

Estimación de la población de áreas subprovinciales con variables sintomáticas. Córdoba (Argentina).

González, Leandro - Celton, Dora - Torres, Eduardo.

Cita:

González, Leandro - Celton, Dora - Torres, Eduardo (Noviembre, 2010). *Estimación de la población de áreas subprovinciales con variables sintomáticas. Córdoba (Argentina). IV Congreso de la Asociación Latinoamericana de Población. Asociación Latinoamericana de Población, La Habana.*

Dirección estable: <https://www.aacademica.org/leandro.m.gonzalez/9>

ARK: <https://n2t.net/ark:/13683/ptoh/g3b>



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>.

Estimación de la población de áreas subprovinciales con variables sintomáticas. Córdoba (Argentina)¹

Leandro Mariano González²
Dora Estela Celton³
Eduardo Torres⁴

Resumen

La ponencia se propone efectuar un análisis del grado de aproximación de las estimaciones de población por departamentos en la provincia de Córdoba (Argentina), efectuadas con variables sintomáticas, a partir de los resultados del censo provincial 2008. Se dispone de datos de nacimientos, defunciones, alumnos primarios y electores. Se efectúan estimaciones de la población de los 26 departamentos con diferentes métodos, y se la coteja con la población censada en 2008. De la comparación de los resultados se obtienen diferencias absolutas y relativas de la población estimada con variables sintomáticas, que permite detectar qué variables son más precisas para la estimación de la población departamental, lo que posibilita efectuar estimaciones confiables para años intercensales. La información de electores y alumnos primarios ofrecen las mejores estimaciones de población, seguidas por los hechos vitales. Se obtiene numerosos métodos de cálculo entre los que se destacan las regresiones lineales, el método compuesto y algunos procedimientos con proporciones. En general las poblaciones intermedias y mayores son mejor estimadas que las menos pobladas, aunque los procedimientos pueden resultar eficaces en diferentes jurisdicciones. La falta de información sintomática para algunos años puede limitar su aplicación o el análisis de su precisión respecto al censo.

¹ Trabajo presentado en el IV Congreso de la Asociación Latinoamericana de Población, realizado en La Habana, Cuba, del 16 al 19 de noviembre de 2010.

² CONICET - Universidad Nacional de Córdoba (Argentina), leandrogonzalez@yahoo.com.ar

³ CONICET - Universidad Nacional de Córdoba (Argentina), dora.celton6@gmail.com

⁴ CONICET, torresedu@gmail.com.

Estimación de la población de áreas subprovinciales con variables sintomáticas. Córdoba (Argentina)

1- Introducción

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

Cuando no se dispone de información censal actualizada de áreas menores se puede recurrir a una variedad de métodos de estimación demográfica basados en funciones matemáticas. La población actual o futura de un área se basa en los resultados de los últimos censos de población y en alguna función matemática (generalmente exponencial o logística). Arriaga advierte que estos métodos no producen *proyecciones* propiamente dichas, sino *estimaciones* de población. La principal diferencia radica en que las estimaciones no proporcionan por lo general la estructura de edad sino el total de una población, aunque la composición etaria puede ser estimada por otros procedimientos (Arriaga, 2001:412-13).

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

Una cantidad no despreciable de registros estadísticos guarda una relación con el cambio demográfico aunque este vínculo no es preciso. Por lo tanto 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). Autores como Guiomar Bay (1998), Teixeira Jardim (2001) y Chaves Esquivel (2001) han aplicado diferentes métodos para estimar la población de áreas subnacionales con variables sintomáticas, aplicando diferentes métodos: distribución por prorrateo, distribución proporcional, tasas vitales, razón censal, diferencia de tasa, método compuesto, correlación de razón, correlación de tasa y correlación de diferencia.

La información referida a nacimientos, defunciones, alumnos matriculados, usuarios de servicios públicos, viviendas construidas, vehículos inscriptos, electores habilitados para votar, entre otros registros, permiten ser asociados con la población de las jurisdicciones subprovinciales. A partir de diversos métodos estadísticos, y sobre la base de una proyección demográfica de la población mayor (provincial o nacional), se procede a la estimación de la población de un área determinada para la cual se dispone de variables sintomáticas, y los censos de población permiten evaluar el grado de exactitud de las estimaciones realizadas.

Este trabajo se propone efectuar un análisis del grado de aproximación de las estimaciones de población por departamentos en la provincia de Córdoba (Argentina), efectuadas con variables sintomáticas, a los resultados del censo provincial 2008.

La provincia de Córdoba se ubica en el centro de la República Argentina y se encuentra dividida en 26 departamentos. En 2008 el gobierno provincial realizó un censo de población, que arrojó una población total de 3.243.621 habitantes, y una tasa de crecimiento anual de

8,3% en el período 2001-08. El departamento Capital, donde se ubica la ciudad de Córdoba, concentra cerca del 41 por ciento de la población provincial. Un 14 por ciento reside en tres departamentos limítrofes a la capital y contienen parte del conurbano de Córdoba (Colón, Santa María y Punilla), que se caracteriza por un marcado crecimiento poblacional impulsado por la migración intrametropolitana. Otro 18 por ciento habita en tres departamentos de la región pampeana con importantes ciudades intermedias (Río Cuarto, San Justo y General San Martín). El resto se distribuye en áreas rurales combinadas con ciudades menores (20 por ciento), y una región noroeste poco poblada y netamente rural (7 por ciento). En el noroeste provincial se encuentran los tres departamentos menos poblados (Pocho, Minas y Sobremonte), por debajo de los 6.000 habitantes cada uno.

2- Metodología

Para este trabajo se cuenta con los antecedentes de una investigación anterior (González, 2007), donde se recopilaron los registros estadísticos susceptibles de ser empleados como variables sintomáticas de los departamentos de la provincia de Córdoba, en el período 1980-2001. En el mismo se reunieron datos de nacimientos, defunciones, alumnos primarios, electores y viviendas construidas. Se aplicaron ocho métodos diferentes de cálculo (distribución por prorrateo y proporcional, razón censal, diferencia de tasas, método compuesto, correlaciones de razón, de tasas y de diferencia) y se analizó su aproximación a los resultados del Censo 2001. En esta oportunidad se repite el análisis con los resultados del último censo provincial. Para ello se recolectó la información disponible por departamentos más recientes y se efectuaron las estimaciones, a fin de cotejarlas con los resultados censales.

En primer lugar, se presentan sintéticamente los procedimientos empleados para calcular la población de poblaciones menores a partir de variables sintomáticas (BAY, 1998:185-89; TEIXEIRA JARDIM, 2001:35-37; CHAVES ESQUIVEL, 2001:54-58):

- Distribución por prorrateo: Supone que la población entre áreas menores se distribuye en la misma proporción registrada para las variables sintomáticas. Es de fácil cálculo y requiere información sobre un momento temporal, aunque es muy sensible a la calidad de los registros sintomáticos.

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

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

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

- Compuesto: Emplea distintos registros sintomáticos para estimar proporcionalmente la población de un área local por grupos de edades. En este caso se emplea la información de nacimientos para estimar la población menor de 5 años; la matrícula escolar primaria para

calcular la población de 5 a 14 años; los electores para estimar el grupo 15-69 años; y las defunciones para el cálculo de la población de 70 años y más. La principal ventaja es la utilización de cada variable sintomática para estimar el grupo etario con el que está más relacionado; la desventaja es la necesidad de contar con un conjunto de registros sintomáticos de manera simultánea.

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

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

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

Se utilizaron las siguientes fuentes de información: los resultados del censo 2008 difundidos por la Dirección de Estadísticas de la Provincia de Córdoba; registros de hechos vitales (nacimientos y defunciones) del Ministerio de Salud de la Nación; electores empadronados en comicios provinciales, ofrecidos por la Justicia Electoral Provincial y el Ministerio del Interior de la Nación; alumnos matriculados en el nivel primario registrados por el Ministerio de Educación de la Provincia de Córdoba. Se descartó la información referida a viviendas construidas, debido a los elevados errores que registraron en las estimaciones realizadas para el año 2001.

Para el cálculo de la población de cada departamento con los diferentes métodos de variables sintomáticas, se requiere contar con la proyección de la población del área mayor que las contiene. En este trabajo se efectuó una proyección de la población de la provincia de Córdoba durante el período 2001-2008, a través del método de los componentes y empleando el programa RUPLEX (desarrollado por el Bureau of the Census de EE.UU.; Arriaga 2001). Se partió de la población censada en 2001, ajustada por omisión censal por sexo y grupos etarios, y se emplearon los registros de nacimientos (por edad de la madre) y defunciones (por sexo y edad) en el mismo período (Ministerio de Salud de la Nación).

En segundo lugar, para evaluar el grado de aproximación de las estimaciones con variables sintomáticas se cotejó con la población censada en el último relevamiento disponible. Dado que el censo en la provincia de Córdoba fue realizado el 28 de agosto de 2008, se interpoló exponencialmente la población total al 1 de julio del mismo año dado que las variables sintomáticas estiman la población de cada área a mitad de año calendario. Esto se debe a que están basadas en información relevada a lo largo de un año calendario, y a su vez emplean la población provincial proyectada a mitad de año. Para ello se calculó la diferencia porcentual absoluta o error de estimación mediante la siguiente fórmula:

$$DA = \left| \frac{PE_i - PC_i}{PC_i} \right| \times 100$$

Donde:

DA = diferencia porcentual absoluta o error de estimación.

PE_i = población estimada con variables sintomáticas en el área i.

PC_i = población censada en el área i a mitad de año calendario.

Se asume que una estimación con variables sintomáticas es aceptable cuando la diferencia porcentual, sin considerar el signo algebraico, es inferior a 10 puntos porcentuales. Este límite es arbitrario y puede resultar excesivo en términos absolutos cuando se trata de poblaciones de mayor magnitud. Por lo tanto se puede reconsiderar el nivel de error de acuerdo a la escala demográfica con la que se trabaje o se tenga mayor interés.

Con relación a la disponibilidad de las variables sintomáticas, cabe aclarar que no siempre se cuenta con los registros estadísticos necesarios para todos los años calendarios. Esto ocurre especialmente con los datos electorales, ya en Argentina las elecciones se realizan regularmente en años impares. Pero como esta información ha demostrado ser muy eficaz para la estimación de poblaciones departamentales para el año 2001, se emplearon los últimos datos electorales disponibles y se cotejó con la población interpolada a mitad del año 2007.

Para el año 2007 también se realizaron estimaciones con la información de la matrícula escolar primaria, debido a que es la última información disponible. Esta variable también ofreció resultados muy satisfactorios en las pruebas realizadas para el año 2001.

Con los restantes registros sintomáticos disponibles (nacimientos y defunciones) se efectuaron estimaciones para el año censal 2008. A continuación se presentan los resultados obtenidos de acuerdo a la disponibilidad de información sintomática.

3- Resultados obtenidos

En primer lugar se presentan los resultados para 2008, el año en que se realizó el último censo de población. Se utilizaron los registros de hechos vitales (nacimientos y defunciones) por departamentos difundidos por el Ministerio de Salud de la Nación, que son los únicos datos disponibles para dicho año al momento de la redacción de este informe. En el Cuadro 1 y Gráficos 1.A-1.B se presentan aquellos métodos que produjeron diferencias inferiores a 10 puntos porcentuales en promedio, respecto a la población censada.

Con los registros de nacimientos y defunciones por departamentos se obtuvieron cinco estimaciones con aceptables grados de error respecto al censo. El método de diferencia de tasas presenta una diferencia media de 7,1 puntos porcentuales cuando se emplean las defunciones como variable sintomática, seguido del método de razón censal (7,6 puntos) también con defunciones. Ambos métodos son muy semejantes y sólo varían en el cálculo del cambio en los valores de las variables a lo largo del tiempo (diferencia y razones respectivamente). En ambos casos las estimaciones obtenidas son más precisas en los departamentos con población mayor a los 35.000 habitantes, con excepción de Punilla y Santa María (jurisdicciones vecinas a la capital provincial) (Gráfico 1.B).

Las defunciones también ofrecen estimaciones aceptables con el método de distribución proporcional, con un error medio de 9,2 puntos porcentuales; las áreas mejor estimadas son las que superan los 55.000 habitantes con excepción de los departamentos cercanos a la capital (Colón, Punilla y Santa María).

Cuadro 1: Población total y errores de estimación de las variables sintomáticas según diferentes métodos. Departamentos de la provincia de Córdoba, 2008. Valores absolutos.

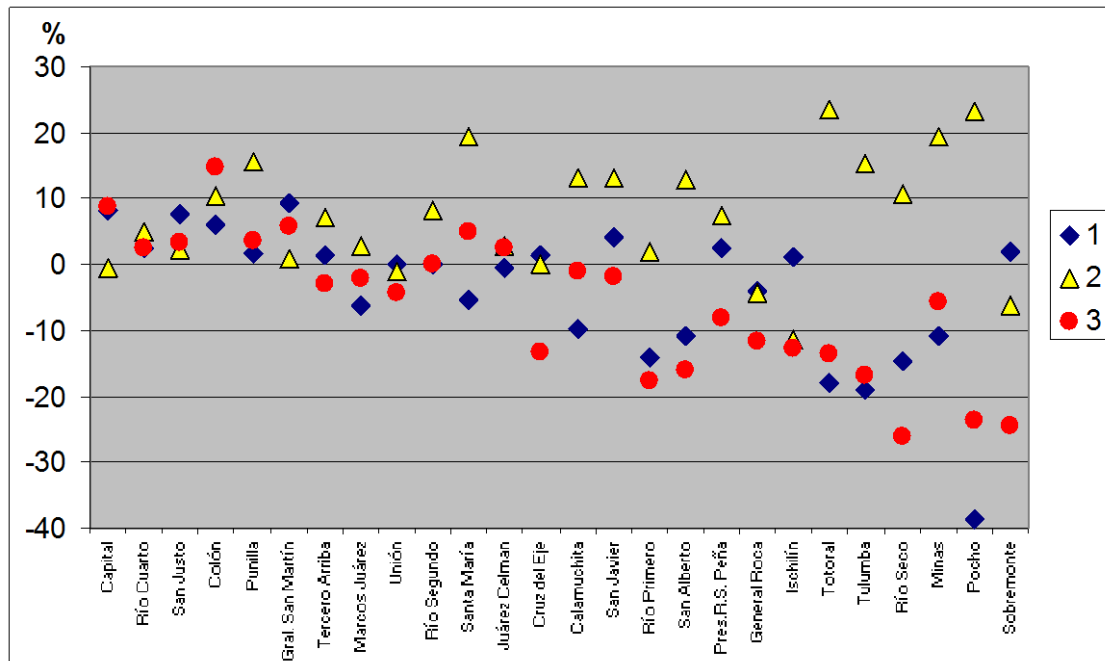
Departamento	Población al 1/7/2008 (1)	Error de estimación según método de cálculo y tipo de variable (%)				
		Distribución por prorrateo - Hechos vitales	Distribución proporcional - Defunciones	Distribución proporcional - Hechos vitales	Razón censal - Defunciones	Diferencia de tasas - Defunciones
Error medio (2)		7,7	9,2	9,5	7,6	7,1
Capital	1.309.078	8,3	-0,4	8,8	-3,4	-3,5
Río Cuarto	242.868	2,6	4,8	2,5	1,6	2,1
San Justo	200.506	7,7	2,3	3,3	-0,8	-0,1
Colón	200.063	5,9	10,5	14,8	7,1	5,7
Punilla	162.281	1,6	15,5	3,5	12,0	11,9
Gral. San Martín	122.614	9,3	0,9	5,8	-2,2	-1,6
Tercero Arriba	109.109	1,5	7,1	-2,9	3,8	4,3
Marcos Juárez	104.093	-6,3	2,7	-2,0	-0,5	0,2
Unión	103.131	0,1	-1,0	-4,2	-4,0	-3,4
Río Segundo	99.248	0,0	8,3	0,0	5,0	4,7
Santa María	92.933	-5,4	19,4	4,9	15,8	14,2
Juárez Celman	58.342	-0,5	2,8	2,5	-0,4	-0,5
Cruz del Eje	55.369	1,5	0,0	-13,2	-3,1	-3,2
Calamuchita	51.625	-9,8	13,0	-1,2	9,6	8,1
San Javier	51.121	4,1	13,2	-1,8	9,8	9,6
Río Primero	45.347	-14,1	2,1	-17,7	-1,1	-1,9
San Alberto	36.154	-10,9	13,0	-16,1	9,5	7,3
Pres.R.S. Peña	36.144	2,5	7,4	-8,2	4,1	4,7
General Roca	35.390	-4,0	-4,3	-11,8	-7,2	-7,2
Ischilín	31.375	1,2	-11,4	-12,8	-14,1	-13,8
Totoral	17.699	-17,8	23,4	-13,5	19,6	17,2
Tulumba	13.356	-19,1	15,3	-16,8	11,8	11,5
Río Seco	13.087	-14,6	10,8	-26,1	7,4	6,6
Minas	5.245	-10,8	19,4	-5,6	15,7	15,1
Pocho	5.144	-38,7	23,2	-23,7	19,4	17,8
Sobremonte	4.919	2,1	-6,2	-24,4	-9,0	-7,3

(1) Población interpolada exponencialmente al 1 de julio de 2008, entre los censos 2001 y 2008.

(2) Promedio de los valores absolutos de los errores de estimación (%).

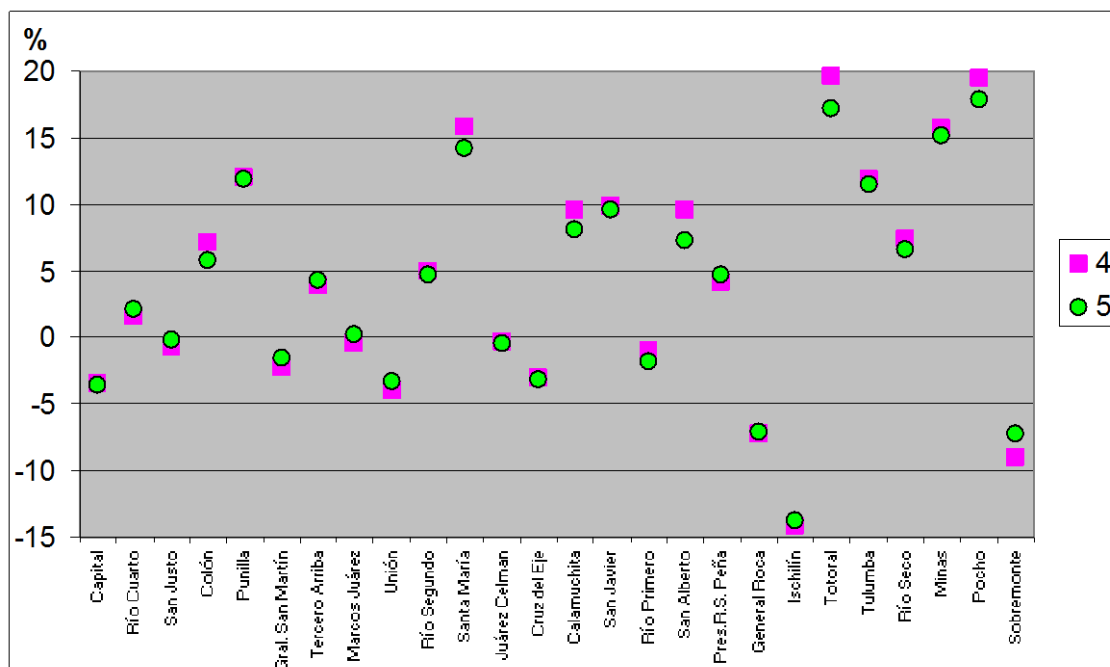
Fuente: Elaboración propia a partir de los censos de población 2001 y 2008 y registros de estadísticas vitales de la provincia de Córdoba.

Gráfico 1A: Errores de estimación de la población con variables sintomáticas según diferentes métodos. Departamentos de la provincia de Córdoba, 2008. Porcentajes.



- Referencias:** 1 – Distribución por prorrato – Hechos vitales.
 2 – Distribución proporcional- Defunciones.
 3 - Distribución proporcional – Hechos vitales.

Gráfico 1B: Errores de estimación de la población con variables sintomáticas según diferentes métodos. Departamentos de la provincia de Córdoba, 2007. Porcentajes.



- Referencias:** 4 – Razón censal – Defunciones.
 5 – Diferencia de tasas- Defunciones.

Fuente: Elaboración propia a partir de los censos de población 2001 y 2008 y registros de estadísticas vitales de la provincia de Córdoba.

Cuando se emplean los hechos vitales en su conjunto (nacimientos más defunciones) se obtienen estimaciones razonables con los métodos de distribución por prorratio (7,7 puntos en promedio) y distribución proporcional (9,5 puntos). El primer procedimiento es más preciso en las poblaciones mayores a 50.000 habitantes; se observa también que tiende a subestimar la población a medida que se reduce la escala demográfica de las jurisdicciones, debido a una mayor cantidad de errores de signo negativo que se obtienen frente a la población censada (Gráfico 1.A). Con el método de distribución proporcional se registran mejores estimaciones en las áreas mayores a 55.000 habitantes (excepto Colón), y también se verifica la tendencia a subestimar las poblaciones más reducidas.

En segundo lugar, se reseñan las estimaciones que se obtienen con la información disponible para el año 2007. Para este año se cuenta con registros de alumnos primarios y electores (además de registros vitales), los que han demostrado buena capacidad predictiva de la población de los departamentos de la provincia de Córdoba en comparación con el censo 2001. Por esta razón se ensaya el grado de aproximación de estas variables con la población interpolada a mitad del año 2007, de acuerdo a las tasas de crecimiento intercensales de cada departamento (2001 y 2008). Si bien las poblaciones interpoladas se alejan temporalmente del último censo, se considera oportuno el ejercicio a fin de evaluar la capacidad predictiva de estos registros y ser utilizados en años posteriores. En el Cuadro 2 y los Gráficos 2.A-B-C se presentan los resultados más precisos, en términos de errores medios en todos los departamentos.

Como se puede apreciar, la información de alumnos primarios y electores ofrecen múltiples modelos de estimación de la población con grados de error muy satisfactorios en comparación con la población interpolada al año 2007. Entre estos procedimientos se encuentran el método compuesto y los procedimientos basados en regresiones lineales (correlación de razón, de tasa o de diferencia). Si bien el método compuesto posee mayor sustento conceptual, ya que relaciona proporcionalmente las variables sintomáticas con los grupos etarios con los que está naturalmente relacionado (por ejemplo, los nacimientos con la población menor de cinco años), no necesariamente arroja mejores estimaciones que los procedimientos con regresiones.

Las estimaciones calculadas con el método compuesto arrojan un error medio de 3 puntos porcentuales (Cuadro 2). Los 26 departamentos de la provincia de Córdoba registran diferencias frente a la población interpolada menores al 10 por ciento; sólo cuatro de ellos superan el 5 por ciento: Capital, Punilla y Santa María (pertenecientes al área metropolitana), y Sobremonte, el departamento menos poblado. Este procedimiento tiende a subestimar levemente la población en las jurisdicciones menores a 50.000 habitantes (Gráfico 2.A).

En relación con las regresiones lineales, la eficacia de las estimaciones depende más de las variables utilizadas más que del procedimiento de cálculo. Así es como los registros electorales logran menores errores en promedio que los datos de matrícula escolar primaria. Cuando se utilizan los electores como única variable sintomática, el método de correlación de tasa muestra un error medio de 1,8 puntos porcentuales y el de correlación de diferencia 2,4 puntos. En ambos procedimientos la totalidad de departamentos registran errores inferiores al 5 por ciento, con la única excepción de Sobremonte (sólo en la correlación de diferencia). En los Gráficos 2.B y 2.C se observa que las correlaciones de diferencias tienden a subestimar levemente las poblaciones en comparación con las correlaciones de tasa, aunque los resultados en ambos casos son muy precisos.

Cuadro 2: Población total y errores de estimación de las variables sintomáticas según diferentes métodos. Departamentos de la provincia de Córdoba, 2007. Porcentajes.

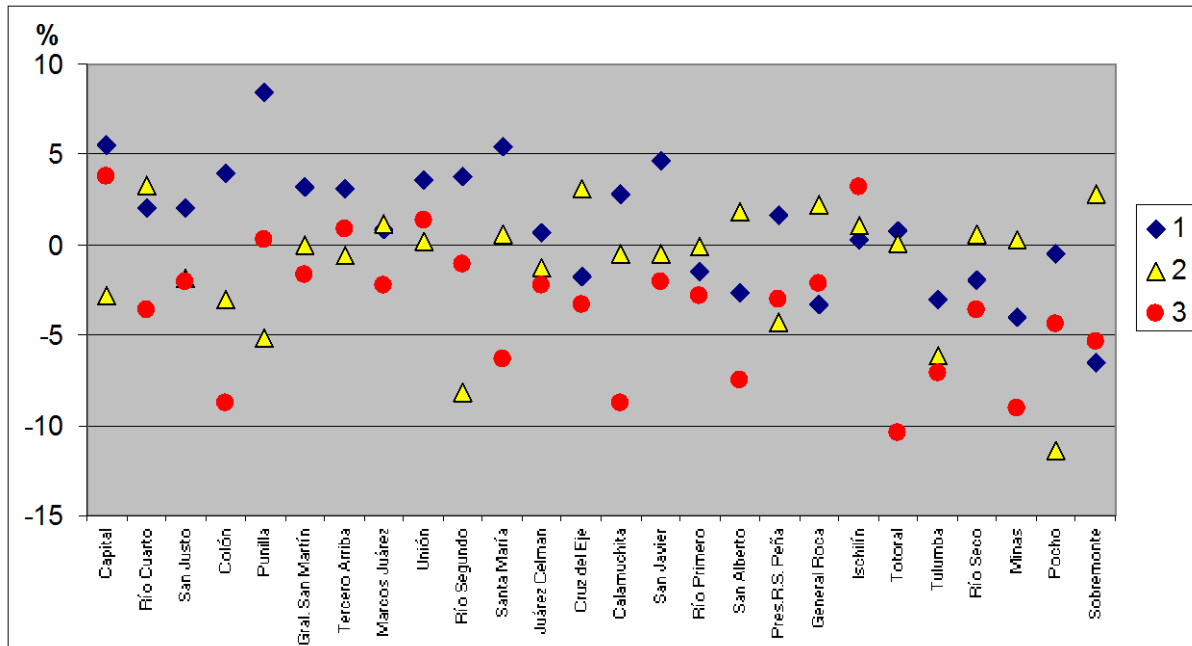
Departamento	Población al 1/1/2007 (1)	Error de estimación según método de cálculo y tipo de variable (%)								
		Método compuesto	Correlación de razón		Correlación de tasa			Correlación de diferencia		
			Nacimientos y alumnos	Defunciones y electores	Alumnos	Electores	Defunciones y electores	Electores	Defunciones y alumnos	Defunciones y electores
Error medio (2)		3,0	2,4	4,1	2,8	1,8	2,1	2,4	5,4	3,3
Capital	1.309.078	5,6	-2,8	3,7	-4,0	3,4	3,8	2,0	-5,2	3,2
Río Cuarto	242.868	2,0	3,3	-3,6	0,5	-1,1	-1,3	-2,9	-3,6	-3,5
San Justo	200.506	2,1	-1,9	-2,1	-8,9	-0,4	-0,5	-2,1	7,7	-2,1
Colón	200.063	4,0	-3,1	-8,8	-3,3	-4,1	-5,0	-2,3	-3,9	-5,5
Punilla	162.281	8,5	-5,2	0,3	-4,8	3,9	3,3	4,1	-9,9	1,8
Gral. San Martín	122.614	3,2	-0,1	-1,6	-2,9	-0,4	-0,4	-1,9	-3,1	-1,7
Tercero Arriba	109.109	3,1	-0,6	0,9	-1,2	2,8	2,7	0,6	-11,9	0,7
Marcos Juárez	104.093	0,9	1,2	-2,2	-1,5	-0,5	-0,5	-2,8	0,3	-2,6
Unión	103.131	3,6	0,2	1,4	-1,2	1,5	1,8	-0,5	-1,3	0,9
Río Segundo	99.248	3,7	-8,2	-1,1	-7,2	1,5	1,3	0,0	1,9	-0,7
Santa María	92.933	5,4	0,6	-6,4	3,2	-0,3	-1,3	-0,7	11,7	-3,6
Juárez Celman	58.342	0,7	-1,3	-2,2	-1,5	-0,5	-0,5	-2,1	1,8	-2,2
Cruz del Eje	55.369	-1,7	3,1	-3,3	0,0	-1,8	-1,8	-3,7	11,3	-3,4
Calamuchita	51.625	2,8	-0,5	-8,7	-2,6	-3,4	-4,3	-2,9	0,4	-5,8
San Javier	51.121	4,7	-0,5	-2,0	-1,6	1,8	1,2	0,8	2,7	-1,0
Río Primero	45.347	-1,5	-0,1	-2,9	0,7	-1,1	-1,2	-2,3	7,6	-2,6
San Alberto	36.154	-2,7	1,9	-7,5	-0,1	-2,5	-3,3	-2,4	4,0	-4,7
Pres.R.S. Peña	36.144	1,6	-4,3	-3,0	-6,0	-0,1	-0,4	-2,3	6,1	-3,0
General Roca	35.390	-3,3	2,2	-2,1	0,0	-1,8	-1,6	-3,8	9,3	-2,7
Ischilín	31.375	0,2	1,1	3,2	-1,7	0,8	1,7	-0,8	-0,4	2,2
Totalal	17.699	0,8	0,1	-10,5	-3,3	-2,5	-3,7	-4,1	10,0	-7,0
Tulumba	13.356	-3,0	-6,2	-7,2	-5,0	-1,7	-2,5	-1,9	-9,9	-5,0
Río Seco	13.087	-2,0	0,6	-3,6	0,4	0,1	-0,3	-2,3	0,3	-3,3
Minas	5.245	-4,0	0,3	-9,0	-1,4	-2,2	-3,3	-4,0	13,8	-7,1
Pocho	5.144	-0,5	-11,4	-4,4	-9,1	2,0	1,2	-1,9	-0,5	-3,9
Sobremonte	4.919	-6,5	2,8	-5,4	1,1	-4,7	-4,5	-7,2	-2,6	-5,7

(1) Población interpolada exponencialmente al 1 de julio de 2007, entre los censos 2001 y 2008.

(2) Promedio de los valores absolutos de los errores de estimación (%).

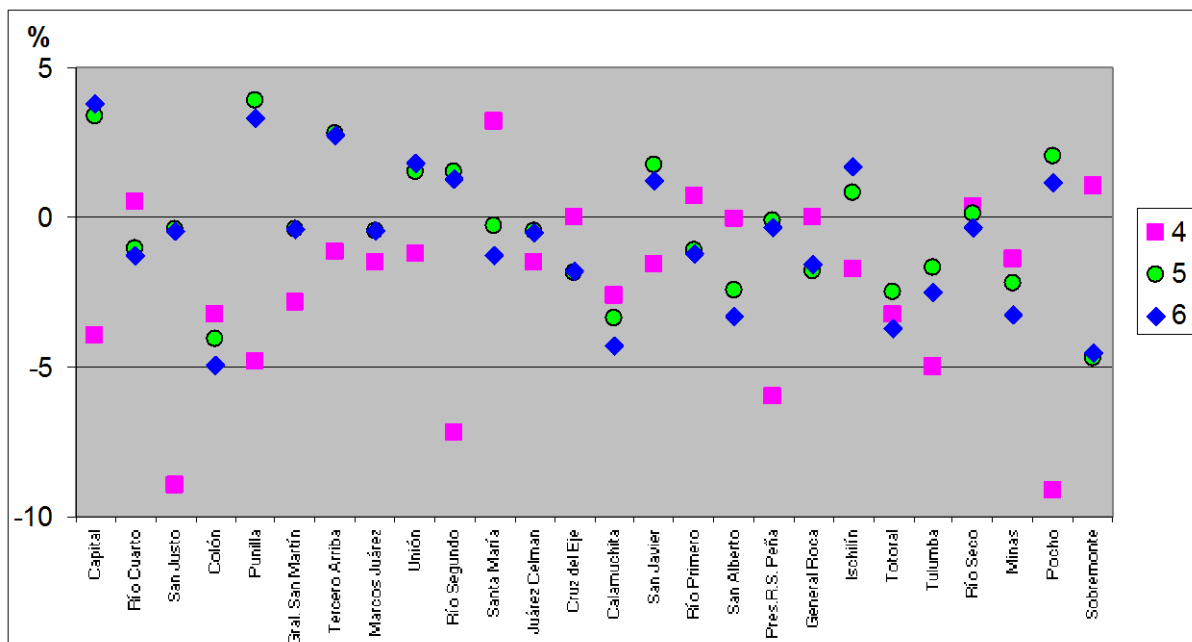
Fuente: Elaboración propia a partir de los censos de población 2001 y 2008, registros de estadísticas vitales y electores de la provincia de Córdoba.

Gráfico 2A: Errores de estimación de la población con variables sintomáticas según el método compuesto y correlaciones de razón. Departamentos de la provincia de Córdoba, 2007. Porcentajes.



- Referencias:** 1 - Método compuesto.
 2 - Correlación de razón – Nacimientos y alumnos.
 3 - Correlación de razón – Defunciones y electores.

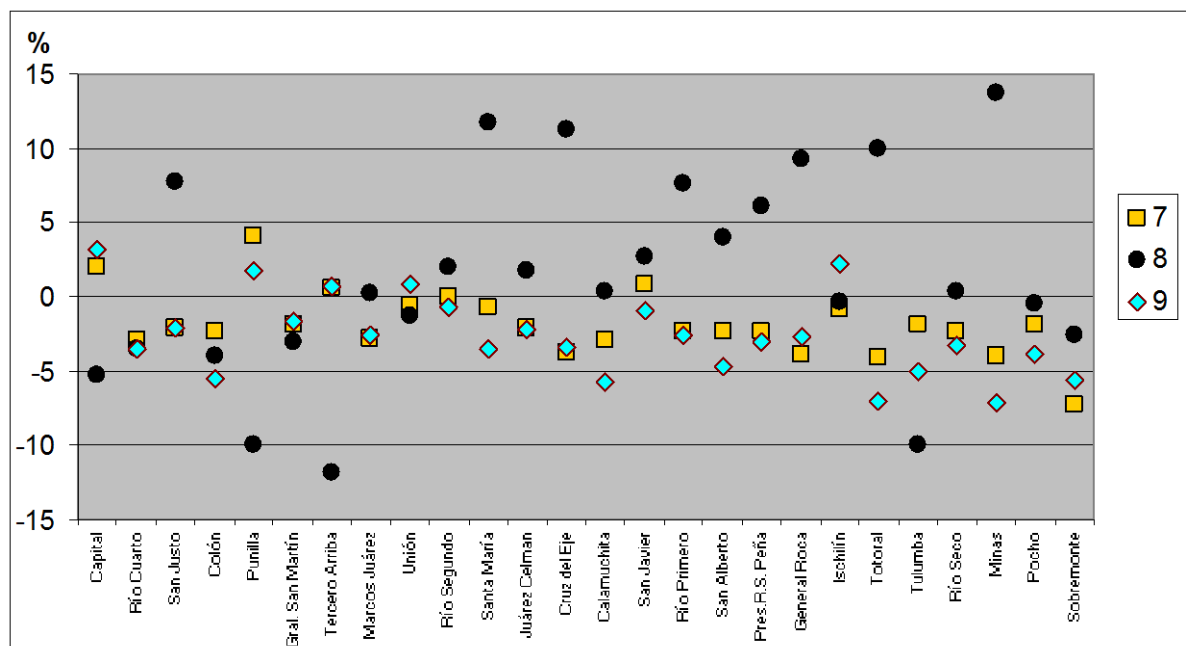
Gráfico 2B: Errores de estimación de la población con variables sintomáticas con el método de correlación de tasa. Departamentos de la provincia de Córdoba, 2007. Porcentajes.



- Referencias:** 4- Correlación de tasa - Alumnos.
 5- Correlación de tasa – Electores.
 6- Correlación de tasa - Defunciones y electores.

Fuente: Elaboración propia a partir de los censos de población 2001 y 2008, registros de estadísticas vitales y electores de la provincia de Córdoba.

Gráfico 2C: Errores de estimación de la población con variables sintomáticas según con el método de correlación de diferencia. Departamentos de la provincia de Córdoba, 2007. Porcentajes.



- Referencias:**
- 7- Correlación de diferencia - Electores.
 - 8- Correlación de diferencia - Defunciones y alumnos.
 - 9- Correlación de diferencia - Defunciones y electores.

Fuente: Elaboración propia a partir de los censos de población 2001 y 2008, registros de estadísticas vitales y electores de la provincia de Córdoba.

Quando los datos electores son complementados con los registros de defunciones se pueden construir modelos lineales con los tres tipos de procedimientos, aunque la eficacia de las estimaciones es menor a la obtenida cuando se usan solamente electores (Cuadro 2). El procedimiento más preciso con defunciones y electores es el de correlación de tasa, con un error medio de 2,1 puntos porcentuales, seguido de las correlaciones de diferencia (3,3 puntos) y correlación de razón (4,1). El primero de ellos produce estimaciones muy ajustadas, con errores inferiores a 5 por ciento en todos los departamentos, y cierta dispersión en el área metropolitana (Colón, Capital y Punilla). Por su parte, las correlaciones de diferencia y de razón tienden a subestimar levemente las poblaciones, aunque las estimaciones son bastante aceptables en general (Gráficos 2.A-B-C).

Con la información de alumnos primarios también es posible realizar estimaciones a través de los tres métodos de regresión lineal, con niveles de aproximación muy aceptables. El procedimiento más eficiente se obtiene con las correlaciones de razón, empleando datos de alumnos y nacimientos, con un error promedio de 2,4 puntos porcentuales (Cuadro 2). Sólo un departamento registra un error superior a los 10 puntos (Pocho), y tres se ubican entre los 5 y 10 puntos (Punilla, Río Segundo y Tulumba) (Gráfico 2A).

Quando se emplea solamente los registros de alumnos con las correlaciones de tasa, se registra un error medio de 2,8 puntos porcentuales, levemente superior al procedimiento anterior. Todos los departamentos obtienen diferencias respecto a la población interpolada

menores a 10 por ciento y sólo cuatro por encima del cinco por ciento. En general este método de cálculo tiende a subestimar las poblaciones, aunque los resultados son muy aceptables.

Finalmente, las estimaciones calculadas con correlación de diferencia a partir de los datos de alumnos y defunciones tienen un grado menor de aproximación, con un error medio de 5,4 puntos porcentuales. Cuatro departamentos muestran errores superiores al 10 por ciento, y por otra parte, 14 áreas se ubican por debajo del 5 por ciento, por lo que la eficacia del procedimiento es aceptable en general y se debe analizar cada caso para su adopción.

Tomando en cuenta los hechos vitales, se puede observar que la información de nacimientos del año 2007 sólo se combina con la matrícula escolar en el método de correlación de razón (2,4 puntos de error medio), ofreciendo un buen procedimiento de estimación de la población departamental. Las defunciones, por su parte, se combinan con los electores en los tres tipos de regresiones lineales pero disminuyen la eficacia de los registros electorales cuando se aplican de manera individual; junto con la información de alumnos ofrecen estimaciones aceptables en general.

4- Comentarios finales

Primeramente se puede afirmar que los procedimientos de estimación con variables sintomáticas ofrecen varias alternativas de cálculo de la población de los departamentos de la provincia de Córdoba, con aceptables grado de aproximación con los últimos resultados censales. Los registros anuales de hechos vitales resultan más fáciles de obtener, a pesar de que se refieren a eventos ocurridos un año antes. La matrícula escolar también es una información anual pero su difusión es más acotada y generalmente debe solicitarse al Ministerio de Educación provincial. Los datos electorales en cambio se generan regularmente en años impares, cuando se efectúan elecciones de autoridades políticas y en el año en curso.

Un elemento importante en la aplicación de variables sintomáticas es la disponibilidad de una adecuada proyección demográfica del área mayor que contiene a las poblaciones de interés. En este caso la proyección provincial construida con el método de los componentes resultó ser ajustada a la población censada (3,7% superior), por lo que las estimaciones basadas en esta proyección cuentan con una buena referencia de la dinámica demográfica provincial.

Respecto a la medición del grado de eficacia de las diferentes estimaciones en comparación con la población censada, debe tenerse en cuenta que el límite de 10 puntos porcentuales es arbitrario. En las poblaciones mayores ese porcentaje puede representar un elevado número absoluto de habitantes, y por esa razón las estimaciones podrían ser imprecisas. Por lo tanto el usuario debe definir el límite más aceptable de acuerdo a la magnitud demográfica de las áreas de interés.

Asimismo es necesario considerar que las estimaciones con variables sintomáticas se basan en una proyección provincial o nacional, cuya población base generalmente está ajustada por omisión censal (2,75 por ciento en el censo 2001 en Argentina). Si los resultados del último censo realizado no cuenta con el mismo ajuste, parte de la diferencia relativa calculada entre las estimaciones con variables sintomáticas y la población censal podría estar afectada por el grado de omisión contemplado en la proyección del área mayor.

Con relación a los resultados obtenidos, puede considerarse que los errores calculados para el año 2008 son más fiables ya que corresponden al año censal. A pesar de la distancia temporal de las estimaciones del año 2007, su comparación con la población interpolada guarda interés metodológico por la mayor precisión que ofrece la información disponible en dicho año.

Para el año 2008 se obtuvieron cinco procedimientos con aceptable grado de estimación. La información de defunciones resultó ser más precisa que los registros de nacimientos, especialmente con el método de diferencia de tasas (7,1 puntos porcentuales en promedio) y razón censal (7,6 puntos). La suma de hechos vitales ofreció resultados aceptables con el método de distribución por prorrato (7,7 puntos). En estos modelos se detectaron menos precisión en algunos departamentos cercanos a la capital (Punilla, Santa María y Colón) y en las poblaciones menores a 50.000 habitantes aproximadamente.

El ejercicio de estimación para el año 2007 con la información electores y alumnos primarios ofrecieron un mayor número de opciones de cálculo, a la vez que un importante grado de aproximación a la población interpolada entre los censos. Los registros electorales ofrecen las mejores estimaciones cuando se los trata individualmente con el método de correlación de tasa (1,8 puntos de diferencia media) y correlación de diferencia (2,4 puntos). Cuando se incorporan las defunciones a estos modelos disminuye la eficacia general de los procedimientos, aunque mejoran las estimaciones para algunas áreas en particular.

Con la matrícula escolar primaria también se obtuvieron modelos de estimación muy precisos al año 2007. Junto con los registros de nacimientos se aplicaron eficazmente en las correlaciones de razón (2,4 puntos de error medio), y de manera individual en las correlaciones de tasa (2,8 puntos). Los hechos vitales no mostraron para el año 2007 mayor eficacia que la información de electores y alumnos primarios, y sólo se integran complementariamente a algunos modelos.

El método compuesto, por su parte, ofrece también una muy buena estimación de la población con un error medio de 3 puntos porcentuales. Este procedimiento goza a su vez de la ventaja de emplear un conjunto de variables sintomáticas y de relacionarlas a los grupos etarios con los que está naturalmente relacionado. Esta característica le confiere una ventaja conceptual respecto a los restantes procedimientos empleados.

En los varios procedimientos aplicados para 2007 se detectó la menor precisión de los departamentos del área metropolitana, como también la tendencia a subestimar la población de las áreas menos pobladas.

Cabe aclarar que en este trabajo se analizó la aproximación conjunta de los 26 departamentos de la provincia de Córdoba, a través del promedio de errores, para obtener un panorama general de la eficacia de las variables sintomáticas y los diferentes métodos de cálculo. Pero es necesario aclarar que una jurisdicción puede ser mejor estimada con algún método en particular, que no resulte eficaz en otra. Por lo tanto, si el interés del usuario se concentra en una población determinada o grupo de áreas, necesitará detectar cuál procedimiento e información sintomática es más útil.

El ejercicio aquí realizado permite ampliar el conocimiento acerca de estos procedimientos como herramienta de estimación de poblaciones menores en la provincia de Córdoba. Hasta aquí se han empleado las variables sintomáticas como estimadoras de la población total de las áreas de interés. Una variante que puede ser ensayada en futuros ejercicios es la aplicación de estos modelos a la estimación de los saldos migratorios, de manera que los registros vitales

puedan dar cuenta del crecimiento natural y las variables sintomáticas estimen el componente migratorio de las poblaciones.

Por otra parte, se considera que el análisis de los resultados de acuerdo al perfil demográfico de los departamentos permite contar con orientaciones acerca de la utilidad de los registros sintomáticos para la actualización de la información demográfica disponible. Resta realizar un análisis en torno a la calidad de los registros sintomáticos utilizados, a fin de controlar los sesgos que pueden imprimir los errores de registros a las estimaciones de la población.

BIBLIOGRAFÍA

- 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, 1998:181-208.
- CELADE (1998): *Demografía II*. México, Programa Latinoamericano de Actividades en Población, Carlos Welti Editor.
- CHAVES ESQUIVEL, Edwin (2001): “Variables sintomáticas en las estimaciones poblacionales a nivel cantonal en Costa Rica”, en *Notas de Población* N° 71. Santiago, CELADE, 2001:51-72.
- GONZÁLEZ, Leandro (2007): *Ajuste de proyecciones de poblaciones menores con variables sintomáticas. Localidades del Gran Córdoba 2001-16*. Córdoba, Tesis de Doctorado en Demografía, Universidad Nacional de Córdoba, inédito.
- INDEC (2003): *Censo Nacional de Población, Hogares y Viviendas 2001*. Buenos Aires, INDEC.
- DIRECCIÓN DE ESTADÍSTICAS Y CENSOS (2009): *Censo Provincial de Población 2008, Resultados Provisionales*, Gobierno de la Provincia de Córdoba, www.cba.gov.ar
- HOWE, Andrew (2004): “Assessing the accuracy of Australia’s small area population estimates, 2001”. Canberra, Australian Population Association.
- TEIXEIRA JARDIM, Maria de L. (2001): “Uso de variables sintomáticas para estimar la distribución espacial de población. Aplicación a los municipios de Rio Grande do Sul, Brasil”, en *Notas de Población* N° 71. Santiago, CELADE, 2001:21-50.