Chimbás	52263	61550	53736	10299	8813	1459	1539	285	313	1744	1852	12,7
Moreno	21746	19936	23970	4419	5266	594	607	166	145	760	751	-20,2
Alberdi	10633	10042	12188	2225	2921	284	335	62	70	346	405	-21,4
Susques	2846	3181	3862	610	771	116	145	27	33	143	178	-21,4
General Güemes	43762	47380	58182	8873	12192	1698	1751	372	353	2070	2103	-22,8
Paso de los Indios	1883	1733	2139	290	342	51	51	16	14	67	65	-23,4
Gastre	1900	1758	2172	224	270	21	28	12	6	34	35	-23,6
Capital	121620	123046	152061	13844	25004	2414	2195	1121	1095	3535	3290	-23,6
Rivadavia	20992	22645	28035	4436	5396	423	675	92	120	515	795	-23,8
Pehuenches	6538	8248	10494	1109	1935	236	341	33	35	269	376	-27,2

IV. Conclusiones y recomendaciones

Las conclusiones más importantes que pueden destacarse de este ejercicio, derivadas de comparaciones entre metodologías de estimación de población, son dos:

- 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;
- la aplicación del método de variables sintomáticas es una opción factible para examinar proyecciones poblacionales basadas en modelos matemáticos.

La primera conclusión, derivada de comparar diversos modelos de estimación de población por variables sintomáticas frente a la población enumerada por un censo, puede ser desglosada en varios apuntes adicionales:

- 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 prorrato, distribución proporcional y método de las tasas vitales);
- entre los métodos de estimación que emplean variables sintomáticas 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 ya que con un valor de error promedio al método de correlación de razones tuvo menor desviación con respecto a dicho valor;
- 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.

En el sentido de la segunda conclusión puede agregarse que:

- los métodos de 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;
- la aplicación del método de variables sintomáticas es una herramienta apropiada 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.

Por último, el desarrollo de este ejercicio ha generado una experiencia que podría replicarse con nuevas cifras de registros administrativos y constituirse en un programa sistemático de investigación sobre las tendencias demográficas de áreas espaciales menores. A fin de extender esta práctica sería importante examinar las siguientes recomendaciones:

- los registros administrativos localizados deberían ser fortalecidos tanto en su cobertura como en su oportunidad, en algunos campos se han hecho progresos importantes en los últimos años que requieren ser sostenidos en el tiempo;
- 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;
- 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;
- este trabajo constituye una primera experiencia de una línea de trabajo que debería profundizarse haciendo un relevamiento aún más minucioso de las variables sintomáticas disponibles y explorando nuevas alternativas metodológicas;

- un examen más atinado de los modelos podría lograrse cuando se concrete el censo nacional de población, viviendas y hogares que se ha programado en la Argentina para octubre del año 2001;
- el método de variables sintomáticas podría ser una herramienta eficaz para ajustar cuestiones logísticas del censo poblacional en preparación puesto que, mediante registros administrativos recientes, podría ofrecer estimaciones de población actualizadas.

Bibliografía

- Basavarajappa, K. G., Rosemary Bender y Ravi B.P. Verma (1982), “New approaches to methods of estimating the population of census divisions”, Ottawa, Demography Division, Statistics Canada.
- Bay, Guiomar (1998), “El uso de variables sintomáticas en la estimación de la población de áreas menores”, Revista *Notas de Población* N° 67-68, CELADE, Santiago.
- Bogue, Donald (1950), “A technique for making extensive population estimates”, *Journal of the American Statistical Association* N° 45.
- Crosetti, A. H. y R. C. Schmitt (1954), “Accuracy of the ratio-correlation method for estimating postcensal population”, *Land Economics* N° 30.
- Chaves Esquivel, Edwin A. (s/f), *Variables sintomáticas en las estimaciones poblacionales a nivel cantonal* (mimeo).
- Grier, John M. y R. C. Schmitt (1966), “A method of estimating the population minor civil divisions”, *Rural Sociology* N° 31.
- INDEC (Instituto Nacional de Estadística y Censos) (1996), *Estimaciones de la población por departamento. Período 1990-2005*, Serie Análisis Demográfico N° 8, INDEC, Buenos Aires.
- _____(1995), Anuario Estadístico de la República Argentina 1995, CD-ROM.
- _____(1992), Censo Nacional de Población y Vivienda 1991, INDEC, Buenos Aires.
- _____(1991), *Censo Nacional de Población y Vivienda 1991. Resultados provisionales*, INDEC, Buenos Aires.
- INDEC-CELADE (Instituto Nacional de Estadística y Censos – Centro Latinoamericano y Caribeño de Demografía) (1995a), “Estimaciones y proyecciones de población. Total del país (versión revisada) 1950-2050”, Serie Análisis Demográfico N° 5, INDEC, Buenos Aires.
- _____(1995b), *Proyecciones de población por sexo y grupos de edad: urbana, rural y económicamente activa 1990-2025 y por provincia 1990-2010 (versión revisada)*”, Serie Análisis Demográfico N° 7, INDEC, Buenos Aires.

- Long, John (1993), "Methods for postcensal population estimation: flow methods vs. stock methods", *International Union for the Scientific Study of Population. International Population Conference, 1993*, Liege, Union Internacional para el Estudio Científico de la Población (UIECP).
- O'hare, William P. (1976), "Report of a multiple regression method for making population estimates", *Demography* N° 7
- PNES (Programa Nacional de Estadísticas de Salud) (1993), *Estadísticas vitales. Información básica*, Serie Estadísticas Vitales N° 35, Ministerio de Salud y Acción Social, Buenos Aires.
- _____(1992), *Estadísticas vitales. Información básica*, Serie Estadísticas Vitales N° 34, Ministerio de Salud y Acción Social, Buenos Aires.
- _____(1984), *Estadísticas vitales. Información básica*, Serie Estadísticas Vitales N° 23, Ministerio de Salud y Acción Social, Buenos Aires.
- Swanson, David A. y Lucky Tedrow (1984), "Improving the measurement of temporal change in regression models used for country population estimates", *Demography* N° 21.
- Teixeira Jardim, María de Lourdes (1992), "El uso de variables sintomáticas para estimar la distribución espacial de población: Aplicación a los municipios de Río Grande del Sur - Brasil", *Anais del VIII Encontro Nacional de Estudos Populacionais*, Asociación Brasileña de Estudios Poblacionales (ABEP).



NACIONES UNIDAS



Serie

Población y desarrollo

Números publicados

- 1 Migración y desarrollo en América del Norte y Centroamérica: una visión sintética (LC/L.1231-P), N° de venta: S.99.II.G.22 (US\$ 10.00), 1999. [www](#)
- 2 América Latina y el Caribe: crecimiento económico sostenido, población y desarrollo (LC/L.1240-P), N° de venta: S.99.II.G.30 (US\$ 10.00), 1999. [www](#)
- 3 Migración internacional de jóvenes latinoamericanos y caribeños: protagonismo y vulnerabilidad (LC/L.1407-P), N° de venta: S.00.II.G.75 (US\$ 10.00), 2000. [www](#)
- 4 El envejecimiento de la población latinoamericana: ¿hacia una relación de dependencia favorable? (LC/L.1411-P), N° de venta: S.00.II.G.80 (US\$ 10.00), 2000. [www](#)
- 5 Vulnerabilidad demográfica: una faceta de las desventajas sociales (LC/L.1422-P), N° de venta: S.00.II.G.97 (US\$ 10.00), 2000. [www](#)
- 6 Juventud, población y desarrollo: problemas, posibilidades y desafíos (LC/L.1424-P), N° de venta: S.00.II.G.98 (US\$ 10.00), 2000. [www](#)
- 7 Población y desarrollo en América Latina y el Caribe: un desafío para las políticas públicas (LC/L.1444-P), N° de venta: S.00.II.G.118 (US\$ 10.00), 2000. [www](#)
- 8 Los problemas en la declaración de la edad de la población adulta mayor en los censos (LC/L. 1442-P), N° de venta: S.00.II.G.117 (US\$ 10.00), 2000. [www](#)
- 9 Adolescencia y juventud en América Latina y el Caribe. Problemas, oportunidades y desafíos (LC/L.1445-P), N° de venta: S.00.II.G.122 (US\$ 10.00), 2000. [www](#)
- 10 La migración internacional y el desarrollo en la era de la globalización e integración: temas para una agenda regional (LC/L.1459-P), N° de venta: S.00.II.G.140 (US\$ 10.00), 2000. [www](#)
- 11 Insumos sociodemográficos en la gestión de políticas sectoriales (LC/L.1460-P), N° de venta: S.00.II.G.141 (US\$ 10.00), 2000. [www](#)
- 12 Informe de relatoría del Simposio sobre migración internacional en las Américas (LC/L.1462-P), N° de venta: S.00.II.G.144 (US\$ 10.00), 2000. [www](#)
- 13 Estimación de población en áreas menores mediante variables sintomáticas: una aplicación en departamentos de la República Argentina (1991 y 1996) (LC/L.1481-P), N° de venta: S.01.II.G.14 (US\$ 10.00), 2001. [www](#)

- El lector interesado en adquirir números anteriores de esta serie puede solicitarlos dirigiendo su correspondencia a la Unidad de Distribución, CEPAL, Casilla 179-D, Santiago, Chile, Fax (562) 210 2069, publications@eclac.cl.
- **www**: Disponible también en Internet: <http://www.eclac.cl>

Nombre:.....
Actividad:
Dirección:
Código postal, ciudad, país:.....
Tel.: Fax: E.mail: