

**UNIVERSIDAD DE COSTA RICA
SISTEMA DE ESTUDIOS DE POSGRADO**

**INDICADORES SINTOMÁTICOS EN LAS ESTIMACIONES
POBLACIONALES PARA ÁREAS MENORES: COSTA RICA**

**TESIS SOMETIDA A LA CONSIDERACIÓN DE LA COMISIÓN DEL PROGRAMA DE
ESTUDIOS DE POSGRADO EN ESTADÍSTICA PARA OPTAR AL GRADO DE
MAGISTER SCIENTIAE**

EDWIN A. CHAVES ESQUIVEL

**CIUDAD UNIVERSITARIA "RODRIGO FACIO"
COSTA RICA**

1997

DEDICATORIA

A Jesús, mi guía y protector de todo momento.

A mis abuelitos José y María Rosa, cuya enseñanza sobrepasa cualquier conocimiento universitario.

A mi esposa Magda y mis hijas Cindy y Karla, quienes, junto al Bebé que viene en camino, son la principal razón de mi vida.

AGRADECIMIENTOS

El más profundo agradecimiento al M.Sc. Johnny Madrigal Pana, por su desinteresada ayuda en la culminación del presente trabajo. Su compromiso, motivación y dedicación son un verdadero ejemplo que trataré de imitar.

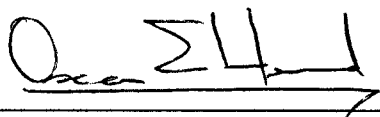
Del mismo modo, a la Dra. Doris Sosa por el apoyo brindado como directora del presente trabajo.

A la M.Sc. Irma Sandoval, al M.Sc. Juan Bautista Chavarría y al M.Sc. Miguel Gómez, por su apoyo y consejo como lectores del estudio.

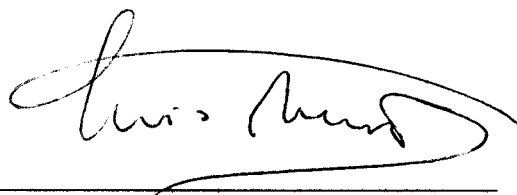
También agradezco profundamente al M.Sc. Oscar Hernández Rodríguez, Director de la Maestría y al Dr. Luis Rosero Bixby, quienes me apoyaron en todo momento brindándome su amistad y confianza.

Finalmente, quiero agradecer al M.Sc. Mario Castillo que además de compañero y amigo, con su ejemplo de sacrificio y voluntad me ha enseñado a no dejarme vencer por la adversidad.

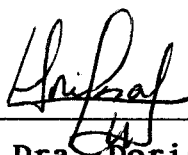
Esta tesis fue aceptada por la Comisión del Programa de Estudios de Posgrado en Estadística de la Universidad de Costa Rica, como requisito parcial para optar al grado de Magister Scientia



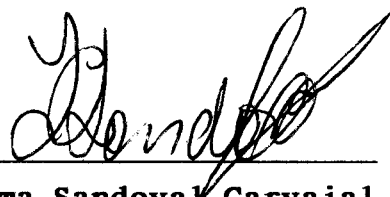
M. Phil. Oscar Hernández Rodríguez
Representante de la Decana
S.E.P



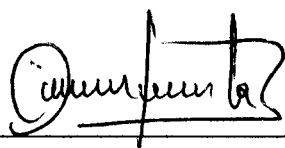
Dr. Luis Rosero Bixby
Representante del Director
Posgrado en Estadística



Dra. Doris Sosa Jara
Directora



M.Sc. Irma Sandoval Carvajal
Asesora



M.Sc. Juan B. Chavarría Chaves
Asesor



Edwin A. Chaves Esquivel
Candidato

ÍNDICE

1. INTRODUCCIÓN.....	1
2. ANTECEDENTES.....	6
2.1 Estimaciones y proyecciones a nivel nacional.....	7
2.2 Estimaciones y proyecciones en áreas menores.....	7
2.3 Limitaciones de las estimaciones poblacionales en áreas menores	9
2.4 Las variables sintomáticas y su aplicación.....	10
2.5 Indicadores más utilizados.....	12
2.6 Procedimientos de estimación.....	13
3. OBJETIVOS DE LA INVESTIGACIÓN.....	14
4. ESTRATEGIA METODOLÓGICA.....	15
4.1 Indicadores sintomáticos disponibles.....	18
4.2 Métodos de estimación.....	20
4.2.1 Método de Razón Censal y de Diferencia de Tasas.....	20
4.2.2 Método de Razón Correlación y de Tasa Correlación.....	23
4.3 Método de viviendas.....	28
4.4 Valoración de las estimaciones.....	32
4.4.1 Precisión de las estimaciones.....	33
4.5 Selección de los indicadores sintomáticos.....	34
4.6 Análisis de las técnicas de estimación.....	35
4.7 Proyecciones de población para áreas pequeñas.....	36
4.4 Análisis de grupos.....	37
4.7 Muestreo de áreas.....	38
5. PRECISIÓN Y APLICACIÓN SIMPLE DE LOS INDICADORES SINTOMÁTICOS.....	44
6. APLICACIÓN MULTIVARIABLE DE LOS MÉTODOS DE ESTIMACIÓN..	49
6.1 Estimaciones a nivel provincial.....	49
6.2 Estimaciones a nivel cantonal.....	52
6.2.1 Análisis de residuos.....	55
6.3 Estimaciones a nivel distrital	58
6.2.1 Análisis de residuos.....	61
6.4 Estimaciones poblacionales para 1996.....	63
6.5 Comparación de las estimaciones con las obtenidas mediante el método de Crecimiento Natural.....	69
6.6 Proyecciones poblacionales.....	70

6.7	Uso de las variables socio-demográficas para mejorar estimaciones.....	75
6.8	Aplicación del Método de Viviendas.....	79
7.	MUESTREO DE ÁREAS.....	82
8.	CONCLUSIONES Y RECOMENDACIONES.....	90
8.1	Conclusiones	91
8.1.2	Aplicación de los indicadores sintomáticos.....	91
8.1.2	Muestreo de áreas.....	95
8.2	Recomendaciones.....	96
9.	BIBLIOGRAFÍA.....	99
10.	ANEXOS.....	106
	ANEXO 1.....	107
	ANEXO 2.....	121
	ANEXO 3.....	130
	ANEXO 4.....	132
	ANEXO 5.....	136
	ANEXO 6.....	138

RESUMEN

Debido a la necesidad de contar con estimaciones y proyecciones poblacionales a nivel de áreas menores, y a la problemática que esto conlleva, ocasionada ante todo por los movimientos migratorios internos, esta investigación analizó la introducción de los indicadores sintomáticos para ayudar a predecir los cambios de la dinámica poblacional. A la vez, se examinó la factibilidad de incluir el muestreo de áreas como un alternativa adicional al problema de estimación.

En la aplicación de los indicadores sintomáticos se utilizaron cuatro métodos básicos: Razón Censal, Diferencia de Tasas, Razón Correlación y Tasa Correlación. Para valorar la calidad de los resultados, se efectuaron estimaciones y proyecciones poblacionales para 1984, comparándolas luego con su valor censal y con las estimaciones publicadas por la Dirección General de Estadística y Censos (DGEC), la que asume un crecimiento natural de la población.

Los resultados mostraron que, a nivel nacional, el número de indicadores sintomáticos es muy limitado; sin embargo, es posible obtener estimaciones de buena calidad. Se observó, además, que la precisión de las estimaciones depende del nivel de agregación del área político-administrativa de interés. A nivel provincial, con el uso del padrón electoral y la aplicación de los métodos de Razón Censal y Diferencia de Tasas, se logró gran eficiencia en las estimaciones. Su error promedio fue cercano al 1,5% con respecto a su valor censal, mostrando una mejora de aproximadamente 6 puntos porcentuales si se compara con las estimaciones de la DGEC. Por otro lado, a nivel cantonal, se obtuvo un error promedio cercano al 4%, donde más del 95% de los cantones presentaron errores inferiores al 10%, y la mejora con respecto a las estimaciones de la DGEC fue de 8,5 puntos porcentuales. A este nivel, las técnicas que emplean regresión lineal dieron los mejores resultados, sobre todo el método de Tasa Correlación con los indicadores padrón electoral y los nacimientos. A nivel distrital no se pudo

obtener toda la información sintomática utilizada para los otros dos tipos de áreas. Los resultados presentaron un error promedio de aproximadamente 12% y alrededor del 45% de las estimaciones indicaron errores superiores al 10%, por lo que la mejora respecto al error promedio en las estimaciones de la DGEC fue de aproximadamente 9 puntos porcentuales. A este nivel, el método de Tasa Correlación dio los mejores resultados.

Utilizando la información sintomática empleada en las anteriores estimaciones, se puso en práctica una estrategia que permite obtener proyecciones poblacionales para los cinco años posteriores a la obtención de los datos. La prueba se efectuó a nivel cantonal y los resultados indicaron que el 75% de los cantones presentan errores menores al 10% con un error promedio de 6,5%, lo que mejora el error de las estimaciones de la DGEC en 6 puntos porcentuales. Esta práctica permite actualizar las proyecciones cada nuevo año.

También se emplearon los abonados del servicio eléctrico para estimar el número de viviendas de las áreas menores, lo que mostró que este indicador ofrece una alternativa adicional en la solución del problema de estimaciones.

Para la aplicación del muestreo de áreas se utilizó la cartografía del Censo de Población y Vivienda de 1984. Los resultados evidenciaron lo valioso de la técnica, pues aparte de aproximar la población total, permite estimar otras características de interés, a la vez que puede ser aplicado a áreas que no necesariamente dependan de los límites político-administrativos. Su principal problema radica en la inversión de tiempo y costos, lo que obliga al interesado a evaluar si los logros alcanzados compensan el esfuerzo de llevarlo a cabo.

El estudio dejó en evidencia que, ante la carencia de información censal, los indicadores y el muestreo de áreas constituyen una alternativa que debe ser tomada en cuenta por los responsables de ofrecer al público datos poblacionales.

ÍNDICE DE CUADROS, GRÁFICOS Y FIGURAS

Cuadro 5.1 Matriz de correlación de las razones de cambio en las proporciones el tamaño poblacional y de ocho variables sintomáticas, para un conjunto de 22 cantones de San José y Heredia, período 73-84.....	45
Cuadro 5.2 Comparación del valor absoluto en el error porcentual de las estimaciones poblacionales por indicador sintomático según método utilizado, 1984.....	47
Cuadro 5.3 Valor de R^2 y significancia dada por la prueba "t" para cinco modelos de regresión simple, por indicador sintomático utilizado, según método de estimación, período 63-73.....	48
Cuadro 6.1 Comparación del valor absoluto en el error porcentual y del error porcentual en las estimaciones poblacionales a nivel provincial por indicador sintomático según método utilizado, 1984.....	50
Cuadro 6.2 Valor de R^2 y significancia dada por la prueba "t" para los tres modelos de regresión múltiple presentados en el Cuadro 6.1, por indicadores sintomáticos utilizados, según método de estimación.....	52
Cuadro 6.3 Comparación del valor absoluto en el error porcentual y del error porcentual en las estimaciones poblacionales a nivel cantonal por indicador sintomático según método utilizado, año 1984.....	53
Cuadro 6.4 Valor de R^2 y significancia dada por la prueba "t" para los cuatro modelos de regresión múltiple presentados en el Cuadro 6.3, por indicadores sintomáticos utilizados, según método de estimación (nivel cantonal).....	55
Cuadro 6.5 Resultados de la prueba Shapiro-Francia para determinar la normalidad de los residuos en los modelos de regresión utilizados para estimar la población de los cantones de Costa Rica en 1984.....	58

Cuadro 6.6	
Comparación del valor absoluto en el error porcentual y del error porcentual en las estimaciones poblacionales a nivel distrital por indicador sintomático según método utilizado, 1984.....	60
Cuadro 6.7	
Valor de R^2 y significancia dada por la prueba “t” para los tres modelos de regresión múltiple presentados en el Cuadro 6.5, por indicadores sintomáticos utilizados, según método de estimación.....	61
Cuadro 6.8	
Resultados de la prueba Shapiro-Francia para determinar la normalidad de los residuos en los modelos de regresión utilizados para estimar la población de los distritos de Costa Rica en 1984.....	63
Cuadro 6.9	
Estimaciones poblacionales para las provincias de Costa Rica según método utilizando como variable sintomática el padrón electoral, 1996.....	64
Cuadro 6.10	
Estimaciones poblacionales de los cantones de Costa Rica según método utilizando el padrón electoral y los nacimientos, 1996.....	65
Cuadro 6.11	
Valor de R cuadrado y significancia dada por la prueba “t” para los dos modelos de regresión múltiple utilizados para estimar la población a nivel provincial y cantonal de Costa Rica, 1996.....	68
Cuadro 6.12	
Resultados de la prueba Shapiro-Francia para determinar la normalidad de los residuos en los modelos de regresión utilizados para estimar la población de los cantones de Costa Rica en 1996.....	69
Cuadro 6.13	
Comparación del valor absoluto en el error porcentual de las estimaciones poblacionales de provincias, cantones y distritos, suponiendo un crecimiento natural, 1984.....	70
Cuadro 6.14	
Comparación del valor absoluto en el error porcentual en las proyecciones poblacionales a nivel cantonal, por indicador sintomático según hipótesis empleada, 1984.....	72

Cuadro 6.15	
Proyecciones poblacionales de los cantones de Costa Rica Período 1997-2001.....	73
Cuadro 6.16	
Comparación del valor absoluto en el error porcentual de las estimaciones poblacionales a nivel cantonal, por grupo seleccionado, según método utilizado, 1984.....	78
Cuadro 6.17	
Estimación de viviendas, del promedio de personas por vivienda y de la población de 25 cantones, por medio del método de Viviendas. 1996.....	81
Cuadro 7.1	
Población total y número de viviendas por segmento censal muestreado San Rafael de Heredia, 1996.....	84
Cuadro 7.2	
Distribución proporcional de la población de San Rafael de Heredia por sexo y grupos de edad.....	89
Cuadro A 1.1	
Población total e información correspondiente a ocho indicadores sintomáticos para 22 cantones de San José y Heredia, años 1973 y 1984.....	107
Cuadro A 1.2	
Población total e información sobre cinco indicadores sintomáticos por cantón años 1963, 1973 y 1984.....	108
Cuadro A 1.3	
Población total e información sobre cinco indicadores sintomáticos por provincia, años 1963, 1973 y 1984.....	110
Cuadro A 1.4	
Población total e información sobre dos indicadores sintomáticos por distrito, años 1963, 1973 y 1984.....	111
Cuadro A 1.5	
Viviendas, población en viviendas simples y colectivas y abonados del servicio eléctrico por cantón, 1984.....	120
Cuadro A 2.2	
Resultados en la aplicación procedimiento de mínimos cuadrados, a las variables X_u^* y Y_u^*, obtenidos por el paquete estadístico SPSS.....	128

Cuadro A 3.1	
Matriz de correlación de las razones de cambio en las proporciones del tamaño poblacional y de cinco indicadores sintomáticos, a nivel provincial, Período 63-73.....	130
Cuadro A 3.2	
Matriz de correlación de las razones de cambio en las proporciones del tamaño poblacional y de cinco indicadores sintomáticos, a nivel provincial, Período 73-84.....	130
Cuadro A 3.3	
Matriz de correlación de las razones de cambio en las proporciones del tamaño poblacional y de cinco indicadores sintomáticos, a nivel cantonal, Período 63-73.....	130
Cuadro A 3.4	
Matriz de correlación de las razones de cambio en las proporciones del tamaño poblacional y de cinco indicadores sintomáticos, a nivel cantonal, Período 73-84.....	131
Cuadro A 3.5	
Matriz de correlación de las razones de cambio en las proporciones del tamaño poblacional y de dos indicadores sintomáticos, a nivel distrital, Período 63-73.....	131
Cuadro A 3.6	
Matriz de correlación de las razones de cambio en las proporciones del tamaño poblacional y de tres indicadores sintomáticos, a nivel distrital, Período 73-84.....	131
Cuadro A 4.1	
Análisis de los AEP y de los EP en las estimaciones poblacionales a nivel provincial, por el método de Razón Censal, según indicadores sintomáticos utilizados, 1984.....	132
Cuadro A 4.2	
Análisis de los AEP y de los EP en las estimaciones poblacionales a nivel provincial, por el método de Diferencia de Tasas, según indicadores sintomáticos utilizados, 1984.....	132
Cuadro A 4.3	
Análisis de los AEP y de los EP en las estimaciones poblacionales a nivel provincial, por el método de Razón Correlación, según indicadores sintomáticos utilizados, 1984.....	133

Cuadro A 4.4	
Análisis de los AEP y de los EP en las estimaciones poblacionales a nivel provincial, por el método de Tasa Correlación, según indicadores sintomáticos utilizados, 1984.....	133
Cuadro A 4.5	
Análisis de los AEP y de los EP en las estimaciones poblacionales a nivel cantonal, por el método de Razón Censal según indicadores sintomáticos utilizados, 1984.....	134
Cuadro A 4.6	
Análisis de los AEP y de los EP en las estimaciones poblacionales a nivel cantonal, por el método de Diferencia de Tasas, según indicadores sintomáticos utilizados, 1984.....	134
Cuadro A 4.7	
Análisis de los AEP y de los EP en las estimaciones poblacionales a nivel cantonal, por el método de Razón Correlación según indicadores sintomáticos utilizados, 1984.....	135
Cuadro A 4.8	
Análisis de los AEP y de los EP en las estimaciones poblacionales a nivel cantonal, por el método de Tasa Correlación según indicadores sintomáticos utilizados, 1984.....	135
Cuadro A 5.1	
Resultados en los AEP de las estimaciones poblacionales a nivel cantonal, por indicador sintomático, según método utilizado, 1984.....	136
Cuadro A 5.2	
Resultados en los AEP de las estimaciones poblacionales a nivel distrital, por indicador sintomático, según método utilizado, 1984.....	137
Gráfico 6.1	
Gráficos de residuos correspondientes a cuatro modelos de regresión utilizados para estimar la población de los cantones de Costa Rica en 1984, según método utilizado.....	57
Gráfico 6.2	
Gráficos de residuos correspondientes a tres modelos de regresión utilizados para estimar la población de los distritos de Costa Rica en 1984, según método utilizado.....	62

Gráfico 6.3	
Gráficos de residuos correspondientes a los dos modelos de regresión utilizados para estimar la población de los cantones de Costa Rica en 1996, según método utilizado.....	68
Gráfico 6.4	
Agrupación de los cantones de Costa Rica según características Socio-demográficas, 1973.....	77
Figura 4.1	
Esquema resumen del proceso de investigación sobre las estimaciones poblacionales para áreas menores.....	17

INTRODUCCIÓN

1. INTRODUCCIÓN

La incorporación de la variable población en la planificación del desarrollo, convirtió a las estimaciones y proyecciones poblacionales, referidas a áreas de diversos tamaños, en un recurso demográfico necesario. Ellas constituyen el elemento básico para la planeación y preparación de los programas de desarrollo, para la formulación de las políticas de acción gubernamental y privada, así como para la ejecución y evaluación de dichos planes (Rincón, 1990).

Las estimaciones poblacionales se convirtieron, entonces, en un instrumento útil para el diagnóstico de necesidades y para estimar la demanda en la formulación de proyectos. Por ejemplo, el número de habitantes debe conocerse para crear nuevos centros educativos, clínicas de atención médica, campos deportivos, puestos de vigilancia o para implementar servicios generales, como electricidad, agua, teléfono, alcantarillado, recolección de basura, etc. Estos elementos revisten especial interés dentro del proceso de democratización de un país (CELADE, MIDEPLAN, y DGEC, 1990).

El sector privado también requiere de estimaciones poblacionales para establecer sus programas y políticas empresariales de desarrollo futuro. Éstas se convierten en la fuente de información básica de clientes y consumidores potenciales y, a la vez, son el fundamento de sus demandas de mano de obra. Por ejemplo, para determinar las áreas de atracción al construirse un nuevo centro comercial, es necesario, entre otras cosas, contar con estimaciones confiables del número de habitantes, con el fin de buscar un balance apropiado entre inversión y número de consumidores potenciales.

Estimaciones poblacionales eficientes a nivel de áreas menores se tienen fundamentalmente para los años censales. Sin embargo, durante el período intercensal, el cual frecuentemente supera los diez años, quienes deben tomar

decisiones con base en el número de habitantes requieren de información precisa, lo que generalmente no se posee.

En esta materia es poco lo que se ha hecho en Costa Rica. Los intentos por estimar la población suponen que las condiciones del pasado se mantienen en el tiempo, llegándose a estimaciones y proyecciones poblacionales poco confiables. El método más común, y que se utiliza desde hace muchos años, consiste en suponer un crecimiento natural de la población. Es decir, se suma a la población censal los nacimientos y se restan las defunciones ocurridas en el período de interés. No obstante, al no considerar el movimiento migratorio, las estimaciones son cada vez menos confiables, conforme se alejan del año censal. Esto provoca una subestimación en aquellas zonas de atracción poblacional y una sobreestimación en las de expulsión. Pese a esto son las más utilizadas por parte de estudiantes, profesores, investigadores o por instituciones públicas y privadas.

El principal problema en considerar los movimientos migratorios, dentro del proceso de estimación, consiste en la carencia de registros, con el detalle y la calidad necesarios.

Debido a la ausencia de información sobre la migración, desde hace muchas décadas, diversos investigadores han propuesto utilizar los indicadores sintomáticos, como una alternativa razonable al problema. Estos indicadores consisten en variables auxiliares ligadas al crecimiento poblacional, que pueden ser utilizadas para medir sus cambios. Muchas series de datos administrativos pueden ser considerados para este fin; por ejemplo, la matrícula escolar, las declaraciones de impuestos, el número de abonados de los servicios básicos, el número de electores inscritos en el padrón electoral, etc. (Shryock y Siegel, 1976).

Las técnicas que han sido utilizadas para trabajar con estos indicadores son diversas. Obviamente, la precisión de las estimaciones depende de la relación existente entre los indicadores sintomáticos y el tamaño de la población, así como del grado de relación que guarda en el tiempo. Pero lógicamente, también influye la disponibilidad de datos con el detalle debido. La carencia de información sintomática es una de las principales razones por las que no se han aplicado estas técnicas en países latinoamericanos.

El presente estudio ofrece una contribución en este sentido. Para el contexto costarricense se determinan los indicadores sintomáticos que más se asocian con el crecimiento poblacional. También se aplican varios métodos de estimación que han dado resultados aceptables en países desarrollados. Se efectúan estimaciones poblacionales para 1984 (año del último censo), las cuales se comparan contra su valor censal para medir la precisión de cada uno de los indicadores.

Puesto que los indicadores sintomáticos aparecen fundamentalmente para áreas político-administrativas, la investigación se limita a este tipo de áreas. De este modo, la estrategia metodológica a discutir hace énfasis en estimaciones poblacionales a nivel provincial, cantonal y distrital, dependiendo de la información disponible.

Puesto que, además de las estimaciones para el último año es necesario contar con proyecciones poblacionales a nivel de área menor, se presenta un método dinámico, con el que es posible proyectar el número de habitantes a cinco años plazo.

También se recurre a un agrupamiento de áreas de acuerdo con sus características socio-demográficas del pasado, tales como comportamiento migratorio, densidad poblacional, estructura por edad, etc. Con lo anterior se

valora si existen mejoras en la precisión de las estimaciones aplicando las técnicas dentro de cada uno de estos grupos, en vez de hacerlo para todas las áreas en conjunto.

Por otra parte, las necesidades de muchos investigadores y planificadores no se limitan a áreas bien definidas administrativamente. Además, muchas veces se requiere estimar otras características de la población, tales como estructura por sexo y edad, población económicamente activa, mujeres en edad fértil, etc. Por tal razón, se aplica un esquema de muestreo de áreas.

ANTECEDENTES

2. ANTECEDENTES

2.1 Estimaciones y proyecciones a nivel nacional

Desde hace varios años, diversas instituciones públicas de Costa Rica se han preocupado por preparar, en forma unificada, las estimaciones y proyecciones nacionales de población. Ellas se elaboran, por lo general, posterior a un censo de población con el objetivo de evaluarlo y actualizar las proyecciones existentes.

El Ministerio de Planificación Nacional y Política Económica (MIDEPLAN), el Centro Latinoamericano de Demografía (CELADE) y la Dirección General de Estadística y Censos (DGEC), fueron las instituciones que coordinaron esfuerzos en la elaboración de las últimas estimaciones y proyecciones nacionales de población. Sus resultados se publicaron en 1988, comprendiendo el período 1950-2025. También en ese año, publicaron proyecciones y estimaciones poblacionales de las seis regiones territoriales del país para el período 1975-2000 (CELADE, MIDEPLAN y DGEC, 1988).

El método usado para obtener estas estimaciones y proyecciones, se denomina "método de los componentes". Tradicionalmente se ha usado como el modelo principal para elaborar estimaciones demográficas (Rincón, 1990).

2.2 Estimaciones y proyecciones en áreas menores

La Dirección General de Estadística y Censos es el ente gubernamental que, a través del tiempo, se ha encargado de realizar estimaciones poblacionales para distintas áreas político-administrativas, fundamentalmente a nivel distrital. Los resultados están disponibles semestralmente. En ellas se supone un crecimiento natural de la población para todas las áreas. Sin embargo, están sujetas a errores graves en áreas que presentan un fuerte movimiento migratorio. Estas

estimaciones constituyen la fuente de información más importante para investigadores e instituciones que requieren de estimaciones poblacionales a nivel desagregado.

Además de estas estimaciones, se han hecho otros intentos por estimar y proyectar la población, fundamentalmente a nivel cantonal. En 1971, el Instituto de Investigaciones de la Escuela de Ciencias Económicas y Sociales, publicó estimaciones poblacionales a nivel cantonal para el período 1970-1980. Éstas fueron elaboradas por los profesores Carlos Raabe y Rodrigo Umaña. Los métodos utilizados suponen que las condiciones que afectan el crecimiento de la población en las áreas de interés, se mantienen constantes o varían en forma gradual. Esto permite utilizar ciertos modelos matemáticos que emplean las tasas anuales de crecimiento poblacional obtenidas en el período intercensal anterior, de acuerdo con los supuestos sobre el tipo de crecimiento que experimenta el área en consideración. Debido a lo restrictivo de este supuesto para ser aplicado a nivel cantonal, con el pasar de los años se observaron muchas deficiencias en estos cálculos (Raabe y Umaña, 1971).

También el CELADE, junto con MIDEPLAN y la DGEC, ha publicado otras estimaciones y proyecciones de población a nivel cantonal por sexo y grupos de edades. La primera de ellas en 1983 y la segunda en 1990. Esta última se fundamenta en el procedimiento que Louis Duchesne presentó en un seminario internacional sobre proyecciones subnacionales realizado en Girardot, Colombia en 1988. Este método se denominó "Relación de Cohortes" y las estimaciones y proyecciones que se derivan de él surgen como una continuación de tendencias pasadas. El supuesto básico radica en que la evolución experimentada por la población de áreas menores está sujeta a un comportamiento que armoniza con el período anterior considerado (Duchesne, 1989). Los resultados de estas estimaciones y proyecciones son también muy utilizadas, sin embargo, su precisión no ha sido evaluada.

Algunas instituciones privadas efectúan sus propias estimaciones poblacionales, con el fin de determinar la ubicación adecuada para los nuevos proyectos comerciales. Aunque no existen publicaciones al respecto, se sabe, por ejemplo, que cadenas de supermercados, utilizan los clientes del servicio eléctrico para estimar la población de ciertas áreas de interés comercial.

2.3 Limitaciones de las estimaciones poblacionales en áreas menores

Para el caso de áreas menores, la aplicación de los métodos tradicionales de estimación presenta una serie de limitantes. La dinámica demográfica en estas áreas resulta compleja y hace difícil la tarea de efectuar pronósticos en un período poscensal. La problemática se fundamenta en cuatro aspectos básicos:

- 1) Existen poblaciones expuestas a fuertes movimientos migratorios internos, con efectos importantes a corto plazo. A esto hay que sumarle que la información sobre el movimiento migratorio es de calidad deficiente o sencillamente no está disponible.
- 2) Hay áreas con poblaciones reducidas. En éstas los hechos demográficos (nacimientos, defunciones, cambios de residencia habitual) varían mucho de un año a otro. Esto hace que el uso de metodologías demográficas tradicionales para analizar tendencias históricas de la dinámica demográfica resulten inapropiadas (Rincón y Hernández, 1989).
- 3) Las estadísticas y registros demográficos que se tienen para estas áreas son limitadas y deficientes en cuanto a su calidad, debido a problemas con la incorrecta declaración del lugar de residencia (CELADE, MIDEPLAN y DGEC, 1988).

4) Cuanto más desagregadas sean las áreas en consideración, mayor es el número de ellas, por lo que la cantidad de datos que debe manejarse puede ser muy grande. Tal situación hace difícil la producción de estimaciones que sean coherentes en todas sus partes (Rincón y Hernández, 1989).

Estos aspectos evidencian la dificultad de formular hipótesis sobre el comportamiento de las variables demográficas en áreas menores. Hacerlo requeriría un estudio exhaustivo en cada área. Por estos motivos, la creencia de que el futuro es una extensión del pasado, no se aplica necesariamente al caso de estimaciones y proyecciones poblacionales para áreas pequeñas. Se hace necesario introducir otras fuentes de información para realizar estimaciones precisas. La atención debe estar dirigida hacia dónde se pueden encontrar estas fuentes y la forma de relacionarlas con los censos (Lee y Goldsmith, 1982).

2.4 Las variables sintomáticas y su aplicación

Debido a la problemática que conllevan las estimaciones poblacionales a nivel de áreas menores, en países como Estados Unidos o Inglaterra se han sugerido otros procedimientos de estimación. Ellos involucran a los indicadores o variables sintomáticas. Estas variables corresponden a información recolectada regularmente por empresas e instituciones públicas o privadas, para propósitos de orden administrativo, propios de sus actividades rutinarias (Serow y Rives, 1995; Simpson y otros, 1996).

La intención de recolectar estos datos no es estimar poblaciones, por lo que se debe estar pendiente de ciertos problemas como el cambio de las definiciones en el tiempo o a las interrupciones en su recolección, que pueden restringir su aplicabilidad para fines de estimación.

Para utilizar un indicador sintomático en el proceso de estimación es necesario que se cumplan cinco requisitos básicos (Feeney, Hibbs y Gillaspay, 1995):

a) Sensibilidad:

El indicador debe presentar una alta correlación con el tamaño poblacional, de manera que muestre gran sensibilidad a cualquier tipo de cambio en él.

b) Disponibilidad:

Los datos deben estar disponibles periódicamente, mínimo año a año, para cada área menor, ya que mucha información se publica para períodos mayores - trienales o quinquenales- y otra solo aparece a nivel de áreas geográficas mayores.

c) Confiabilidad:

Se debe entender la naturaleza de los datos que se van a utilizar, pues podrían tener problemas ocultos que conllevarían a errores posteriores. Por lo anterior, se deben analizar tendencias históricas en su comportamiento, con el fin de determinar la existencia de patrones fuera de lo común.

d) Estabilidad en definiciones de conceptos y en métodos de recolección:

Muchas veces, a través del tiempo, el cambio en los métodos de recolección de los datos o en sus definiciones, obliga a realizar un análisis que verifique la uniformidad de criterios a lo largo del tiempo. Además, esta misma coherencia debe mantenerse cuando más de una institución es la encargada de recolectar la información.

e) Facilidad de obtención:

Algunas fuentes de datos presentan limitaciones de acceso, debido a su trascendencia, como por ejemplo, las declaraciones de impuestos. Otras, por su

parte, no están disponibles en forma oportuna, sino que su acceso solo es posible hasta años después de su recolección.

Un problema que confrontan muchas, si no todas las variables sintomáticas, es la exactitud de la codificación geográfica, especialmente la especificación precisa del lugar donde vive una persona (residencia habitual). Muchos registros que contienen variables sintomáticas toman, por ejemplo, las direcciones que se brindan para el correo. Desafortunadamente, estas direcciones a veces no coinciden con los límites político-administrativos utilizados en los censos. Esto puede ocasionar que los registros de las variables sintomáticas y el lugar de residencia censal de una persona sean diferentes. Estos problemas pueden tener un impacto sustancial en la calidad de las estimaciones porque sus efectos usualmente son, por naturaleza, sistemáticos y no aleatorios. A menos que los problemas de geocodificación sean corregidos, la población en ciudades grandes tenderá a sobrestimarse, mientras que en ciudades pequeñas es probable que su población se subestime (Feeney, Hibbs y Gillaspay, 1995).

2.5 Indicadores más utilizados

En países desarrollados como Estados Unidos e Inglaterra, existe una gama de indicadores sintomáticos a los cuales se puede recurrir para realizar las estimaciones. Los más utilizados para efectos de estimaciones poblacionales en áreas pequeñas son: matrícula escolar, datos de impuestos, estadísticas vitales, registros de vehículos motorizados, registros de empleo, registros electorales y registros de salud. También existen varias posibilidades para efectuar estimaciones del número de viviendas. Las más comunes son: permisos de construcción de viviendas, clientes del servicio eléctrico, registros de valoración de propiedades y fotografías aéreas (Feeney, Hibbs y Gillaspay, 1995; Simpson y otros, 1996).

2.6 Procedimientos de estimación

La disponibilidad de los indicadores sintomáticos para las áreas de interés y la manera particular en que sean manejados da lugar a técnicas específicas de estimación.

Se han propuesto variaciones al método de los componentes con el objetivo de estimar el saldo migratorio para cada área. Estos procedimientos generalmente usan estimaciones en subgrupos poblacionales, como por ejemplo, población en edad escolar o población electoral. Con ellos se estima, por lo general, la migración de la población total, asumiendo fija la estructura por edad en la población (Serow y Rives, 1995; Murdock, Hwang y Hamm, 1995).

Otras técnicas están dirigidas a determinar el número de viviendas del área menor considerada y efectuar el cálculo con una estimación del número promedio de personas por vivienda para esa área. Muchos procedimientos se recomiendan para estimar las viviendas; los más utilizados hacen uso de los permisos de construcción o clientes de servicios básicos. También se recomiendan otros, tales como fotografías aéreas, utilización de registros de valoración de propiedades e incluso el conteo directo de las viviendas (Smith y Lewis, 1983; Smith, 1995).

Técnicas que utilizan razones y tasas poblacionales han sido muy utilizadas. Las más comunes son el método de razón censal, que recurre al uso de un indicador sintomático para obtener las estimaciones o la técnica de razón correlación, que presenta una aproximación multivariada al problema, utilizando el análisis de regresión múltiple con un conjunto de indicadores sintomáticos previamente seleccionados (Shryock y Siegel, 1976).

3. OBJETIVOS DE LA INVESTIGACIÓN

Los objetivos que se propusieron en la investigación son los siguientes:

- a) Comparar los principales métodos de estimación de la población de áreas pequeñas utilizando indicadores sintomáticos.
- b) Recomendar los métodos y los indicadores sintomáticos más apropiados en el contexto nacional para la realización de las estimaciones poblacionales.
- c) Proponer un esquema dinámico para producir proyecciones poblacionales a nivel de área menor, usando indicadores sintomáticos.
- d) Examinar la viabilidad del uso de un método basado en la cartografía censal y en el muestreo de áreas, para realizar estimaciones del número de viviendas y de la población de las áreas pequeñas.

ESTRATEGIA METODOLÓGICA

4. ESTRATEGIA METODOLÓGICA

En el presente capítulo se describe la estrategia metodológica utilizada para lograr los objetivos del estudio.

El análisis comprendió dos fases básicas. La primera consideró la aplicación de los indicadores sintomáticos y la segunda, la utilización del muestreo de áreas.

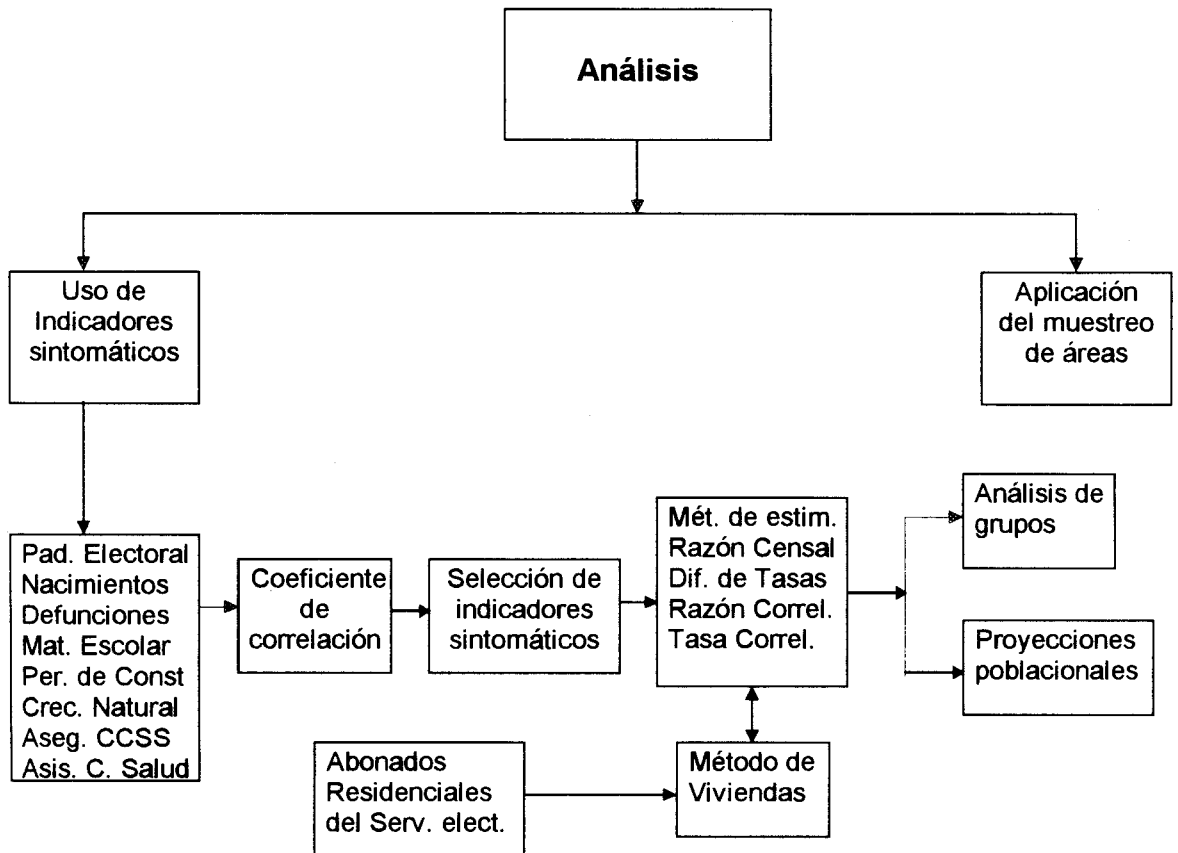
La estrategia con el uso de los indicadores sintomáticos siguió varias etapas. En la primera, se consideraron las fuentes de información existentes en el país y que constituyeron las variables candidatas a indicadores sintomáticos. Dichas variables fueron sometidas a un proceso de selección, de modo que se escogieron aquellas que mostraron estar más fuertemente correlacionadas con el tamaño poblacional. En la segunda etapa, se emplearon algunas técnicas de estimación utilizando los indicadores sintomáticos seleccionados. En este caso se practicaron dos procedimientos que utilizan tasas poblacionales: el método de Razón Censal y el método de Diferencia de Tasas. Además, dos procedimientos que emplean regresión lineal con mínimos cuadrados: el método de Razón Correlación y el método de Tasa Correlación. También se utilizó el método de Viviendas basado en el número de abonados del servicio eléctrico. En la tercera etapa, se realizó una agrupación de áreas de acuerdo con las características socio-demográficas de sus habitantes. En este caso se usó la técnica estadística de análisis de grupos o "Cluster Analysis" para efectuar el agrupamiento y se realizaron estimaciones poblacionales en cada grupo de áreas. Finalmente, mediante las técnicas de estimación que utilizan regresión lineal, se puso en práctica un procedimiento que permitió obtener proyecciones poblacionales a corto plazo.

Por otro lado, en la aplicación del muestreo de áreas, se procedió a seleccionar un cantón que cumpliera ciertas características urbano-rurales, de crecimiento

poblacional y que hubiera presentado un comportamiento migratorio importante en los últimos años. Aquí se utilizó la técnica para analizar su factibilidad.

La Figura 4.1 presenta un esquema del proceso de investigación que se desarrolló en el estudio.

Figura 4.1
Esquema resumen del proceso de investigación sobre las estimaciones poblacionales para áreas menores



A continuación se detallan cada uno de los elementos presentados en el diagrama anterior.

4.1 Indicadores sintomáticos disponibles

Para cumplir con los objetivos del estudio es requisito fundamental contar con información sintomática para las diferentes áreas administrativas con que cuenta el país. Dicha información debió ser sometida a un proceso de valoración para seleccionar aquellas que calificaran como indicadores sintomáticos aceptables.

Debido a las características de los métodos de estimación que se emplearon, se necesitó que los datos utilizados estuvieran disponibles a partir del censo de población de 1963.

Se determinaron nueve fuentes de datos. A continuación se detalla cada uno de ellos.

Matrícula escolar en I y II ciclos: Corresponde al número de estudiantes de la enseñanza primaria matriculados al inicio de cada período lectivo. Esta información estuvo disponible para los diferentes años en que los métodos la requirieron. Sin embargo, para 1963 apareció solo a nivel cantonal, mientras que en los años siguientes se consiguió a nivel de centro educativo.

Nacimientos: Representa el número de nacimientos registrados anualmente en cada área menor. Para los años setenta, la información se encontró a nivel cantonal, mientras que para los años siguientes, a nivel de distrito.

Defunciones: Representa el número de muertes registradas anualmente para cada área menor. Al igual que los nacimientos, en los años setenta se obtuvo a nivel de cantón y para los años restantes, a nivel distrital.

Crecimiento natural: Corresponde a la diferencia entre nacimientos y defunciones. Representa el aumento poblacional de cada área, bajo la hipótesis

de un saldo migratorio nulo. Esta información apareció a nivel distrital para todo el período de interés.

Permisos de construcción de viviendas: Indica el número anual de permisos de construcción de viviendas particulares, los cuales se solicitan en las diferentes municipalidades del país. Su información estuvo disponible únicamente a nivel cantonal.

Clientes residenciales del servicio eléctrico: Se refiere al número de abonados residenciales del servicio eléctrico, por área menor. En las publicaciones del Instituto Costarricense de Electricidad (ICE), referentes a la cobertura eléctrica del país, se indica el número de clientes residenciales a nivel distrital. La información que allí aparece no corresponde necesariamente a los abonados reales, pues el argumento que se maneja es que si el cantón se encuentra 100 % electrificado, entonces el número de clientes es igual al número de viviendas del cantón, lo cual no siempre es cierto, como se observa al comparar estos datos con la información del censo de vivienda de 1984 (ICE, 1986). Por tal razón, la información real sobre esta variable apareció únicamente para los años noventa.

Registros de asegurados de la Caja Costarricense de Seguro Social: Representa el número de asegurados, por área menor, en al menos uno de los regímenes de la Caja Costarricense de Seguro Social (CCSS). Tal información se obtuvo únicamente a nivel cantonal.

Registros de atención en los centros de salud: Corresponde a la información reportada por los diferentes centros del Ministerio de Salud, sobre el número anual de consultas por dicho centro. Estos datos se encontraron a nivel cantonal y no se obtuvieron registros para la década de los sesenta.

Padrón electoral: Señala el número de personas, mayores de edad, que aparecen inscritas en el registro de electores del Tribunal Supremo de Elecciones (TSE). Dicha información está disponible anualmente a partir de los años setenta, a nivel cantonal. Sin embargo, para los años electorales, se cuenta con registros más detallados, con los datos a nivel de distrito administrativo y de distrito electoral.

4.2 Métodos de estimación

A continuación se detallan aquellos procedimientos que, por las características del medio y de la información disponible, pueden dar los mejores aportes en esta materia.

Los cuatro métodos que se discuten producen estimaciones dependientes, es decir, requieren del conocimiento del tamaño poblacional del área mayor, así como del comportamiento de sus indicadores sintomáticos. Además, ellas no se realizan en forma individual, sino para el conjunto de todas las áreas menores que constituyen la unidad administrativa mayor. Por tal razón, las estimaciones obtenidas deben prorratearse, de manera que el conjunto de ellas reproduzca la población total.

4.2.1 Método de Razón Censal y de Diferencia de Tasas

Método de Razón Censal

Se le atribuye a Donald Bogue; sin embargo, él mismo reconoció que la idea del método fue presentada por Whelpton, publicado en 1938 (Bogue, 1950).

El "Método de las Tasas Vitales", como se llamó, fue visto como una manera tosca y rápida de estimar la población en áreas pequeñas. No obstante, llegó a ser una herramienta muy popular en los años 50s y 60s para la confección de

estimaciones poblacionales poscensales de estados, condados y municipios en los Estados Unidos (Voss, Palit, Kale y Krebs; 1995). Varias extensiones de la metodología básica han sido desarrolladas; éstas comienzan con la introducción de otros indicadores sintomáticos junto con los nacimientos y las defunciones (Ericksen, 1973).

El método de Razón Censal considera las tasas de ocurrencia de un determinado indicador sintomático. Supone que las áreas locales presentan una razón de cambio proporcional a la del área mayor, en el período comprendido entre el último censo y el momento de la estimación.

En primer lugar, el procedimiento requiere determinar la tasa de ocurrencia de cada variable sintomática para las distintas áreas menores durante el último período censal. Supóngase que se desea estimar la población basada en la variable sintomática S para una área u , t años después del último censo. Se calcula:

$$r_0(u) = \frac{S_0(u)}{P_0(u)}$$

$r_0(u)$ es la tasa de ocurrencia del indicador sintomático S para el área u en el año correspondiente al último censo.

$S_0(u)$ es la información de la variable sintomática S observado para el área u en el año correspondiente al último censo.

$P_0(u)$ es la población del área u medida en el último censo.

Posteriormente, se busca aproximar la tasa del indicador sintomático al momento t , por medio de un parámetro ϕ , el cual representa la razón de cambio en el indicador sintomático entre el año censal y el momento t , tal que:

$$r_t(u) = \phi r_0(u)$$

Para estimar el valor de ϕ es necesario tener información del área mayor que contenga a u , para la cual se pueden obtener estimaciones independientes de la población en el momento t . Esta área se denota con M y su población en el momento t por $P_t(M)$, donde la estimación de la población de M en t es $\hat{P}_t(M)$. De este modo, se calculan las tasas de ocurrencia de S en el área M , al momento censal y al momento t , por las fórmulas:

$$r_0(M) = \frac{S_0(M)}{P_0(M)} \qquad \hat{r}_t(M) = \frac{S_t(M)}{\hat{P}_t(M)}$$

Suponiendo que la razón de cambio en el indicador sintomático para el área mayor M puede aproximar efectivamente al valor de ϕ , se tiene:

$$\hat{\phi} = \frac{\hat{r}_t(M)}{r_0(M)}$$

Entonces, la estimación de $r_t(u)$ es:

$$\hat{r}_t(u) = \hat{\phi} r_0(u) = \frac{\hat{r}_t(M)}{r_0(M)} r_0(u)$$

Por lo tanto, la estimación poblacional para cada área menor u , en el momento t ,

viene dada por:

$$\hat{P}_t(u) = \frac{D_t(u)}{\hat{r}_t(u)}$$

Usualmente se obtiene más de una estimación de $P_t(u)$, utilizando para ello diferentes indicadores sintomáticos. La estimación final es, generalmente, el promedio simple de las estimaciones individuales, aunque se han propuesto métodos para ponderarlas (Ericksen, 1973; Ericksen, 1974).

La principal ventaja de este método es la simplicidad de aplicación. Además, produce estimaciones que han probado ser bastante exactas (Voss y otros, 1995). Tiene la desventaja que la dirección de las series sintomáticas podría estar

afectada por factores distintos a los que justifican el cambio en la población. Esto significa que la razón para una área pequeña puede cambiar en una proporción diferente y realmente podría, incluso, moverse en dirección contraria a la razón del área mayor; de este modo se incumpliría el supuesto básico del modelo.

Método de Diferencia de Tasas

Este procedimiento es una variación del método de Razón Censal. La principal diferencia entre los dos métodos radica en la forma de estimar la tasa de ocurrencia del indicador sintomático para cada área menor. El método supone que el crecimiento de esta tasa, para el área menor u , es igual al crecimiento experimentado en ella para el área mayor M .

Considerando nuevamente el caso de las estimaciones por medio del indicador sintomático S , la aproximación para la tasa de ocurrencia este indicador en el momento t , para el área menor u , viene dada por la fórmula:

$$\hat{r}_t(u) = r_0(u) + [\hat{r}_t(M) - r_0(M)]$$

donde sus componentes se definen tal y como se especificó anteriormente.

4.2.2 Método de Razón Correlación y de Tasa Correlación

Método de Razón Correlación

Esta técnica fue introducida por Schmitt y Crosetti en 1954. En su forma básica, consiste en distribuir la población de un área mayor en sus áreas menores constituyentes. Este método ha demostrado ser uno de los más exactos en la estimación de población poscensal (Namboodiri, 1972; Mandell y Tayman, 1982; Bureau of the Census, 1973, 1980, 1985, Fenney, Hibbs y Guillaspy, 1995).

El modelo Razón Correlación fue desarrollado utilizando regresión con mínimos cuadrados ordinarios (Fenney, Hibbs y Guillaspy, 1995). Requiere información de la población de un área mayor y sus áreas menores constituyentes, y los datos de uno o más indicadores sintomáticos.

El supuesto básico del modelo radica en la existencia de una relación lineal entre la proporción que representa la población del área menor dentro de la población total del área mayor y la proporción correspondiente a una o más variables sintomáticas. En su forma más simple el modelo se puede expresar como:

$$\frac{P_0(u)}{P_0(M)} = a_0 + a_1 \frac{S_0(u)}{S_0(M)}$$

$P_0(u)$ es la población censal correspondiente a cada área menor u

$P_0(M)$ es la población censal correspondiente al área mayor M

$S_0(u)$ es el valor de algún indicador sintomático en cada área u en el año censal

$S_0(M)$ es el valor del indicador sintomático para el área mayor M en el año censal

a_i ($i = 0,1$) corresponde a los coeficientes del modelo

Los coeficientes anteriores se pueden estimar, considerando la ecuación como un modelo de regresión, mediante el método de mínimos cuadrados. Una vez estimados, el modelo se utiliza para estimar las poblaciones de las diferentes áreas menores, t años después del censo. Para ello se utiliza la información correspondiente al indicador sintomático en ese año.

El criterio anterior supone que la relación entre las proporciones poblacionales y las proporciones de los indicadores sintomáticos, reflejan coeficientes de regresión que permanecen constantes en el tiempo. Sin embargo, debido a lo

restrictivo del supuesto, la experiencia ha mostrado que las estimaciones obtenidas no son de buena calidad (Fenney, Hibbs y Guillaspy, 1995).

Un supuesto más factible es el que supone una relación lineal entre las razones de cambio de las proporciones poblacionales y las razones de cambio de las proporciones de los indicadores, entre dos períodos. Considerando, además, que los cambios en el tiempo en esta relación afectan uniformemente todas las áreas geográficas, se obtiene un modelo más refinado que puede formularse como:

$$Y_u = a_0 + a_1 X_u$$

$$\text{donde: } Y_u = \frac{\frac{P_1(u)}{P_0(u)}}{\frac{P_1(M)}{P_0(M)}} \quad \text{y} \quad X_u = \frac{\frac{S_1(u)}{S_0(u)}}{\frac{S_1(M)}{S_0(M)}} \quad \text{para toda área menor } U$$

P_1 y S_1 corresponden a los valores poblacionales y sintomáticos en un segundo censo.

Esta ecuación indica que la razón de cambio en el tiempo de la proporción poblacional está sistemáticamente relacionada con la correspondiente razón de cambio en las proporciones de los indicadores.

Se pueden introducir al modelo otros indicadores sintomáticos de manera que:

$$Y_u = a_0 + \sum_{j=1}^m a_j X_{u,j} \quad \text{donde} \quad X_{u,j} = \frac{\frac{S_{j,1}(u)}{S_{j,0}(u)}}{\frac{S_{j,1}(M)}{S_{j,0}(M)}}$$

para cada área menor U y cada indicador sintomático j .

Una vez que el modelo de regresión ha sido estimado, resulta muy simple obtener las estimaciones de población para un período poscensal.

El Método de Tasa Correlación

El método de Razón Correlación estima el cambio temporal o razón de cambio en las proporciones poblacionales de las áreas menores. Para ello, utiliza los cambios temporales en las proporciones de los indicadores sintomáticos, los cuales se miden tomando una razón de la proporción entre dos puntos en el tiempo para cada indicador y para la población. Puesto que una información confiable para estos elementos solo se obtiene de los censos, el modelo utiliza la diferencia en el tiempo entre los censos para medir dicho cambio.

Se le critica a esta técnica una inconsistencia en su lógica interna, ya que los coeficientes del modelo se obtienen tomando como referencia el tiempo entre los censos y, sin embargo, las estimaciones se efectúan en tiempos diferentes del período base, el cual es, por lo general, menor al tiempo entre los censos. Esto puede afectar la estabilidad natural de los coeficientes y las estimaciones producidas (Mandell y Tayman, 1982; O'Hare, 1980; Swanson, 1980).

Una solución inmediata a esta crítica consiste en construir un modelo que considere a la fracción de razón del cambio anual, en vez de la fracción utilizada antes. Este modelo es consistente, pues la razón anual de cambio se puede utilizar para "cualquier" número de años.

Se ha propuesto una transformación al método de Razón Correlación, llamada "Tasa Correlación", que se basa en una aproximación exponencial de cambio, y que se realiza tomando el logaritmo natural de cada razón y dividiéndola entre el tiempo transcurrido entre los censos utilizados como referencia (Swanson y Tedrow, 1984). Melec y Katzoff, dieron evidencia de que esta transformación logarítmica reduce la autocorrelación espacial (Malec y Katzoff, 1983).

Formalmente, el modelo se puede expresar por:

$$Y_u^* = a_0 + \sum_{j=1}^m a_j X_{u,j}^*$$

donde : $Y_u^* = \frac{\ln(Y_u)}{k}$ y $X_{u,j}^* = \frac{\ln(X_{u,j})}{k}$

Y_u y $X_{u,j}$ están definidos tal como se indicó para la técnica de Razón Correlación y k es el intervalo de tiempo transcurrido entre los censos.

Después de un experimento en condados del estado de Washington, Swanson y Katzoff señalan tres razones por las cuales se puede suponer que la técnica Tasa Correlación mejora la exactitud, si se la compara con la de Razón Correlación:

- 1) Como la primera se obtiene de una transformación de la segunda, ambas tienen la misma tendencia, y si esta tendencia es determinada durante varios años de estimaciones, entonces el método razón correlación presentará mayores inconsistencias, pues estas extrapolaciones están basadas en una relación estructural de cambio sobre un período mayor de años.
- 2) La transformación de las tasas ofrece una reducción en las variancias, por lo que se puede esperar una mayor exactitud al aplicar la Tasa Correlación.
- 3) Puesto que se ha reducido el efecto de la autocorrelación espacial es de esperar que las estimaciones sean más precisas (Swanson y Tedrow, 1984).

Verificación de supuestos

Una consideración importante que se debe hacer al trabajar con los modelos de regresión que utilizan el método de mínimos cuadrados, consiste en verificar sus supuestos básicos.

Si un modelo $Y_j = a_0 + a_1 X_j + \varepsilon_j$ es válido, el error asociado con cada valor X_j , se supone estadísticamente independiente, es decir $Cov(\varepsilon_j, \varepsilon_k) = 0$, para $j \neq k$.

Por otro lado, los ε_j se supone que se distribuyen normalmente con media 0 y variancia constante σ^2 (Johnson y Wichern, 1988). Los anteriores constituyen los supuestos más importantes de este método. El incumplimiento de alguno de ellos podría traer problemas a las estimaciones. A estos supuestos se debe agregar el de independencia entre las variables predictoras, cuando se trabaja con más de una de ellas. Por tal razón, se debe controlar que las variables utilizadas en conjunto no presenten altas correlaciones.

Por lo anterior, es importante que al aplicar los métodos de Razón Correlación y Tasa Correlación, se realice un análisis de residuos que permita, al menos empíricamente, validar el cumplimiento de dichos supuestos.

4.3 Método de Viviendas

Por medio de este método se pueden obtener estimaciones individuales para cada área menor. Ha sido muy utilizado en los Estados Unidos (U.S. Bureau of the Census, 1978 y 1983), y las pruebas demuestran que, a pesar de su simplicidad, no presenta desventajas con respecto a otros métodos de su misma naturaleza (Smith y Lewis, 1983; Smith y Mandell, 1984).

El supuesto fundamental del método es que la población de una área geográfica se supone igual al número de viviendas ocupadas, multiplicado por el número medio de personas por vivienda, más el número de personas que viven en residencias colectivas (residencias estudiantiles, prisiones, asilos, etc.), es decir:

$$P_t = (V_t \cdot PPV_t) + RC_t$$

P_t es la población total en el momento t

V_t es el número de viviendas ocupadas en el momento t

PPV_t es el número medio de personas por viviendas en el momento t

RC_t es la población en residencias comunes

Puesto que los diferentes elementos de la ecuación anterior rara vez se conocen con exactitud, se deben crear estrategias para estimar sus valores. Para ello, se hace uso de los indicadores sintomáticos disponibles.

Estimación del número de viviendas

Se han utilizado varios métodos para estimar el número de viviendas. La información puede provenir de los permisos de construcción, de los clientes de servicios públicos (abonados de servicios de agua, de electricidad, etc.), registros de valoración de propiedades y fotografías aéreas (Feeney, Hibbs y Gillaspay, 1995; Simpson y otros, 1996). Sin embargo, las variables más utilizadas son los permisos de construcción y los clientes del servicio eléctrico.

Para la aproximación por medio de los permisos de construcción de viviendas, el número estimado de viviendas viene dado por:

$$\hat{V}_t = (V_c + PC_t - D_t) \cdot O_t$$

V_c es el número de viviendas obtenidas en el último censo

PC_t es el número de permisos de construcción otorgado entre el último censo y el momento t

D_t es el número de demoliciones entre el último censo y el momento t

O_t es la razón de ocupación de viviendas en el momento t

Varias investigaciones concluyen que las mejores estimaciones del número de viviendas se obtienen de los abonados del servicio eléctrico (Starsinic y Zitter, 1968; Smith y Lewis, 1980,1983). Ello se debe, principalmente, a la calidad de la información, ya que estos datos son a menudo más precisos que los de permisos de construcción, lo cual permite relacionar directamente las familias con los clientes residenciales activos.

Hay varias formas de estimar el número de familias de acuerdo con los clientes residenciales activos del servicio eléctrico. Una técnica es la que utiliza el cambio neto en clientes como un indicador del cambio neto en familias:

$$\hat{V}_t = V_c + (CE_t - CE_c)$$

V_c corresponde al número de viviendas ocupadas en el último censo

CE_t es el número de clientes residenciales activos del servicio eléctrico al momento t

CE_c es el número de clientes residenciales activos del servicio eléctrico al último censo

La fórmula supone una correspondencia uno a uno entre los cambios en las viviendas y los cambios en los clientes activos del servicio eléctrico. No obstante, esta relación uno a uno no siempre se presenta, por la existencia de viviendas sin servicio eléctrico, medidores residenciales compartidos por más de una familia, medidores calificados como residenciales dedicados a otras actividades de uso no familiar, etc. Todos estos factores pueden causar que el número de clientes residenciales activos del servicio eléctrico difiera del número de familias. Para solventar este problema se puede utilizar la fórmula siguiente:

$$\hat{V}_t = \left(\frac{V_c}{CE_c} \right) CE_t$$

La validez de la fórmula supone que la razón de hogares a clientes residenciales activos permanece constante. Esta razón está sujeta a cambios en el tiempo, aunque en períodos cortos se puede asumir constante. Esta fórmula ha producido buenos resultados, particularmente cuando la razón de clientes a familias es considerablemente diferente de uno.

Otra alternativa para estimar el número de viviendas consiste en utilizar fotografías aéreas y efectuar un conteo simple de viviendas en ellas. Obviamente, presenta una serie de problemas al utilizarlo a nivel urbano, por la dificultad de discriminar las viviendas; sin embargo es una alternativa muy adecuada a nivel rural. El problema fundamental que presenta, radica en el costo económico que conlleva su aplicación.

Personas por vivienda

Existen diversas técnicas para estimar el número de personas por hogar. La más simple consiste en utilizar el mismo número medio de personas por vivienda obtenido en el último censo. Otra estimación surge de una extrapolación lineal de la tendencia en el número promedio de personas por vivienda observado entre los últimos dos censos. Esta relación se muestra con la siguiente fórmula:

$$\hat{PPV}_t = PPV_c + \frac{x}{y}(PPV_c - PPV_{c-1})$$

\hat{PPV}_t corresponde al número promedio de personas por vivienda en el momento t

PPV_c corresponde al número promedio de personas por vivienda en el último censo.

PPV_{c-1} corresponde al número promedio de personas por vivienda en el penúltimo censo.

x es el tiempo transcurrido entre c y t

y es el tiempo transcurrido entre $c-1$ y c

Una tercera alternativa consiste en incorporar datos poscensales y relacionarlos con los resultados del último censo. Este procedimiento se basa en el conocimiento del PPV local del último censo para estimar la tendencia de PPV desde el censo:

$$\hat{PPV}_t = D_t \cdot PPV_c$$

donde: D_t corresponde a la razón de cambio en el PPV desde el último censo.

Estimación del RC_t

El número de residentes en viviendas colectivas se puede obtener de registros administrativos o se puede suponer que no presenta cambios drásticos desde el último censo.

4.4 Valoración de las estimaciones

La mejor manera de evaluar la bondad de las estimaciones producidas por los diferentes métodos, consiste en compararlas con sus valores reales. Sin embargo, debido a que este valor es desconocido, se procedió a efectuar esa comparación con los valores censales. Para ello, fue necesario obtener estimaciones poblacionales para 1984, año del último censo poblacional. Dichas estimaciones se realizaron aplicando los diferentes métodos de estimación, con el uso de cada uno de los indicadores sintomáticos seleccionados y combinaciones de éstos.

Para medir la calidad de las estimaciones producidas se calculó su error porcentual (EP) o porcentaje de desviación respecto al valor censal. Esta medida viene dada por la fórmula:

$$EP = \frac{(\text{Pobl. estimada} - \text{Pobl. censal})}{\text{Pobl. censal}} \cdot 100$$

Un error porcentual positivo indica que la estimación sobreestimó la población censal, mientras que un valor negativo indica una subestimación.

Puesto que las estimaciones son prorrateadas para ajustar la población total del país, el error porcentual promedio toma valores cercanos a cero. Por tal razón, para analizar el comportamiento global de las estimaciones, se requirió calcular el valor absoluto del error porcentual (AEP), determinado por la fórmula:

$$AEP = \left| \frac{(\text{Pobl. estimada} - \text{Pobl. censal})}{\text{Pobl. censal}} \right| \cdot 100$$

De este modo, el promedio de los AEP fue una buena medida para evaluar la calidad de un conjunto de estimaciones, y un buen parámetro de comparación, de los resultados de dos o más grupos de ellas.

Otro elemento importante en el estudio es el sesgo que pueden presentar las estimaciones con respecto al número de áreas sobreestimadas o subestimadas. Se debió comparar las estimaciones que presentaron errores porcentuales negativos y positivos. La existencia de diferencias importantes en estos valores, las cuales se repitieran para las diferentes técnicas, podría inducir la existencia de un sesgo sistemático en las estimaciones.

Por otro lado, se analizó el comportamiento de los AEP con respecto al tamaño del área en consideración. Para medir el efecto que tiene este tamaño, se calculó el promedio de los AEP ponderando por la población censal, esto es:

$$\frac{\sum_j (AEP)_j \cdot P_j}{\sum_j P_j} \text{ donde } P_j \text{ es la población censal del área } j.$$

Si el tamaño del área no afecta las estimaciones, dicho promedio no debe ser muy diferente del promedio simple de AEP.

4.4.1 Precisión de las estimaciones

De acuerdo con lo anterior, la precisión de las estimaciones se midió comparando su valor con la información censal. Para calificar una estimación como buena o mala, primero se consideró que la información censal está sujeta a errores de cobertura. Para el último censo, este error se fue estimado en aproximadamente un 5% (CELADE, DGEC y MIDEPLAN, 1988). En consecuencia, no parece lógico que a las estimaciones poblacionales de áreas menores se les exigiera una extrema calidad. Por esta razón, para efectos del presente trabajo, se consideraron como “buenas” aquellas estimaciones que presentaron errores menores al 5%, “aceptables” las que estuvieran entre 5 y 10% y “malas” las que sobrepasaron el 10%.

4.5 Selección de los indicadores sintomáticos

Como se mencionó al inicio del capítulo, de las nueve variables candidatas a indicadores sintomáticos, se debió descartar a los abonados del servicio eléctrico por problemas de disponibilidad de información. Para valorar el aporte de las restantes, se efectuó un análisis a nivel cantonal en dos etapas.

Primero, se calculó el coeficiente de correlación con el fin de determinar la relación existente entre los tamaños poblacionales de los cantones y la información correspondiente a cada variable, según lo requieran los métodos de estimación.

El coeficiente de correlación entre dos variables X e Y , es una medida del grado de asociación lineal entre ellas. Este valor generalmente se denota con el símbolo ρ y se define en términos de la covariancia entre X e Y y sus variancias correspondientes.

El coeficiente de correlación es un valor entre -1 y 1 ; si ρ es cercano a 1 indica una fuerte asociación lineal positiva entre las variables y si es próximo a -1 , señala una fuerte asociación lineal negativa. Por otro lado, si su valor es cercano a 0 , evidencia que no existe asociación lineal entre ellas o que ésta es muy débil.

Para el caso muestral, el coeficiente de correlación se representa con r y puede ser estimado por la fórmula:

$$r = \frac{S_{XY}}{\sqrt{S_{XX} \cdot S_{YY}}}$$

donde S_{XY} simboliza la covariancia muestral entre X e Y , S_{XX} y S_{YY} corresponden a sus respectivas variancias muestrales (Freund y Walpole, 1990).

Otro elemento importante de esta medida es que su cuadrado, r^2 , permite determinar la proporción de variación de una de las variables explicada por la variación de la otra (Freund y Walpole, 1990). Por ejemplo, si el valor de r^2 entre X e Y es $0,75$, significa que el 75% de la variación de Y es explicada por la variación en X .

Debido a que los indicadores sintomáticos deben ser muy sensibles a los cambios en el tamaño poblacional, únicamente se seleccionaron para el estudio aquellas variables que presentaron un alto coeficiente de correlación con la población. Se utilizó como criterio valores de r superior a $0,60$.

En una segunda etapa se procedió a valorar el comportamiento de cada una de las variables consideradas dentro del proceso de estimación poblacional mediante un análisis de los AEP para cada indicador.

4.6 Análisis de las técnicas de estimación

Para comparar los resultados de las estimaciones se efectuó un análisis de los errores EP y AEP, tal y como se hizo con los indicadores sintomáticos. Para este fin, se realizaron estimaciones a nivel provincial, cantonal y distrital. Se seleccionaron, para cada tipo de área, los modelos de estimación y las combinaciones de indicadores sintomáticos que dieron las estimaciones más precisas. Los mejores modelos del proceso anterior, se utilizaron para obtener estimaciones para 1996.

Desafortunadamente, por problemas con la información sintomática necesaria para utilizar el método de Viviendas, no fue posible someterlo a una prueba en sus estimaciones, como sí se hizo con las otras cuatro técnicas. Sin embargo, a través de los clientes del servicio eléctrico, también se efectuaron estimaciones poblacionales para 1996, las cuales se comparan con las obtenidas por los otros métodos.

4.7 Proyecciones de población para áreas pequeñas

Poseer buenas estimaciones poblacionales, a nivel de pequeñas áreas, resuelve el problema de muchas personas e instituciones que las requieren para sus actividades laborales. Sin embargo, el problema de otros va más allá, pues además necesitan saber lo que se espera en el futuro. Por esta razón, las proyecciones poblacionales cobran especial importancia. En esta sección se presenta un método que permite obtener proyecciones poblacionales en áreas menores, al menos a corto plazo y que fue aplicado para 1984.

El planteamiento utiliza los procedimientos de Razón Correlación y Tasa Correlación con los indicadores sintomáticos que proporcionan los resultados más precisos. Puesto que se pretende obtener proyecciones a corto plazo (no mayores a 5 años), el problema de aplicar estos métodos consiste en buscar la mejor manera de proyectar los valores de los indicadores sintomáticos hasta el año de interés.

Por la dificultad que conlleva trabajar con los datos globales de esos indicadores, se sugiere utilizar una medida más estable en el tiempo, como lo es la proporción que representa cada área menor dentro del total del país, para cada indicador sintomático utilizado. Así, para proyectar el comportamiento de las proporciones del indicador sintomático, se pueden considerar dos posibilidades:

- a) Asumir que para períodos cortos no se producen grandes cambios en ellas. Si esta suposición es válida, se pueden mantener constantes las proporciones para cada uno de los años a proyectar, lo que equivale a repetir los valores obtenidos del último año con que se cuenta información.
- b) Suponer que su comportamiento es regular, por lo que se puede extrapolar su valor por medio de alguna función matemática, con base en la tendencia que ellas han venido presentando.

Realizada la estimación de las proporciones para cada uno de los indicadores sintomáticos, en los diferentes años en que se van a realizar las proyecciones, se siguen las mismas pautas empleadas para obtener las estimaciones.

4.8 Análisis de grupos

Una vez aplicados los diferentes métodos y probados sus resultados, se pueden buscar alternativas tendientes a mejorar su calidad. Un procedimiento que ha

dados buenos resultados en otros países, consiste en realizar un agrupamiento de las áreas menores, de acuerdo con sus características socio-demográficas. Así lo revela un estudio realizado en la Universidad de Wisconsin, en Estados Unidos (Voss y Perry, 1995). Para este propósito, se pueden considerar variables tales como comportamiento migratorio pasado, ruralidad (porcentaje de población en área rural), densidad poblacional, estructura por edad, etc.

Para llevar a cabo el agrupamiento se utilizó la técnica estadística multivariada de análisis de grupos o "Cluster Analysis" (Dillon y Goldstein, 1984). Todos los criterios de agrupamiento de este método están basados en una matriz de las distancias entre pares de áreas menores. El criterio que se usa en el estudio se conoce como el del "vecino más lejano" (o enlace completo). Por medio de éste, las primeras dos áreas que se combinan son las que presentan la menor distancia entre sí. La distancia entre el nuevo grupo formado y las áreas individuales se calcula como la mayor distancia entre el área individual y un área en el grupo. La distancia entre áreas que no han sido juntadas no cambia. En cada paso, la extensión entre dos grupos es la obtenida entre sus dos puntos más lejanos. La medida de distancia utilizada es la "euclidiana", por medio de la cual, la extensión entre dos áreas se obtiene de la raíz cuadrada de la suma de los cuadrados de las diferencias en los valores de cada variable. La distancia entre las áreas menores X e Y viene dada por:

$$\sqrt{\sum_i^n (X_i - Y_i)^2}$$

donde n representa el número de variables en el estudio, X_i e Y_i corresponde al valor de la variable socio-demográfica i correspondiente las áreas X e Y respectivamente (para más información ver Dillon y Goldstein, 1984).

Una vez efectuado el agrupamiento, se procedió a obtener estimaciones en cada uno de ellos, por medio de los modelos que presentaron la mayor consistencia a

nivel general. Para medir el aporte que el procedimiento de agrupamiento dio a la eficiencia de las estimaciones, se compararon sus resultados con los obtenidos por dichos modelos a nivel general.

4.9 Muestreo de áreas

Las técnicas estudiadas anteriormente son formas indirectas de estimar la población. Sin embargo, una forma directa de aproximar el valor poblacional de una área menor es a través de la observación. A continuación se describe un método para estimar la población utilizando el muestreo de áreas. Éste se puso en práctica en un cantón previamente determinado.

Una de las principales ventajas de este método es que, aparte de producir cálculos a nivel global, permite estimar la población para diferentes grupos, tales como sexo, edad, ocupación, población en edad escolar, etc.. Otra ventaja radica en que el área de estimación no se circunscribe a límites político-administrativos, lo que permite obtener estimaciones poblacionales a otros niveles geográficos, fundamentalmente útiles para el comercio.

El esquema de la aplicación del muestreo de áreas en el proceso de estimación de la población total, involucra cinco etapas (Rives, 1982); a saber:

- 1- En la primera se requiere dividir el área geográfica en un número grande "N" de unidades menores. Estas unidades menores constituyen grupos de viviendas y corresponden a las llamadas "unidades primarias de muestreo" (Cochran, 1987). Dependiendo del área en consideración, es posible tomar manzanas o segmentos censales, para los cuales existe cartografía, de manera que simplifique el trabajo. Por las características de Costa Rica, para esta etapa lo ideal es trabajar con los segmentos censales.

2- La segunda etapa requiere seleccionar una muestra de “n” grupos del total del área (N). La selección debe efectuarse con base en los grupos establecidos al momento del último censo, puesto que el modelo base involucra el cambio entre los datos censales y la información obtenida para el período poscensal.

3- La tercera etapa del proceso de estimación requiere determinar la información censal sobre el número de viviendas y de personas que residían en dicho grupo.

4- La cuarta etapa consiste en el denominado “trabajo de campo” (Gómez, 1985). Se debe recorrer cada uno de los grupos o segmentos seleccionados y recoger la información del número de viviendas. Además, cada vivienda debe proveer información sobre el número de personas que usualmente residen en ella, siguiendo la misma definición censal de residencia habitual. Si se requiere información para grupos particulares, entonces se debe pedir la información específica para cada miembro de la vivienda.

5- La etapa final del procedimiento consiste en efectuar los cálculos. Sea y_i la población correspondiente al i-ésimo segmento muestreado y sea X_i la población del mismo segmento obtenida en el censo, entonces:

$$Y = \sum_{i=1}^n y_i \quad \text{y} \quad X = \sum_{i=1}^n X_i$$

corresponden a los totales de la población para los segmentos muestreados al momento de la observación y al momento censal respectivamente.

Si P_c corresponde a la población del área en consideración para el último censo y P_e a la población estimada por medio de la muestra, entonces se tiene que:

$$P_e = \frac{Y}{X} P_c$$

El valor $\frac{Y}{X}$ corresponde al estimador de razón del crecimiento poblacional del período poscensal.

La población estimada P_e está sujeta a errores muestrales y no muestrales del estimador de razón y de los errores de cobertura del censo en el cálculo de P_c . Los errores de muestreo pueden ser controlados efectivamente por medio del diseño de una muestra eficiente y los no muestrales a través del control de la calidad del procedimiento. Los errores de cobertura censal traen mayores problemas, pues no es posible saber cuál fue el error cometido en cada área.

Cuando el número de grupos o segmentos de viviendas es grande, generalmente mayor que treinta, y estos producen coeficientes de variación para los valores promedio de los X_i y de los y_i , relativamente pequeños, generalmente menos del 10%, entonces el estimador de razón $\frac{Y}{X}$ se distribuye aproximadamente normal en muestras repetitivas (Cochran, 1987). Bajo estas condiciones, la variancia para muestras del estimador P_e está dado por la expresión:

$$\frac{N^2(1-f)}{n(N-1)} \sum_{i=1}^n (y_i - Rx_i)^2$$

donde R es la razón verdadera de crecimiento poblacional para el período poscensal y f es la fracción de muestreo ($\frac{n}{N}$). La ecuación anterior supone un muestreo aleatorio simple.

El estimador simple para la variancia poblacional de P_e puede expresarse como:

$$\frac{N^2(1-f)}{n(n-1)} \sum_{i=1}^n (y_i - rx_i)^2$$

Donde r es la razón de crecimiento muestral para el período poscensal. La última ecuación es una expresión aproximada, la cual presenta un sesgo del orden

$$\frac{1}{n} \text{ (Cochran, 1987).}$$

Si se consideran características particulares de la población o de las viviendas, el estimador de la proporción de unidades pertenecientes a una de esas categorías viene dado por:

$$\hat{p} = \frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n Y_i}$$

donde Y_i es el número de elementos en el i -ésimo segmento y a_i corresponde al número de elementos pertenecientes a la categoría de interés en este segmento.

Todo lo expuesto, está dirigido a la población residente en viviendas. Si en el área de interés se encuentran personas que residen en residencias colectivas, tales como prisiones, residencias estudiantiles, conventos, etc., deben ser considerados por separado y agregados a la estimación efectuada.

Dependiendo de las condiciones del área seleccionada, se pueden introducir algunas variaciones al método, como estratificar por áreas, por crecimiento residencial en los últimos años, por características de urbano-rural, etc., para evitar sesgos en las estimaciones.

RESULTADOS

5. PRECISIÓN Y APLICACIÓN SIMPLE DE LOS INDICADORES SINTOMÁTICOS

En el presente capítulo se estudian los diferentes indicadores sintomáticos disponibles. Para analizar en el período 73-84 el comportamiento de cada uno, se calculó, en primer lugar, su coeficiente de correlación con el tamaño poblacional. Para aquellos que mostraron las mayores correlaciones, se midió su precisión comparando el resultado de cada estimación poblacional con los resultados del censo de 1984.

Los cantones representan la menor área administrativa para la cual existe información más completa en el período de interés. En una primera etapa, se consideraron 22 cantones de las provincias de San José y Heredia. Para cada uno de ellos se obtuvo la información sintomática correspondiente a: padrón electoral, nacimientos, defunciones, crecimiento natural, matrícula escolar, asegurados del régimen de Enfermedad y Maternidad de la Caja Costarricense de Seguro Social, asistencia a Centros de Salud del Ministerio de Salud y permisos de construcción de viviendas, para los últimos dos años censales (1973 y 1984). Esta información aparece en el Anexo 1, Cuadro A 1.1. Se determinó, para la población y para cada variable, su razón de cambio con respecto a la proporción que representa cada cantón en el total del país, para el período 73-84.

El Cuadro 5.1 presenta la matriz de correlación para estas razones de cambio. Los mayores coeficientes de correlación con la población se obtienen de los indicadores padrón electoral, nacimientos y crecimiento natural, que están arriba de 0,80. Las defunciones y la matrícula escolar también presentaron un coeficiente de correlación superior a 0,60, por lo que es importante considerar su aporte al proceso de estimación. Las otras tres variables presentan un coeficiente de correlación muy bajo, por lo que se decidió excluirlas del estudio.

Cuadro 5.1

Matriz de correlación de las razones de cambio en las proporciones del tamaño poblacional y de ocho variables sintomáticas, para 22 cantones de San José y Heredia, período 73-84

	Pob.	Pad. Elec.	Nac.	Crec. Nat.	Def.	Matr. Esc.	Asis. C. Salud	Perm. Cons.	Aseg. CCSS
Población	1,00								
Padr. Electoral	0,88	1,00							
Nacimientos	0,87	0,79	1,00						
Crec. Natural	0,83	0,75	0,99	1,00					
Defunciones	0,65	0,58	0,56	0,44	1,00				
Matr. Escolar	0,61	0,36	0,48	0,48	0,21	1,00			
Asis. C. Salud	0,35	0,38	0,10	0,04	0,47	0,12	1,00		
Per. Construc.	0,23	0,07	0,00	-0,02	0,04	-0,02	0,30	1,00	
Aseg. CCSS	-0,31	-0,18	-0,27	-0,27	-0,23	-0,48	-0,07	0,09	1,00

Los indicadores sintomáticos seleccionados se utilizaron en la segunda etapa, con el objetivo de lograr estimaciones poblacionales para todos los cantones del país.

Como ya fue mencionado, los métodos de Razón Correlación y Tasa Correlación, requieren información de los dos censos anteriores al momento de realizar las estimaciones. El cálculo de las estimaciones poblacionales para 1984 exige utilizar los censos de 1963 y 1973. Este hecho obligó a trabajar con los 68 cantones que existían en el país en 1963. Por esta razón, la información correspondiente a los cantones creados después de este año, se incluyó en el lugar de procedencia de cada uno. Este problema no afecta a los métodos de Razón Censal y Diferencia de Tasas, pues ellos solo requieren información del censo de 1973, cuando el número de cantones era de 79. Sin embargo, para efectos comparativos en el análisis, se decidió trabajar con los 68 cantones para todos los métodos.

Por otro lado, buscando mayor estabilidad en los nacimientos, defunciones y crecimiento natural, se decidió calcular un promedio trienal, en vez del valor

simple. Este hecho reduce la variabilidad de estos indicadores sintomáticos. La información correspondiente a los cinco indicadores sintomáticos y a la población en el período de interés, se presenta en el Anexo 1, Cuadro A 1.2.

El Anexo 2 presenta ejemplos que ilustran la aplicación del padrón electoral en el cálculo de estimaciones poblacionales, mediante los métodos Razón Censal, Diferencia de Tasas, Razón correlación y Tasa correlación, para los 68 cantones.

El Cuadro 5.2 presenta el comportamiento de los errores producidos en las estimaciones, para los cinco indicadores sintomáticos considerados. Consistente con los resultados obtenidos en la primera etapa, el padrón electoral, los nacimientos y el crecimiento natural dan los mejores resultados a este nivel. El padrón electoral presentó una clara superioridad con respecto a los otros indicadores. Al utilizar el padrón electoral en las estimaciones poblacionales, el error producido fue a lo sumo de 5%, y ninguno de los otros indicadores se aproximó a este valor. Los nacimientos y el crecimiento natural muestran también buenos resultados. Las estimaciones obtenidas de la matrícula escolar y de las defunciones, producen grandes errores. Porcentajes muy altos de ellas tienen errores porcentuales que superan el 10%, por lo que se debe tener cuidado al utilizar estas estimaciones.

Los resultados del Cuadro 5.2 se ratifican con el comportamiento de las matrices de correlación correspondientes al cambio en las proporciones del tamaño poblacional y de cinco indicadores sintomáticos, para los períodos 63-73 y 73-84 presentados en el Anexo 3.

El Cuadro 5.3 presenta el comportamiento del coeficiente de determinación (R^2) para los métodos de Razón Correlación y Tasa Correlación, de acuerdo con el indicador sintomático utilizado por el modelo. También se señala, mediante una prueba "t" la significancia de cada indicador. El indicador sintomático en cada

caso mostró ser altamente significativo (menor de 0,01). Los resultados de R^2 simplemente ratifican lo expuesto anteriormente, de nuevo el padrón electoral, los nacimientos y el crecimiento natural, presentan los mayores valores.

Cuadro 5.2

Comparación del valor absoluto en el error porcentual de las estimaciones poblacionales por indicador sintomático según método utilizado, 1984.

Método	indicador sintomático				
	Pad. Elec.	Nac.	Crec. Nat.	Mat. Esc.	Def.
Razón Censal					
Promedio de AEP	5,0	8,8	9,5	11,3	15,9
Porcentaje de AEP inf. a 5%	61,8	42,6	36,8	30,9	14,7
Porcentaje de AEP sup. a 10%	13,2	33,8	32,4	51,5	67,6
Diferencia de Tasas					
Promedio de AEP	4,3	8,3	8,3	13,1	17,9
Porcentaje de AEP inf. a 5%	66,2	44,1	42,6	22,1	16,2
Porcentaje de AEP sup. a 10%	5,9	32,4	30,9	58,8	72,1
Razón Correlación					
Promedio de AEP	5,0	6,7	6,8	9,5	10,5
Porcentaje de AEP inf. a 5%	58,8	51,5	42,6	32,4	29,4
Porcentaje de AEP sup. a 10%	13,2	20,6	14,7	41,2	36,8
Tasa Correlación					
Promedio de AEP	4,9	6,4	6,7	9,6	10,4
Porcentaje de AEP inf. a 5%	55,9	51,5	44,1	36,8	27,9
Porcentaje de AEP sup. a 10%	11,8	23,5	23,5	42,6	44,1

AEP representa el valor absoluto del error porcentual

Cuadro 5.3

Valor de R^2 y significancia dada por la prueba "t" para cinco modelos de regresión simple, por indicador sintomático utilizado, según método de estimación, periodo 63-73.

Método	Indicador sintomático utilizado en el modelo				
	Pad Elec	Nac	Crec Nat	Mat Esc	Def
Razón Correlación					
R^2	0,76	0,75	0,69	0,63	0,54
Prueba "t" (valor de p)	0,00	0,00	0,00	0,00	0,00
Tasa Correlación					
R^2	0,74	0,74	0,68	0,62	0,58
Prueba "t" (valor de p)	0,00	0,00	0,00	0,00	0,00

6. APLICACIÓN MULTIVARIABLE DE LOS MÉTODOS DE ESTIMACIÓN

En este capítulo se presentan los resultados cuando se utiliza más de un indicador sintomático a la vez. Para los casos de Razón Censal y Diferencia de Tasas, las estimaciones obtenidas de la combinación de dos o más indicadores se consiguen al promediar las estimaciones simples. Para los métodos de Razón Correlación y Tasa Correlación basta con incluir los indicadores correspondientes en el modelo de regresión múltiple.

Además de la comparación a nivel cantonal, se presenta un análisis de las estimaciones poblacionales a nivel provincial y distrital, las cuales se cotejan con las estimaciones hechas por la DGEN para el mismo año. Con los mejores modelos obtenidos para 1984, se obtienen también estimaciones poblacionales a nivel provincial y cantonal para 1996. También se incluye una aplicación de las técnicas, con la finalidad de producir proyecciones poblacionales y mejorar las estimaciones mediante un agrupamiento geográfico de acuerdo con sus características sociodemográficas. Finalmente, se presenta una aplicación del método de viviendas con base en datos de abonados del servicio eléctrico.

6.1 Estimaciones a nivel provincial

Las provincias fueron las primeras áreas en las que se centró el interés de la investigación (Cuadro A 1.3 del Anexo 1). Se aplicaron los cuatro métodos en estudio. Luego de comparar los resultados con la información censal, se procedió a analizar los errores producidos en las estimaciones. El Cuadro 6.1 muestra las combinaciones de indicadores sintomáticos que ofrecieron la mayor consistencia. Todos los métodos dan resultados aceptables, principalmente cuando se utiliza la información del padrón electoral.

Cuadro 6.1
Comparación del valor absoluto en el error porcentual y del error porcentual en las estimaciones poblacionales a nivel provincial por indicador sintomático según método utilizado, 1984.

Indicadores sintomáticos utilizados	Método			
	Razón Censal	Diferencia de Tasas	Razón Correlación	Tasa Correlación
Padrón Electoral				
Promedio de AEP	1,2	1,4	1,8	1,6
Prom. ponde. de AEP ⁽¹⁾	1,1	1,7	1,3	1,2
Valor Máximo de AEP	2,7	3,1	4,1	3,7
Porcentaje de AEP inf a 1%	28,6	14,3	42,9	42,9
Porcentaje de AEP sup a 5%	0,0	0,0	0,0	0,0
Porcentaje de EP positivos	42,9	28,6	28,6	28,2
Padrón electoral y Crecimiento Natural				
Promedio de AEP	1,3	1,7	2,8	2,7
Prom. ponde. de AEP ⁽¹⁾	1,7	2,4	2,7	2,5
Valor Máximo de AEP	3,7	4,2	6,0	5,7
Porcentaje de AEP inf a 1%	57,1	43,9	28,6	28,6
Porcentaje de AEP sup a 5%	0,0	0,0	14,3	14,3
Porcentaje de EP positivos	42,9	28,6	42,9	42,9
Padrón electoral y Nacimientos				
Promedio de AEP	1,5	2,4	3,6	3,2
Prom. ponde. de AEP ⁽¹⁾	1,4	1,7	1,9	1,8
Valor Máximo de AEP	3,6	4,1	6,8	6,2
Porcentaje de AEP inf a 1%	28,6	28,6	14,3	28,6
Porcentaje de AEP sup a 5%	0,0	0,0	28,6	14,3
Porcentaje de EP positivos	57,1	42,9	42,9	42,9

⁽¹⁾ Prom. ponde. de AEP corresponde al promedio de los AEP ponderados por el tamaño poblacional

Todos los errores de las estimaciones son inferiores al 5%, con un error promedio muy bajo. Los métodos de Razón Censal y Diferencia de Tasas presentan los menores errores. Si se consideran los resultados de las distintas combinaciones de indicadores sintomáticos, el procedimiento de Razón Censal muestra una mayor consistencia. Para los tres casos expuestos, los resultados son muy similares, con un error promedio casi inalterado.

En lo que se refiere a la influencia del tamaño poblacional no se observan grandes diferencias entre el error promedio simple de AEP o el ponderado por el tamaño del área, por lo que a este nivel no parece existir influencia del tamaño poblacional. Los resultados reflejan un sesgo en las estimaciones, pues la mayoría de los modelos presentaron más errores porcentuales negativos que positivos, lo cual implica que hubo más provincias subestimadas que sobreestimadas.

Para los modelos del Cuadro 6.1 que utilizaron regresión, el Cuadro 6.2 indica el valor del R^2 y el resultado de la prueba "t" para determinar la significancia de los indicadores sintomáticos utilizados. Lo limitado del número de provincias da como consecuencia que los resultados no sean muy convincentes. Aunque los R^2 son altos, los coeficientes de regresión no son significativos en los casos en que se utiliza más de un indicador. Esto es evidencia de multicolinealidad entre padrón electoral con los nacimientos y con el crecimiento natural, lo que justifica que a este nivel resulta adecuado utilizar los métodos de Razón Censal y Diferencia de Tasas, en vez de aquellas técnicas que emplean modelos de regresión.

El Anexo 4 complementa la información anterior, presentando un detalle de los resultados en la aplicación de los métodos con otras combinaciones de indicadores sintomáticos.

Cuadro 6.2

Valor de R^2 y significancia dada por la prueba "t" para los tres modelos de regresión múltiple presentados en el Cuadro 6.1, por indicadores sintomáticos utilizados, según método de estimación

Indicadores sintomáticos incluidos	Razón Correlación		Tasa Correlación	
	R cuadrado	Prueba "t" (valor p)	R cuadrado	Prueba "t" (valor p)
Modelo 1	0,78		0,76	
Padrón electoral		0,01		0,01
Modelo 2	0,84		0,85	
Padrón electoral		0,75		0,78
Nacimientos		0,28		0,19
Modelo 3	0,80		0,79	
Padrón electoral		0,46		0,43
Crecimiento Natural		0,56		0,55

6.2 Estimaciones a nivel cantonal

A nivel de cantones, los métodos de Razón Censal y Diferencia de Tasas no presentan los problemas de los otros dos métodos, es decir, que debía limitarse a los 68 cantones existentes en 1963 (El Anexo 5 presenta el análisis de los AEP en las estimaciones poblacionales para los 81 cantones por medio de estas dos técnicas). Sin embargo, en aras de realizar una comparación de la precisión de las cuatro técnicas, se procedió a obtener estimaciones poblacionales para estos 68 cantones, tal y como estaban distribuidos territorialmente en 1963. El Cuadro 6.3 indica la combinación de los indicadores que mostraron la mayor consistencia.

Cuadro 6.3

Comparación del valor absoluto en el error porcentual y del error porcentual en las estimaciones poblacionales a nivel cantonal por indicador sintomático según método utilizado, año 1984

Indicadores sintomáticos utilizados	Método			
	Razón Censal	Diferencia de Tasas	Razón Correlación	Tasa Correlación
Padrón Electoral				
Promedio de AEP	5,0	4,3	5,0	4,9
Prom. ponde. de AEP	4,8	4,3	4,6	4,6
Valor Máximo de AEP	21,2	16,6	18,5	19,4
Porcentaje de AEP inf a 5%	61,8	66,2	58,8	55,9
Porcentaje de AEP sup a 10%	11,8	5,9	13,2	11,8
Porcentaje de EP positivos	36,8	39,7	45,6	45,6
Padrón electoral y Crecimiento Natural				
Promedio de AEP	5,0	4,6	4,1	4,0
Prom. ponde. de AEP	3,7	3,7	3,8	3,8
Valor Máximo de AEP	23,4	21,7	13,0	14,2
Porcentaje de AEP inf a 5%	64,7	67,6	66,2	63,2
Porcentaje de AEP sup a 10%	13,2	10,3	2,9	4,4
Porcentaje de EP positivos	39,7	41,1	50,0	48,5
Padrón electoral y Nacimientos				
Promedio de AEP	4,7	4,7	4,1	4,0
Prom. ponde. de AEP	3,8	4,0	3,9	3,9
Valor Máximo de AEP	21,6	20,7	12,1	12,5
Porcentaje de AEP inf a 5%	66,2	64,7	72,1	72,1
Porcentaje de AEP sup a 10%	10,3	10,3	2,9	2,9
Porcentaje de EP positivos	44,1	39,7	48,5	48,5
Padrón electoral , Nacimientos y Matrícula Escolar				
Promedio de AEP	4,8	5,4	3,8	3,8
Prom. ponde. de AEP	4,1	4,8	3,7	3,6
Valor Máximo de AEP	17,5	17,2	13,5	14,2
Porcentaje de AEP inf a 5%	55,9	57,4	72,1	72,1
Porcentaje de AEP sup a 10%	8,8	14,7	2,9	1,5
Porcentaje de EP positivos	32,4	30,9	47,1	41,2

No hay muchos cambios en los modelos que a nivel provincial ofrecieron las mejores aproximaciones. Al incluir la matrícula escolar junto con el padrón electoral y los nacimientos reduce el promedio de AEP en 0,3 puntos porcentuales y para el caso de Tasa Correlación reduce el número de cantones con AEP mayores al 10%.

La mayoría de los modelos representados en el cuadro señalan un bajo porcentaje de áreas con errores grandes, particularmente, aquellos que se fundamentan en los métodos de Razón Correlación y de Tasa Correlación.

Los resultados revelan cierta influencia del tamaño poblacional del área en las estimaciones obtenidas, para todos los modelos se reduce el error promedio cuando se pondera por el tamaño de la población.

Por otro lado, a nivel cantonal también se presenta un sesgo (errores negativos), pues casi en todos los modelos, el porcentaje de áreas con errores positivos es inferior al porcentaje de áreas que presenta errores negativos. Este elemento se evidencia más entre los métodos de Razón Censal y Diferencia de Tasas.

El Cuadro 6.4 presenta los valores de R^2 para los cuatro modelos del Cuadro 6.3, y la significancia de los coeficientes de cada indicador, mediante la prueba estadística "t".

Todos los modelos considerados muestran un R^2 alto y los indicadores utilizados son significativos al nivel del 5%, como lo muestra la prueba "t", con la excepción de la matrícula escolar. Este resultado señala que el aporte de la matrícula escolar al modelo con el padrón electoral y los nacimientos es mínimo, el R^2 únicamente aumenta únicamente en un punto porcentual.

Cuadro 6.4

Valor de R^2 y significancia dada por la prueba "t" para los cuatro modelos de regresión múltiple presentados en el Cuadro 6.3, por indicadores sintomáticos utilizados, según método de estimación (nivel cantonal)

Indicadores sintomáticos incluidos	Razón Correlación		Tasa Correlación	
	R^2	Prueba "t" (valor p)	R^2	Prueba "t" (valor p)
Modelo 1	0,76		0,74	
Padrón electoral		0,00		0,00
Modelo 2	0,81		0,81	
Padrón electoral		0,00		0,00
Nacimientos		0,00		0,00
Modelo 3	0,80		0,80	
Padrón electoral		0,00		0,00
Crecimiento Natural		0,00		0,00
Modelo 4	0,82		0,82	
Padrón electoral		0,00		0,00
Nacimientos		0,00		0,00
Matrícula Escolar		0,07		0,06

El Anexo 4 complementa la información presentada, señalando el comportamiento de los errores con otras combinaciones de indicadores sintomáticos.

6.2.1 Análisis de residuos en las estimaciones

Por lo expuesto anteriormente, los métodos de Razón Correlación y Tasa Correlación son los que presentan la mayor eficiencia en las estimaciones poblacionales a nivel cantonal. Puesto que ambos procedimientos emplean el análisis de regresión múltiple con mínimos cuadrados, fue necesario analizar los supuestos de este modelo.

Un buen método para juzgar el cumplimiento de los supuestos de linealidad, como el de variancia constante (homogeneidad), consiste en representar gráficamente los residuos contra los valores predichos. Si los supuestos se cumplen, no debe

existir ninguna relación entre los residuos y los valores predichos. Los residuos deben distribuirse aleatoriamente a lo largo de una franja horizontal centrada en cero (Hernández y Bolaños, 1993).

El Gráfico 6.1 muestra los diagramas de residuos contra los valores predichos para cada uno de los modelos señalados en el Cuadro 6.2. Al respecto, no se aprecian tendencias que sugieran algún tipo de patrón sistemático entre los residuos y los valores predichos, principalmente con el método de Tasa Correlación, por lo que se aceptó como válido el supuesto de linealidad. Tampoco se observan incrementos o disminuciones en los residuos que sugieran problemas de heterocedasticidad (o variancia no constante). Por otro lado, los resultados obtenidos no reflejan problemas de multicolinealidad, lo que se confirma al saber que los modelos que utilizan más de un indicador, señala un R^2 superior al 80% y la mayoría de las variables utilizadas son altamente significativas.

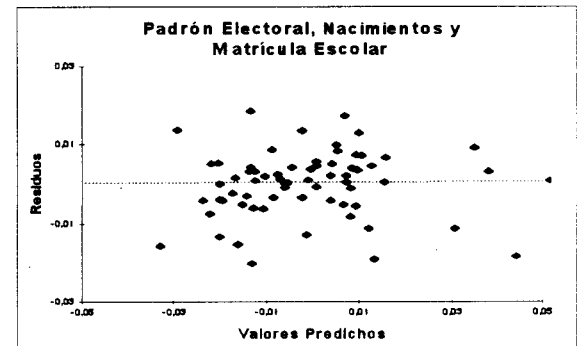
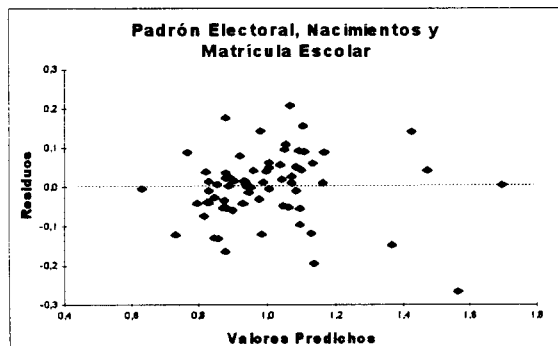
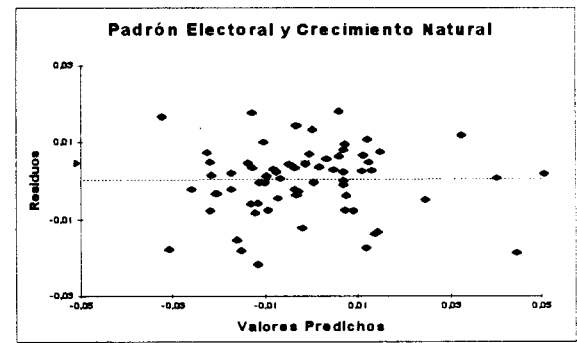
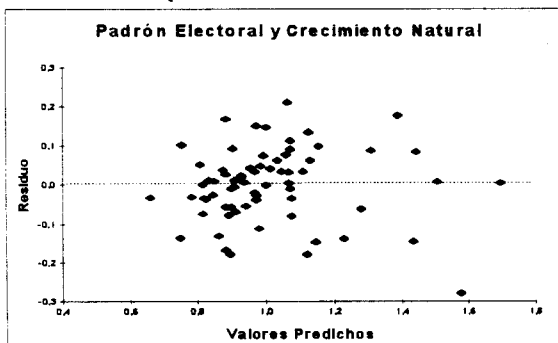
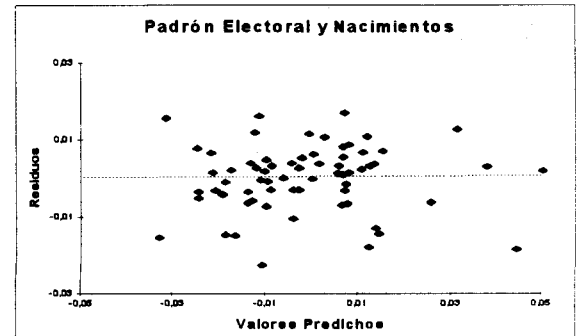
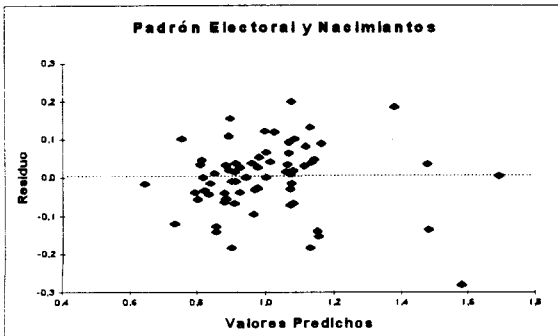
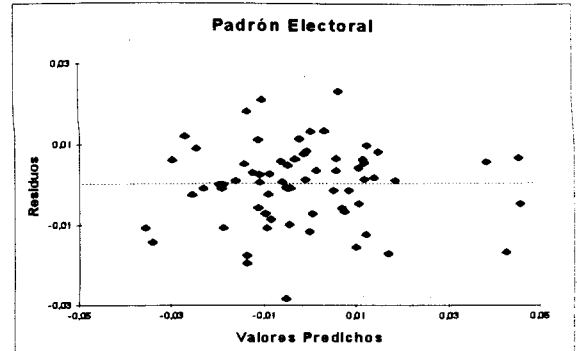
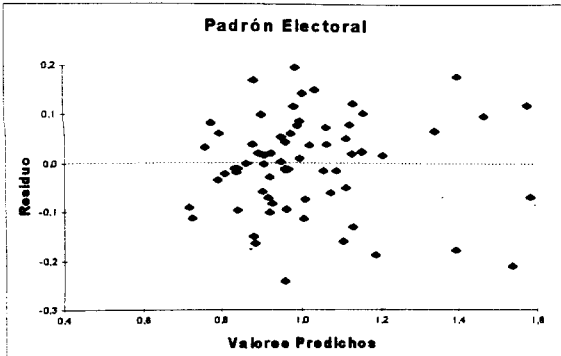
Con respecto al supuesto de independencia estadística de los residuos, el número de datos es grande si se compara con el número de variables utilizadas, por lo que la independencia puede ser ignorada (Hernández y Bolaños, 1993). Para examinar la normalidad de los residuos, el Cuadro 6.5 presenta los resultados de la aplicación del procedimiento “sfrancia” del paquete estadístico “STATA”. Este procedimiento proporciona la prueba Shapiro-Francia para determinar la normalidad de un conjunto de datos (Shapiro y Francia, 1972; Stata Press, 1995). Aunque los residuos no pasaron la prueba de normalidad, con la excepción de los modelos que utilizan únicamente el padrón electoral, los restantes valores de “Pr > z” son bajos, menores a 0,20, por lo que se asume un comportamiento que no dista mucho de la curva normal, ya que son menores a 0,20.

Gráfico 6.1

Gráficos de residuos correspondientes a cuatro modelos de regresión utilizados para estimar la población de los cantones de Costa Rica en 1984, según método utilizado

Método de Razón Correlación

Método de Tasa Correlación



Cuadro 6.5

Resultados de la prueba Shapiro-Francia para determinar la normalidad de los residuos en los modelos de regresión utilizados para estimar la población de los cantones de Costa Rica en 1984

Método de Razón Correlación					
Shapiro-Francia W' test for normal data					
Variable	Obs	W'	V'	z	Pr>z
-----+	-----	-----	-----	-----	-----
paname	68	0.97442	1.692	1.031	0.15121
pacre	68	0.97557	1.616	0.942	0.17301
pana	68	0.97422	1.705	1.046	0.14767
pad	68	0.98205	1.187	0.340	0.36687

Método de Tasa Correlación**Shapiro-Francia W' test for normal data**

Variable	Obs	W'	V'	z	Pr>z
-----+	-----	-----	-----	-----	-----
paname	68	0.97588	1.595	0.918	0.17936
pacre	68	0.97432	1.698	1.025	0.15268
pana	68	0.97595	1.591	0.912	0.18088
pad	68	0.98792	0.799	-0.450	0.67352

paname: Residuos del modelo que utiliza el padrón electoral, los nacimientos y la matrícula escolar

pacre: Residuos del modelo que utiliza el padrón electoral y el crecimiento natural

pana: Residuos del modelo que utiliza el padrón electoral y los nacimientos

pad: Residuos del modelo que utiliza el padrón electoral

6.3 Estimaciones a nivel distrital

Por ser los distritos unidades de menor tamaño, están sujetos a mayores variaciones relativas en los componentes del cambio poblacional, con importantes efectos a corto plazo. Lo anterior, asociado a la carencia de datos, así como a las dudas sobre la calidad de los disponibles, dificulta el proceso de estimación.

A nivel distrital no se contó con toda la información sintomática utilizada para las provincias y los cantones, únicamente se utilizaron dos indicadores: el crecimiento natural de la población y el padrón electoral. Los datos de éste fueron obtenidos por interpolación lineal de los datos referentes a los años electorales.

Por la misma razón discutida para el caso de los cantones, se debió trabajar únicamente con los distritos existentes en 1963, por lo que la información correspondiente a los distritos creados después de ese año debió ser incluida en el distrito de procedencia. De este modo se contó con 333 distritos, para los cuales se estimó la población de 1984. La información correspondiente aparece en el Cuadro A 1.4 del Anexo 1. El Cuadro 6.6 muestra el resumen del comportamiento de los errores producidos con respecto al valor censal.

En términos generales, las estimaciones obtenidas presentan errores bastante elevados, algunos de los cuales superan el 100%. El promedio de AEP supera el 10% en todos los modelos y, aun los mejores resultados, indican que más del 15% de los errores son superiores al 20%.

A pesar de que la información del padrón electoral tuvo que ser estimada, es por medio de este indicador sintomático que se obtienen las mejores estimaciones. No existen grandes diferencias en los resultados obtenidos por los diferentes métodos; sin embargo, las mejores estimaciones se obtienen mediante la aplicación de los métodos de Razón Correlación y Razón Censal.

Una vez más es evidente la influencia que ejerce el tamaño poblacional del área sobre la calidad de las estimaciones. El promedio de los EAP es menor cuando se pondera por el tamaño poblacional, que cuando se calcula en forma simple. Igualmente se repitió el fenómeno del sesgo en las estimaciones, nuevamente el porcentaje de áreas que presentan un error positivo es inferior al porcentaje de ellas con un error negativo.

Cuadro 6.6

Comparación del valor absoluto en el error porcentual y del error porcentual en las estimaciones poblacionales a nivel distrital por indicador sintomático según método utilizado, 1984.

Indicadores sintomáticos utilizados	Método			
	Razón Censal	Diferencia de Tasas	Razón Correlación	Tasa Correlación
Padrón Electoral				
Promedio de AEP	12,1	11,1	11,9	11,5
Prom. ponde. de AEP	11,0	10,3	10,5	10,3
Valor Máximo de AEP	59,2	43,3	46,7	44,9
Porcentaje de AEP inf a 5%	26,1	30,9	27,0	24,9
Porcentaje de AEP sup a 10%	48,0	43,8	46,8	40,8
Porcentaje de AEP sup a 20%	17,1	15,9	18,0	15,0
Porcentaje de EP positivos	40,5	38,4	48,6	45,9
Crecimiento Natural				
Promedio de AEP	25,1	24,0	15,4	16,0
Prom. ponde. de AEP	19,6	18,7	14,6	14,7
Valor Máximo de AEP	114,1	107,2	66,7	69,6
Porcentaje de AEP inf a 5%	14,1	14,1	24,0	21,0
Porcentaje de AEP sup a 10%	75,1	73,0	58,0	57,1
Porcentaje de AEP sup a 20%	27,3	46,8	27,9	30,0
Porcentaje de EP positivos	37,2	37,8	45,6	45,9
Padrón electoral y Crecimiento Natural				
Promedio de AEP	15,9	15,4	11,7	11,7
Prom. ponde. de AEP	12,2	12,1	10,5	10,0
Valor Máximo de AEP	86,1	69,8	46,6	47,5
Porcentaje de AEP inf a 5%	21,9	21,0	30,9	30,9
Porcentaje de AEP sup a 10%	57,1	58,0	45,0	43,8
Porcentaje de AEP sup a 20%	29,1	27,9	16,8	17,1
Porcentaje de EP positivos	38,4	36,3	43,2	43,2

El Cuadro 6.7 muestra algunas estadísticas de interés. El valor del R^2 no fue tan alto como el obtenido para provincias y cantones, pero los indicadores utilizados son altamente significativos, tal como lo muestra la prueba "t".

Cuadro 6.7

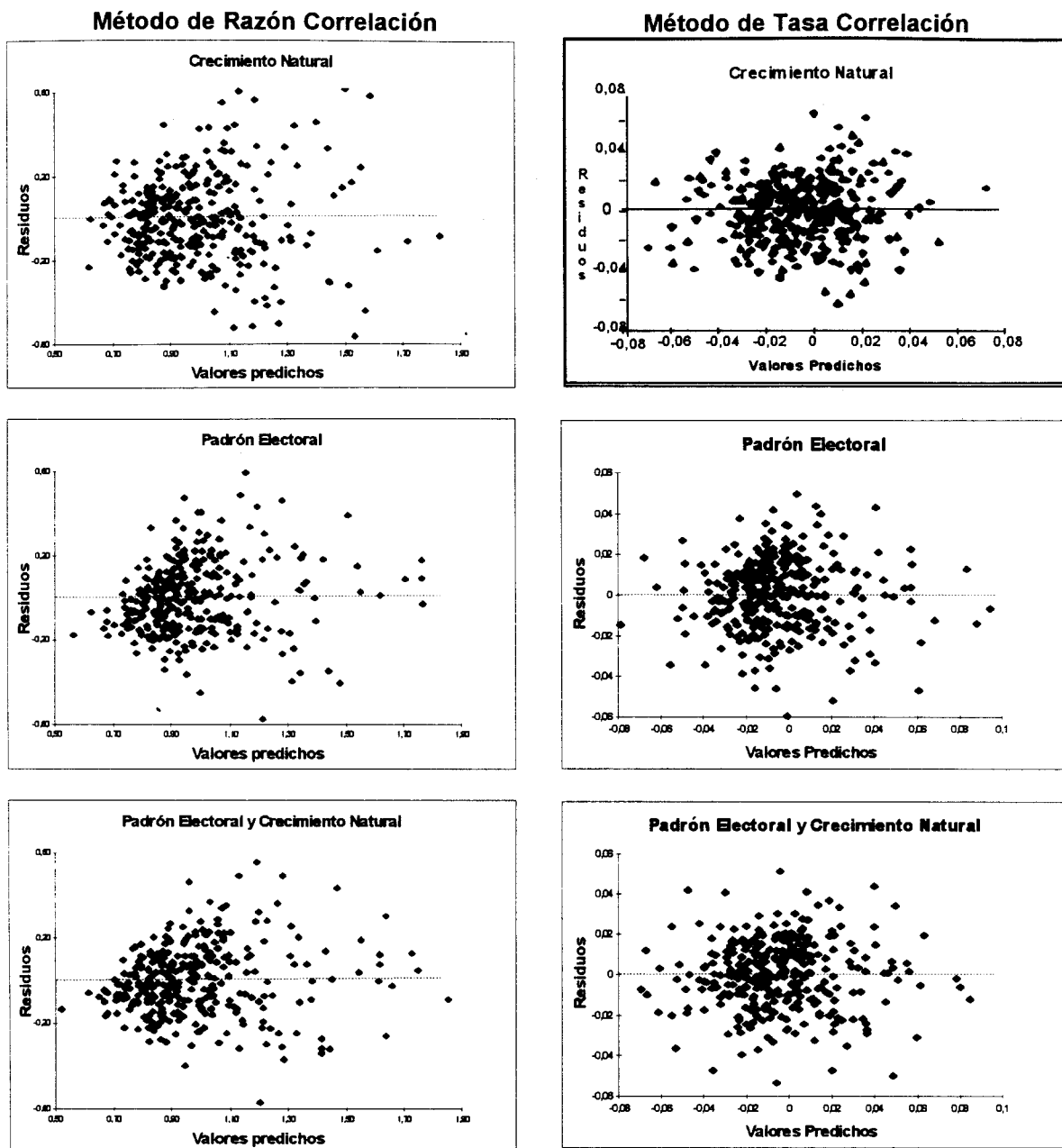
Valor de R^2 y significancia dada por la prueba "t" para los tres modelos de regresión múltiple presentados en el Cuadro 6.5, por indicadores sintomáticos utilizados, según método de estimación

Indicadores sintomáticos incluidos	Razón Correlación		Tasa Correlación	
	R^2	Prueba "t" (valor p)	R^2	Prueba "t" (valor p)
Modelo 1	0,50		0,51	
Crecimiento Natural		0,00		0,00
Modelo 2	0,63		0,63	
Padrón electoral		0,00		0,00
Modelo 3	0,67		0,67	
Crecimiento Natural		0,00		0,00
Padrón electoral		0,00		0,00

6.3.1 Análisis de residuos

Un análisis de residuos, similar al efectuado para los cantones, se llevó a cabo para los distritos. El Gráfico 6.2 presenta los diagramas de residuos contra los valores predichos para cada uno de los modelos del Cuadro 6.6. En este caso se vislumbran varios problemas con los modelos, particularmente con el método de Razón Correlación. Los gráficos no sugieren patrones sistemáticos en el comportamiento de los residuos y los valores predichos, que señalen problemas de falta de linealidad. Sin embargo, todos los modelos del método de Razón Correlación indican cierto incremento en la variabilidad de los residuos conforme el valor predicho aumenta, por lo que se incumple el supuesto de variancia constante para los residuos. Este problema se reduce cuando se utiliza la técnica de Tasa Correlación, y el gráfico correspondiente al último modelo muestra un cambio sustancial en dicho comportamiento.

Gráfico 6.2
Gráficos de residuos correspondientes a tres modelos de regresión
utilizados para estimar la población de los distritos de Costa Rica
en 1984, según método utilizado



Por otro lado, con referencia al supuesto de independencia estadística de los residuos, el número de datos es grande con respecto a la cantidad de variables

utilizadas, por lo que el supuesto puede ser ignorado. Para evaluar la normalidad de los residuos, el Cuadro 6.8 presenta los resultados en la aplicación del procedimiento "sfrancia" del paquete estadístico "STATA. Con la excepción del modelo que utiliza el crecimiento natural, para la técnica de Tasa Correlación, en todas las restantes la prueba fue altamente significativa, apoyando el supuesto de que los residuos se distribuyen normalmente.

Cuadro 6.8

Resultados de la prueba Shapiro-Francia para determinar la normalidad de los residuos en los modelos de regresión utilizados para estimar la población de los distritos de Costa Rica en 1984

Método de Razón Correlación					
Shapiro-Francia W' test for normal data					
Variable	Obs	W'	V'	z	Pr>z
-----+	-----	-----	-----	-----	-----
cre	333	0.96271	9.266	4.692	0.00001
pad	333	0.94845	12.809	5.342	0.00001
pacre	333	0.95673	10.752	4.991	0.00001

Método de Tasa Correlación					
Shapiro-Francia W' test for normal data					
Variable	Obs	W'	V'	z	Pr>z
-----+	-----	-----	-----	-----	-----
cre	333	0.99752	0.615	-1.078	0.85946
pad	333	0.98425	3.913	2.923	0.00173
pacre	333	0.98401	3.972	2.954	0.00157

cre: Residuos del modelo que utiliza el crecimiento natural
pad: Residuos del modelo que utiliza el padrón electoral
pacre: Residuos del modelo que utiliza el padrón electoral y el crecimiento natural

6.4 Estimaciones poblacionales para 1996

Determinados los mejores indicadores y procedimientos metodológicos, resulta de interés aplicar estos elementos en la producción de estimaciones para 1996 y revisar la consistencia de los diferentes métodos. Se escogió un modelo por método, dentro de los que ofrecieron las mejores estimaciones poblacionales para

1984. Además, se trabajó únicamente a nivel provincial y cantonal, ya que estas subdivisiones fueron las que mostraron los mejores resultados.

Puesto que al aplicar estos métodos se requiere una estimación de la población para el área mayor que las comprende, se decidió utilizar el valor de las proyecciones de población para Costa Rica, elaboradas por CELADE, MIDEPLAN y DGEC según la "hipótesis recomendada", que estiman la población de Costa Rica al 30 junio de 1996 en 3442918 personas (CELADE, MIDEPLAN y DGEC, 1988).

El Cuadro 6.9 muestra las estimaciones poblacionales a nivel provincial obtenidas por medio de cada uno de los cuatro métodos desarrollados, utilizando como variable sintomática el padrón electoral. También se señala la mayor diferencia porcentual entre los diferentes pares de estimaciones para cada provincia. Se nota una gran coherencia entre las estimaciones siendo la mayor diferencia presentada apenas un 2,7%.

Cuadro 6.9
Estimaciones poblacionales de las provincias de Costa Rica según método, utilizando como variable sintomática el padrón electoral, 1996

Provincia	Método				
	Razón Censal	Diferencia de Tasas	Razón Correlación	Tasa Correlación	Mayor dif. porcentual
San José	1221902	1218175	1216252	1215247	0,5
Alajuela	620662	621033	621996	621787	0,2
Cartago	403433	403865	405425	405425	0,0
Heredia	311234	311238	314845	315207	0,5
Guanacaste	242816	243065	238537	238393	2,0
Puntarenas	347891	349204	344146	343836	1,6
Limón	294980	296338	301717	303023	2,7
Total	3442918	3442918	3442918	3442918	

Si las estimaciones para 1984 son consistentes, se debería esperar también poca variabilidad en las estimaciones poblacionales de los 81 cantones del país en 1996. El Cuadro 6.10 muestra las estimaciones poblacionales a nivel cantonal obtenidas por medio de cada uno de los cuatro métodos, así como la mayor diferencia porcentual entre los diferentes pares de estimaciones.

Cuadro 6.10
Estimaciones poblacionales de los cantones de Costa Rica según método, utilizando el padrón electoral y los nacimientos. 1996

Cantón	Método				
	Razón Censal	Diferencia de Tasas	Razón Correlación	Tasa Correlación	Mayor dif. porcentual
San José	287878	284962	292512	292250	1,5
Escazú	46906	47397	46923	47020	1,0
Desamparados	169689	169531	168321	168604	0,8
Puriscal	27288	27423	27838	27848	2,0
Tarrazú	13540	13549	13242	13155	2,9
Aserri	43235	43250	43817	44053	1,9
Mora	19101	19539	18861	18838	3,6
Goicoechea	108248	108880	109472	109932	1,5
Santa Ana	27519	27845	27726	27840	1,2
Alajuelita	54727	53669	52925	52513	3,3
V. de Coronado	45809	45985	44407	44192	3,9
Acosta	18709	18793	18971	19008	1,6
Tibás	71181	70705	73289	73684	3,4
Moravia	45508	47301	46373	46641	2,4
Montes de Oca	45990	45671	47502	47714	3,6
Turubares	5179	5259	5395	5422	4,5
Dota	6960	7421	6752	6608	11,0
Cumidabat	58987	60705	57623	57515	5,3
Pérez Zeledón	110225	107410	112302	112920	2,6
León Cortés	10973	11137	10873	10807	3,0
Alajuela	192595	195828	189985	189626	3,2
San Ramón	59720	60487	58811	58608	3,1
Grecia	59452	60379	58103	57734	4,4
San Mateo	4830	5470	4847	4835	1,4
Atenas	19698	20734	19767	19757	4,7
Naranjo	31246	31406	31455	31507	0,8
Palmares	24527	24815	24474	24461	1,4
Poás	21444	21917	21152	21128	3,6
Orotina	14436	15076	14497	14532	3,8

Cuadro 6.10
Continuación

Cantón	Método				
	Razón Censal	Diferencia de Tasas	Razón Correlación	Tasa Correlación	Mayor dif. porcentual
San Carlos	114390	111967	112527	112144	2,1
Alfaro Ruiz	9364	9502	9357	9342	1,7
Valverde Vega	14937	14998	14830	14790	1,4
Upala	40918	39956	41059	41235	2,4
Los Chiles	18585	18296	18666	18736	1,6
Guatuso	13394	13302	13132	13106	2,0
Cartago	124108	125122	123922	124105	1,0
Paraíso	46132	46373	44990	44788	3,4
La Unión	72236	73109	71084	71108	2,8
Jiménez	13864	14032	14290	14342	3,3
Turrialba	62745	63448	64479	64817	3,2
Alvarado	11683	12148	11585	11549	4,9
Oreamuno	34486	34545	34467	34535	0,2
El Guarco	33209	33148	33001	33086	0,6
Heredia	88415	88744	87770	87980	1,1
Barba	27100	26845	27700	27848	2,7
Santo Domingo	30617	31095	31519	31705	3,4
Santa Bárbara	24942	25343	24905	24980	1,7
San Rafael	33622	34643	33407	33427	3,6
San Isidro	12560	12738	12591	12640	1,2
Belén	18864	19513	18527	18480	5,3
Flores	12186	12168	12360	12420	1,9
San Pablo	20551	22712	19929	19807	12,8
Sarapiquí	37682	37331	35598	34871	6,6
Liberia	38190	36790	38455	38570	3,7
Nicoya	41970	42426	43249	43355	3,2
Santa Cruz	37850	38774	38512	38561	1,8
Bagaces	12754	12789	12992	13042	2,2
Camillo	24008	24171	24091	24062	0,5
Cañas	21308	20911	21859	21962	3,0
Abangares	14804	14820	15196	15237	2,8
Tilarán	15369	15102	16268	16323	5,8
Nandayure	10149	10333	10527	10516	3,5
La Cruz	14392	13731	14534	14575	4,6
Hojancha	6707	6820	6897	6907	2,9
Puntarenas	99038	98781	98821	98570	0,3
Esparza	22169	22855	21671	21467	6,1
Buenos Aires	37029	35873	37676	37875	3,1
Motes de Oro	9960	10157	9928	9898	2,5

Cuadro 6.10
Continuación

Cantón	Método				
	Razón Censal	Diferencia de Tasas	Razón Correlación	Tasa Correlación	Mayor dif. porcentual
Osa	27002	26730	28229	28220	4,3
Aguirre	17641	17240	17665	17646	2,3
Golfito	34316	33732	35701	35893	4,4
Coto Brus	39551	38223	41073	41303	4,2
Pamita	10611	10573	11124	11166	5,0
Corredores	36742	35729	37373	37536	2,8
Garabito	7834	7772	7429	7308	6,0
Limón	77424	74758	76725	76674	3,4
Pococí	85633	83161	83982	83858	2,9
Siquirres	51925	52079	50578	50415	3,2
Talamanca	22161	21105	21510	21414	4,8
Matina	30007	29454	28328	27757	5,8
Guácimo	34184	34411	32613	32215	6,4
Total	3442918	3442918	3442918	3442918	

Las diferencias porcentuales, en su mayoría, son inferiores al 5%. Únicamente se presentan los casos extremos de Dota y San Pablo de Heredia, en los que el método de Diferencia de Tasas da una estimación poblacional muy superior a la proporcionada por los restantes tres métodos.

El Cuadro 6.11 muestra los valores R^2 y la significancia de los indicadores sintomáticos utilizados en las estimaciones poblacionales de las provincias y los cantones para 1996. Estos valores reflejan una mayor consistencia que la que se obtuvo para 1984. El R^2 se incrementó en todos los casos y mejoró la significancia de los indicadores. Por ejemplo, para el mismo modelo a nivel provincial, en el caso de los nacimientos, la prueba "t" mostró valores "p" de 0,75 y 0,78 (ver Cuadro 6.2) para las técnicas de Razón Correlación y de Tasa Correlación respectivamente; mientras que para 1996 estos valores bajaron a 0,12 y 0,13 respectivamente.

Cuadro 6.11

Valor de R cuadrado y significancia dada por la prueba "t" para los dos modelos de regresión múltiple utilizados para estimar la población a nivel provincial y cantonal de Costa Rica, 1996.

Indicadores sintomáticos incluidos	Razón Correlación		Tasa Correlación	
	R^2	Prueba "t" (valor p)	R^2	Prueba "t" (valor p)
Provincias	0,99		0,99	
Nacimientos		0,12		0,13
Padrón electoral		0,00		0,00
Cantones	0,92		0,93	
Nacimientos		0,00		0,00
Padrón electoral		0,00		0,00

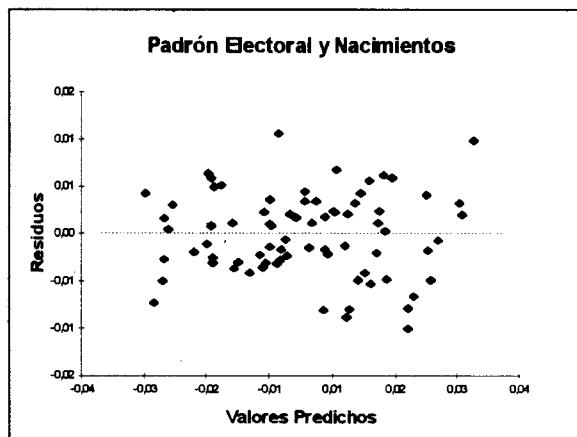
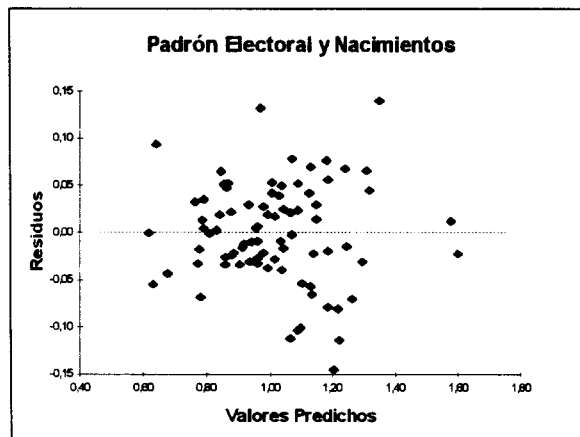
El Gráfico 6.3 muestra el comportamiento de los residuos contra los valores predichos para los modelos de regresión utilizados en las estimaciones a nivel cantonal. Para el caso de Razón Correlación se presentaron problemas de heteroscedasticidad mientras que la técnica de Tasa Correlación no presenta problemas.

Gráfico 6.3

Gráficos de residuos correspondientes a los dos modelos de regresión utilizados para estimar la población de los cantones de Costa Rica en 1996, según método utilizado

Método de Razón Correlación

Método de Tasa Correlación



Para el supuesto de normalidad de los residuos, el procedimiento de Tasa Correlación mostró un mejor comportamiento, tal como lo muestra el Cuadro 6.11. Aunque ninguno de los modelos pasa la prueba de normalidad, los valores de la tabla señalan que los residuos obtenidos por Tasa Correlación se aproximan más a una normal.

Cuadro 6.12

Resultados de la prueba Shapiro-Francia para determinar la normalidad de los residuos en los modelos de regresión utilizados para estimar la población de los cantones de Costa Rica en 1996

Shapiro-Francia W' test for normal data					
Variable	Obs	W'	V'	z	$Pr > z$
Razón correlación	81	0.98852	0.874	-0.273	0.60756
Tasa Correlación	81	0.97901	1.598	0.933	0.17534

Las estimaciones obtenidas, tanto a nivel provincial como a nivel cantonal, son más consistentes que las obtenidas para 1984. Sin embargo, esta afirmación es hipotética hasta tanto no se pueda confrontar con los datos reales, o al menos con los resultados de un censo.

6.5 Comparación de las estimaciones con las obtenidas mediante el método de Crecimiento Natural.

Se ha mencionado que existe una fuerte demanda de estimaciones poblacionales que suponen un crecimiento natural de la población, específicamente las producidas por la DGECC. Por esta razón, se decidió comparar las estimaciones de la DGECC en 1984 con los resultados obtenidos de los indicadores sintomáticos. El Cuadro 6.13 muestra un resumen del comportamiento de los errores producidos al comparar dichas estimaciones con la información censal, para las diferentes áreas administrativas.

Cuadro 6.13

Comparación del valor absoluto en el error porcentual de las estimaciones poblacionales de provincias, cantones y distritos, suponiendo un crecimiento natural, 1984

	Provincias	Cantones	Distritos
Promedio de AEP	7,1	12,4	20,3
Valor Máximo de AEP	19,9	60,7	194,8
Porc. de AEP inf. a 5%	57,1	28,4	15,0
Porc. de AEP sup. a 10%	28,6	44,4	67,3

Al analizar los resultados se pueden observar grandes diferencias en precisión a favor de los que usan información sintomática (Cuadros 6.1, 6.3 y 6.6). Analizando únicamente los valores de AEP, se tiene que para las provincias, con el uso de los indicadores sintomáticos, el Cuadro 6.1 mostró que ese valor fluctuó entre 1 y 2% dependiendo de la técnica de estimación, el cual es muy diferente a 7%, suponiendo un crecimiento natural. En el caso de los cantones, el Cuadro 6.3 indicó que este error promedio fue de aproximadamente 4%, mientras que con el supuesto de crecimiento natural es de 12,4%. A pesar de los inconvenientes obtenidos a nivel distrital, el Cuadro 6.6 mostró que el promedio de AEP aproximadamente de 12%, y con el supuesto de crecimiento natural prácticamente duplica su valor. Esto señala claramente que, al menos para períodos que se alejen notoriamente del año censal, las estimaciones producidas por los indicadores sintomáticos superan las que suponen un crecimiento natural.

6.6 Proyecciones poblacionales

En la presente sección se amplían las técnicas utilizadas, con el fin de obtener proyecciones poblacionales a corto plazo. Para valorar su precisión, se realizaron dos proyecciones poblacionales a nivel cantonal para 1984. A diferencia de las estimaciones anteriores basadas en la información sintomática de 1984, en este caso se utilizó la información de los indicadores sintomáticos hasta el año 1979, y

se proyectó el comportamiento de los indicadores sintomáticos cinco años después, hasta el año 1984. Para tal efecto, se usaron los dos supuestos mencionados en el Capítulo 4, con respecto al comportamiento de la proporción de cada cantón dentro del total del país, para cada indicador sintomático en ese año. El primero consistió en mantener constantes dichas proporciones en el valor de 1979. El segundo supuso que estas proporciones tienen un comportamiento lineal y, que por lo tanto, se puede extrapolar su valor para 1984. Luego de efectuar las consideraciones anteriores, la aplicación de los métodos es la misma que se utilizó para anteriormente.

Se aplicó el método de Tasa Correlación con la información del padrón electoral y de los nacimientos. Debido a la fuerte variación que a nivel cantonal presentan los nacimientos año a año, se optó por calcular el promedio de las proporciones para 1977, 1978 y 1979.

El Cuadro 6.14 muestra un resumen del comportamiento de los AEP para las proyecciones obtenidas.

La calidad de las proyecciones difiere de la obtenida para las estimaciones y que se mostró en el Cuadro 6.3; sin embargo, algunos de los modelos dieron un grado aceptable de precisión. Por ejemplo, si se supone que las proporciones permanecen constantes en ese período de cinco años, y con la aplicación de ambos indicadores para proyectar la población, se obtuvo un AEP promedio de 6,5%, y el 73,6% de los cantones con errores inferiores al 10%.

Estas cifras se pueden considerar aceptables porque superan las estimaciones de la DGEC en ese año, para las cuales solo el 56% de los cantones presentó errores inferiores al 10% y un promedio de AEP de 12,4%.

Cuadro 6.14
Comparación del valor absoluto en el error porcentual en las proyecciones
poblacionales a nivel cantonal, por indicador sintomático
según hipótesis empleada, 1984.

	Hipótesis Constante			Hipótesis Lineal		
	Nac.	Pad.	Nac.- Pad.	Nac.	Pad.	Nac.- Pad.
Promedio de AEP	7,1	8,0	6,5	11,7	11,0	8,2
Valor Máximo de AEP	20,8	35,4	25,1	55,7	36,6	45,6
% de AEP inf. a 5%	52,9	42,6	51,5	39,7	25,0	47,1
% de AEP sup. a 10%	29,4	26,4	26,4	41,2	27,9	32,4

Utilizando el principio anterior y la información sintomática referente a los nacimientos y al padrón electoral hasta el año 1996, el Cuadro 6.15 muestra proyecciones poblacionales para los 81 cantones del país, en el período 97-2001. Puesto que para 1984 se obtuvieron mejores resultados suponiendo que las proporciones de los indicadores se mantenían constantes, se decidió realizar dichos cálculos manteniendo constantes los valores de 1996. Al igual que para las estimaciones poblacionales, la aplicación del método de Tasa Correlación requiere del conocimiento de la población de Costa Rica para cada año a proyectar. Dichos valores se estiman por medio de las proyecciones de población, elaboradas por el CELADE, MIDEPLAN y DGEC según la "hipótesis recomendada", que estiman la población de Costa Rica al 30 junio para los distintos años involucrados (CELADE, MIDEPLAN y DGEC, 1988).

Cuadro 6.15
Proyecciones poblacionales de los cantones de Costa Rica
Período 1997-2001

Cantón	Año				
	1997	1998	1999	2000	2001
San José	294131	303681	312821	315218	330692
Escazú	47905	48859	49797	50715	51728
Desamparados	172926	175198	177538	181854	182677
Puriscal	28016	28937	29817	30036	31539
Tarrazú	13451	13670	13890	14189	14356
Aserri	44931	45776	46613	47515	48347
Mora	19270	19575	19882	20319	20535
Goicoechea	111727	114232	116676	118572	121629
Santa Ana	28364	28929	29485	30028	30629
Alajuelita	54176	54567	55014	56640	56133
V. de Coronado	45853	45921	46068	47666	46622
Acosta	19212	19751	20271	20502	21298
Tibás	74464	76566	78590	79475	82592
Moravia	47496	48465	49416	50307	51367
Montes de Oca	48055	49580	51041	51463	53903
Tumubares	5458	5634	5803	5848	6133
Dota	6703	6867	7026	7127	7346
Currubabat	59671	59765	59963	62036	60695
Pérez Zeledón	114735	117337	119874	121795	125005
León Cortes	10959	11229	11492	11656	12017
Alajuela	193882	197042	200213	204529	206928
San Ramón	59859	60900	61939	63214	64115
Grecia	59083	59992	60911	62272	62873
San Mateo	4886	5024	5156	5215	5419
Atenas	20012	20530	21030	21310	22029
Naranjo	31950	32739	33504	33983	35038
Palmares	24858	25418	25965	26384	27072
Poás	21633	21954	22279	22788	22980
Orotina	14777	15101	15417	15674	16059
San Carlos	114639	116530	118425	120958	122430
Alfaro Ruiz	9472	9708	9936	10077	10392
Valverde Vega	15036	15369	15694	15952	16354
Upala	42360	42847	43359	44475	44512
Los Chiles	19302	19468	19652	20208	20091
Guatuso	13673	13618	13598	14136	13656
Cartago	126471	128959	131412	133859	136465
Paraíso	46071	46540	47041	48308	48201
La Unión	73562	73890	74318	76697	75532

Cuadro 6.15
Continuación

Cantón	Año				
	1997	1998	1999	2000	2001
Jiménez	14433	14903	15352	15469	16231
Tumialba	65521	67352	69116	69911	72606
Alvarado	11744	12001	12252	12457	12763
Oreamuno	35204	35885	36559	37249	37947
El Guarco	34002	34380	34779	35686	35685
Heredia	90469	91421	92434	94895	94758
Barba	28449	28938	29425	30037	30449
Santo Domingo	32125	32945	33739	34197	35324
Santa Bárbara	25558	25957	26359	26944	27217
San Rafael	34131	34734	35335	36054	36590
San Isidro	12919	13134	13348	13633	13802
Belén	18946	19203	19467	19932	20042
Flores	12624	12906	13180	13396	13738
San Pablo	20439	20581	20746	21363	21160
Sarapiquí	36245	36235	36297	37612	36643
Liberia	39196	40078	40938	41601	42679
Nicoya	43560	45051	46475	46763	49251
Santa Cruz	38877	40069	41212	41591	43456
Bagaces	13191	13552	13900	14067	14591
Carrillo	24354	25003	25631	25954	26879
Cañas	22186	22821	23432	23688	24640
Abangares	15337	15833	16309	16435	17238
Tilarán	16329	16961	17563	17605	18726
Nandayure	10502	10927	11331	11342	12110
La Cruz	14784	15145	15496	15720	16199
Hojancha	6936	7177	7408	7450	7856
Puntarenas	99920	102425	104851	106316	109708
Esparza	21885	22307	22724	23155	23586
Buenos Aires	38474	39356	40215	40851	41952
Motes de Oro	10033	10285	10529	10675	11016
Osa	28141	29324	30448	30438	32614
Aguirre	17883	18337	18776	19034	19652
Golfito	36181	37297	38366	38714	40465
Coto Brus	41809	42918	43990	44549	46121
Parrita	11188	11603	11998	12043	12764
Corredores	38036	39004	39941	40486	41812
Garabito	7732	7594	7491	7882	7373
Limón	78273	79673	81064	82699	83970
Pococí	87354	87138	87121	90448	87674

Cuadro 6.15
Continuación

Cantón	Año				
	1997	1998	1999	2000	2001
Siquirres	52154	52387	52692	54377	53553
Talamanca	22217	22252	22307	22363	23062
Matina	28856	28843	29073	28918	29894
Guácimo	33410	33475	33525	33658	34695
Costa Rica	3510637	3577579	3644120	3710656	3783879

6.7 Uso de variables socio-demográficas para mejorar las estimaciones

Luego de realizar estimaciones poblacionales considerando, en conjunto, todos los cantones del país, surgió la inquietud si al realizar algún tipo de agrupamiento de áreas con ciertas particularidades comunes, mejoran las estimaciones. Por tal motivo, se consideró importante efectuar un agrupamiento de áreas de acuerdo con ciertas características socio-demográficas, las cuales intervienen directa o indirectamente en su crecimiento poblacional. Las variables consideradas son: tasas de inmigración, tasa de emigración, porcentaje de población que reside en zona rural, grado de alfabetismo, densidad poblacional, porcentaje de población inactiva, porcentaje de población activa desocupada y porcentaje de personas con edad entre 5 y 30 años. Toda esta información se obtuvo de los censos de población.

Para dicho análisis, se decidió, por razones de disponibilidad de información sintomática, trabajar con los cantones.

Las técnicas utilizadas hasta ahora hacen referencia a dos principios metodológicos básicos; uno que utiliza las tasas de ocurrencia de los indicadores sintomáticos y otro que aplica el análisis de regresión múltiple. Por esta razón,

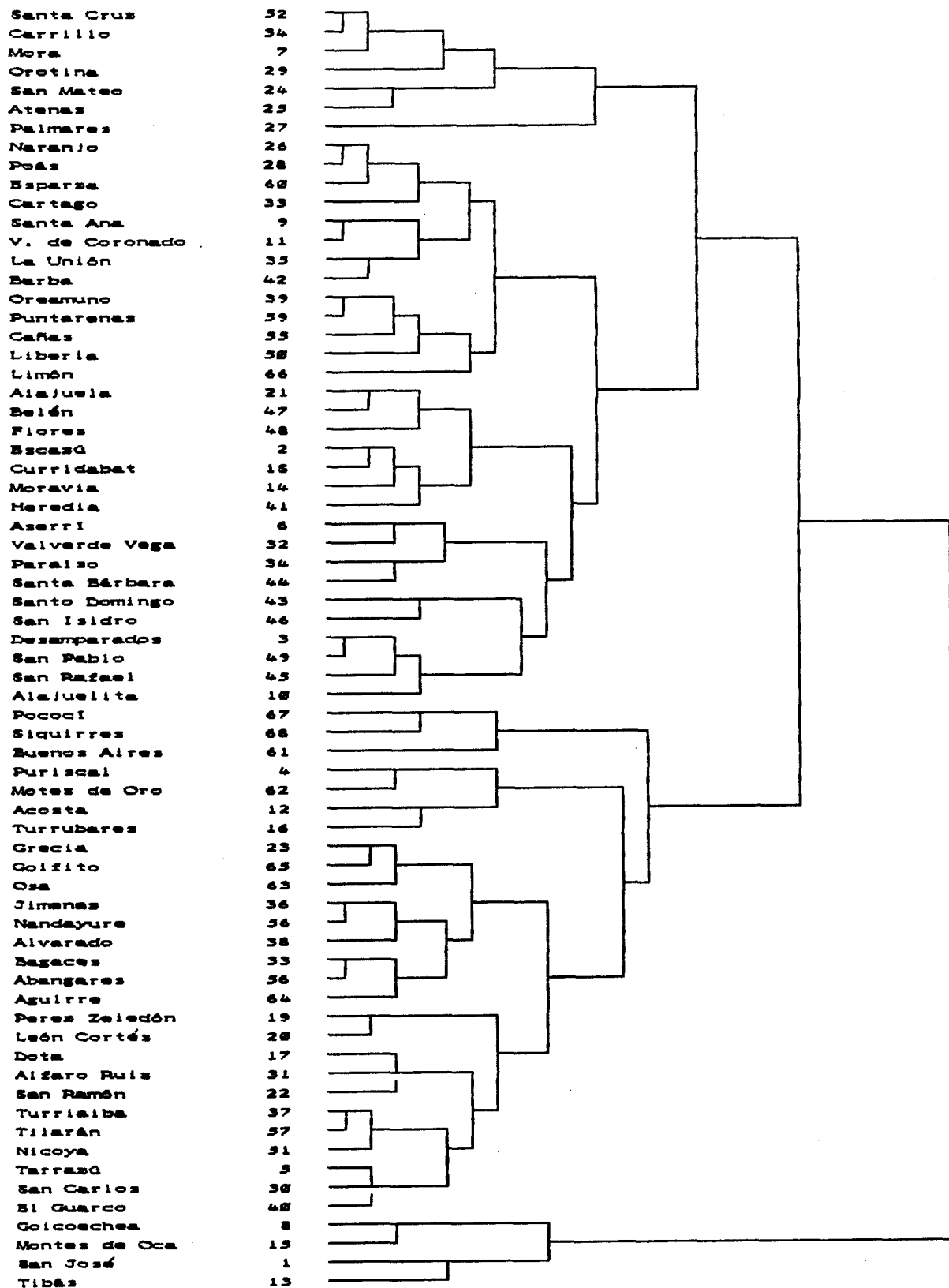
para el análisis de grupos, se decidió usar un método de cada tipo con el fin de simplificar el trabajo. De los métodos que utilizan tasas de ocurrencia se escogió la técnica de Diferencia de Tasas y de los que utilizan regresión múltiple, el de Tasa Correlación.

Debido a que la cantidad de cantones es pequeña, no sería adecuado separarlos en un número muy grande de grupos, pues no se podrían correr regresiones si los grupos tienen pocos datos. Si el agrupamiento fuese efectivo, las estimaciones poblacionales deberían ser más exactas con respecto a su valor censal, que aquellas obtenidas sin agrupar. Para lograr mejores resultados en la aplicación del procedimiento, se estandarizaron las variables. Luego se utilizó el procedimiento "Cluster" del paquete estadístico "SPSS. El gráfico 6.4 muestra el esquema de agrupamiento dado por la computadora. Dentro de las alternativas de agrupación posibles, se escogieron tres grupos. El Cuadro 6.16 muestra las estimaciones obtenidas por medio del proceso de agrupamiento y, además, los resultados para los mismos grupos cuando se trabajó sin agrupar.

A pesar de que las estimaciones producidas dentro de los grupos son aceptables, los resultados de las estimaciones obtenidas sin agrupar son, en términos generales, mejores.

Por otro lado, el hecho de trabajar con grupos y por las características de dependencia de las estimaciones producidas, se tiene la dificultad de tener que efectuar una estimación poblacional previa para cada uno de los grupos, lo cual hace incrementar los errores individuales para los diferentes cantones.

Gráfico 6.4
Agrupación de los cantones de Costa Rica según características
Socio-demográficas, 1973



Cuadro 6.16

Comparación del valor absoluto en el error porcentual de las estimaciones poblacionales a nivel cantonal, por grupo seleccionado, según método utilizado, 1984.

Grupo	N	En grupos		Sin agrupar	
		Diferencia de Tasas	Tasa Correlación	Diferencia de Tasas	Tasa Correlación
Grupo 1	(4)				
Promedio de AEP		2,7	(1)	2,4	4,5
Valor Máximo de AEP		4,0	(1)	3,0	7,4
Porcentaje de AEP inf a 5%		100,0	(1)	100,0	75,0
Porcentaje de AEP sup a 10%		0,0	(1)	0,0	0,0
Grupo 2	(37)				
Promedio de AEP		4,6	5,3	4,6	4,0
Valor Máximo de AEP		17,0	18,6	16,7	12,53
Porcentaje de AEP inf a 5%		67,6	48,6	67,6	75,5
Porcentaje de AEP sup a 10%		8,1	13,5	10,8	2,7
Grupo 3	(27)				
Promedio de AEP		4,8	4,8	5,2	3,9
Valor Máximo de AEP		18,6	16,5	21,6	11,4
Porcentaje de AEP inf a 5%		63,0	44,4	55,5	66,7
Porcentaje de AEP sup a 10%		11,1	22,2	11,1	3,7

⁽¹⁾ Por constar solamente de cuatro cantones no es conveniente correr un modelo de regresión. Los modelos sin agrupar corresponden a estos métodos, cuando se trabajó con todos los cantones a la vez.

Aunque se realizaron otros esfuerzos de mejorar las estimaciones poblacionales, por medio del análisis de grupos, ensayando diferentes combinaciones de variables socio-demográficas, los resultados fueron de menor calidad que la expuesta.

6.8 Aplicación del método de viviendas

La ventaja principal de este método sobre los procedimientos anteriores, radica en que produce estimaciones independientes, por lo que no requiere del tamaño poblacional de una área mayor.

Desafortunadamente, por problemas con la información sintomática, no fue posible examinarlo tal como se hizo con los otros métodos.

Para su aplicación se utilizaron los abonados del servicio eléctrico de 1996, a nivel de cantones. Sin embargo, no se pudo encontrar registros sobre el número real de abonados del servicio eléctrico por cantón para 1984. Se aproximó este valor con los datos del censo de vivienda del mismo año sobre el tipo de alumbrado utilizado en la residencia. Para propósitos de estimación se debe suponer que el número de abonados residenciales para cada cantón, en 1984, fue igual al número de viviendas con energía eléctrica proveniente del servicio público. No obstante, esta relación “uno a uno” no necesariamente es válida, como ya se mencionó en el Capítulo 4.

Para 1996 no hubo problemas con la información y se utilizó el reporte al 30 de junio, que contiene los datos que las distintas empresas distribuidoras de energía eléctrica suministran al ICE sobre el número de clientes residenciales por cantón.

Se procedió a determinar la razón de viviendas a abonados eléctricos para 1984, que al multiplicarla por el número de abonados eléctricos del año 1996, da una estimación de viviendas para dicho año.

La información del año 1984, no permitió trabajar con el 100% de los cantones del país, porque la cobertura eléctrica en muchos de ellos era sumamente baja. Por

esta razón, se decidió utilizar únicamente aquellos cantones que, según el ICE, en 1984 tenían una cobertura eléctrica mayor al 95% (ICE, 1986).

Debido a que todos los cantones seleccionados corresponden a la Región Central del país, se utilizó la Encuesta de Hogares de 1995, que estima en 4,19 el número de personas por vivienda. El mismo valor, para 1984, fue de 4,69. De este modo, la razón de crecimiento para 1996 se estima en 0,89. Dicha razón se aplicó a los promedios de personas por vivienda obtenidos para cada cantón, según el censo de 1984.

Como el número de personas en viviendas colectivas es de esperar que no se modifique drásticamente, se decidió utilizar la misma información del censo de 1984. En el Anexo 6 se muestra un ejemplo que ilustra la aplicación del método de viviendas para el cantón Central de San José.

El Cuadro 6.17 muestra los resultados de las estimaciones poblacionales. Las estimaciones presentaron cierta similitud con las del método de Tasa Correlación; sin embargo, para algunos cantones la diferencia es relativamente grande. No obstante, se evidencia que dicho indicador puede dar grandes aportes en materia de estimación.

Cuadro 6.17

Estimación de viviendas, del promedio de personas por vivienda y de la población de 25 cantones⁽¹⁾, por medio del método de Viviendas. 1996

Cantón	Método de Viviendas			Tasa Correlación		Diferencia porcentual
	Viviendas	Promedio Personas por Viv.	Personas en Viv. colec.	Población estimada	Población estimada	
San José	74459	3,89	2039	291516	292250	-0,3
Escazú	12084	4,25	124	51498	47020	9,5
Desamparados	40534	4,22	193	171035	168604	1,4
Goicoechea	26684	3,97	493	106473	109932	-3,1
Santa Ana	7157	4,28	71	30690	27840	10,2
Alajuelita	11910	4,26	23	50700	52513	-3,5
Coronado	11180	4,32	121	48366	44192	9,4
Tibás	16918	4,21	53	71223	73684	-3,3
Moravia	11364	4,06	140	46317	46641	-0,7
Montes de Oca	14538	3,65	292	53334	47714	11,8
Curidabat	13832	4,12	77	57116	57515	-0,7
Alajuela	48130	4,18	2109	203044	189626	7,1
Cartago	26193	4,60	497	121000	124105	-2,5
Paraíso	10979	4,64	0	50943	44788	13,7
La Unión	16080	4,49	578	72793	71108	2,4
Oreamuno	7246	4,98	52	36161	34535	4,7
Heredia	23545	4,21	376	99463	87980	13,1
Barba	6742	4,47	32	30135	27848	8,2
Santo Domingo	7445	4,24	12	31611	31705	-0,3
Santa Bárbara	5842	4,48	33	26178	24980	4,8
San Rafael	8056	4,41	0	35484	33427	6,2
San Isidro	3653	4,40	0	16070	12640	27,1
Belén	4249	4,69	71	19998	18480	8,2
Flores	2760	4,43	35	12248	12420	-1,4
San Pablo	3911	4,41	49	17312	19807	-12,6

⁽¹⁾ Cantones que en 1984 poseían una cobertura eléctrica superior al 95%

7. MUESTREO DE ÁREAS

El presente capítulo expone un ejemplo de la aplicación del muestreo de áreas para la estimación de la población de un área pequeña. A través del mismo, se pretende analizar sus ventajas y desventajas.

Se decidió seleccionar un cantón que cumpliera ciertas características de urbanidad, migración y tamaño, de modo que sus resultados permitieran valorar su factibilidad. Después de analizar varias posibilidades se escogió el cantón de San Rafael de Heredia, considerando que cumplía los requisitos mínimos aceptables. En San Rafael, para 1984, el 68% de su población vivía en zona rural. Su tasa de emigración para el período 1979-1984 fue de 20% y la de inmigración de 28%, lo que indica un aceptable movimiento migratorio. Su densidad poblacional se ha estimado que cambiará de 455 h/km² en 1984 a 786 h/km² en el 2000, mientras que su tasa de crecimiento se ha proyectado para el período 1985-2000 en 2,9% (CELADE, MIDEPLAN y DGEC, 1990).

Para la aplicación de la técnica se utilizó el marco muestral de 84 segmentos censales que poseía San Rafael para 1984. Sin embargo, según información suministrada por personeros de la Municipalidad del cantón, en tres de estos segmentos se ha dado un crecimiento poblacional mucho más elevado que en los otros, debido a la creación de proyectos urbanísticos. Por tal razón, se tomó la decisión de excluirlos del marco muestral y considerarlos por separado. De no dárseles un trato independiente a estos segmentos, y de haber sido seleccionado al menos uno en la muestra, se presentaría una sobreestimación de la población. Por otro lado, de no seleccionar alguno, se obtendría entonces una subestimación. Con los 81 segmentos restantes se seleccionó una muestra simple al azar de 25, considerándose que con ellos se lograban los objetivos del estudio.

Cada segmento muestreado se recorrió completamente, actualizando el número de viviendas; en cada vivienda se preguntó por el número de personas que allí residía. Además, para cada segmento, en forma sistemática, se seleccionó una de cada cinco viviendas, con el fin de preguntar el sexo y la edad de sus residentes. Este mismo procedimiento se llevó a cabo en los tres segmentos que fueron separados del marco, con el fin de agregarlos a los resultados de la muestra.

El trabajo de campo se llevó a cabo entre el 18 de octubre y el 9 de noviembre de 1996. Se contó con solo un enumerador y el auxilio de un chofer ocasional. Para la muestra de 25 segmentos se invirtió un total aproximado de 96 horas de trabajo de enumeración, para un promedio de 3,8 horas por segmento. Para los restantes tres segmentos se duró en total 22 horas.

Los resultados se muestran en el Cuadro 7.1. Se muestra un fuerte crecimiento en casi todos los segmentos, con la excepción de algunos que, por sus características de urbanización, no crecieron poblacionalmente. Esto se debe a que algunas de sus viviendas fueron demolidas o transformadas en comercio o industria. También se corrobora el fuerte crecimiento que se dio en los tres segmentos tratados por separado, lo cual justifica la decisión tomada. Los 25 segmentos muestreados produjeron 2116 viviendas con un total de 9315 personas, para un promedio de 4,40 personas por vivienda. Del mismo modo, en los tres segmentos separados se contaron 901 viviendas con 3974 personas, para un promedio de 4,41 personas por vivienda.

Cuadro 7.1
Población total y número de viviendas por segmento censal muestreado
San Rafael de Heredia, 1996

Muestra	Viviendas		Población	
	Censo 1984	Conteo 1996	Censo 1984	Conteo 1996
1-03	61	85	303	426
1-05	68	104	319	451
1-07	53	55	226	242
1-09	49	53	201	200
1-13	69	89	301	433
1-16	65	82	314	387
1-18	48	43	242	163
1-21	86	137	375	560
2-01	69	76	326	302
2-05	60	78	290	319
2-09	70	64	355	311
2-19	64	79	298	313
2-20	71	92	358	453
3-01	60	70	261	269
3-05	65	92	343	374
3-06	65	85	274	338
3-07	65	138	207	667
3-10	55	100	262	486
4-01	61	101	259	489
4-05	32	77	129	361
4-11	70	117	320	463
4-16	34	52	138	223
5-01	75	112	351	479
5-03	35	62	119	259
5-04	55	73	241	347
Total	1505	2116	6812	9315
Segmentos separados				
1-25	85	199	381	849
2-15	54	319	271	1474
3-08	36	383	207	1651
Total	175	901	859	3974

A continuación se detallan los cálculos realizados. Sean:

P_c : Población según censo de 1984

P_e : Población estimada

V_c : Número de viviendas según censo de 1984

V_e : Número de viviendas estimado

r : Razón de crecimiento poblacional

r_1 : Razón de crecimiento en viviendas

y_i : Población actual del i -ésimo segmento

x_i : Población censal del i -ésimo segmento

a_i : Número de personas muestreadas en el i -ésimo segmento con característica de interés

Y_i : Número de personas muestreadas en el i -ésimo segmento

El estimador de razón para el crecimiento total de la población, de los 81 segmentos considerados en el marco, viene dado por:

$$r = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n x_i} = \frac{9315}{6812} = 1,3674$$

La aplicación de la razón de crecimiento estimada a la población censal de 1984 (21212), da la estimación poblacional siguiente para los 81 segmentos:

$$P_e = r \cdot P_c = \frac{9315}{6812} \cdot 21212 = 29006$$

Si se agrega la población de los tres segmentos separados, se obtiene que la estimación poblacional para San Rafael de Heredia a octubre de 1996 es:

$$P = 29006 + 3974 = 32980$$

Dicho valor presenta mucha semejanza con los obtenidos por medio del empleo de indicadores sintomáticos (ver Cuadro 6.9).

El coeficiente de variación para el promedio de la población censal de los 25 segmentos es 5,3%, mientras que para el promedio poblacional de la muestra es 7,3%. Ambos valores son menores al 10%, y si se supone que la muestra de 25 segmentos es grande, de manera que pueda considerarse el estimador poblacional utilizado como normalmente distribuido (Cochran, 1987), se tiene que la variancia estimada de P_e está dada por:

$$\text{Var}(P_e) = \frac{N^2(1-f)}{n(n-1)} \sum_{i=1}^n (y_i - rx_i)^2 = (81)^2 \left(1 - \frac{25}{81}\right) \frac{355998,31}{25 \cdot 24} = 2691347,22$$

Con error estándar igual a: $ee(P_e) = \sqrt{2691347,22} = 1641$

Los límites del intervalo de confianza del 90% para la estimación de la población de San Rafael de Heredia a octubre de 1996 son:

$$\text{Límite inferior: } P - t_{n-1, \alpha/2} \cdot ee(P_e) = 32980 - 1,71 \cdot 1641 = 30174$$

$$\text{Límite superior: } P + t_{n-1, \alpha/2} \cdot ee(P_e) = 32980 + 1,71 \cdot 1641 = 35786$$

Por otro lado, la estimación de la Razón de cambio en viviendas para los 25 segmentos muestreados es:

$$r_1 = \frac{\sum_{i=1}^n t_i}{\sum_{i=1}^n w_i} = \frac{2116}{1505} = 1,4060$$

Y así, la estimación del número de viviendas de los 81 segmentos correspondientes a la población considerada es:

$$V_e = r_1 \cdot V_c = \frac{2116}{1505} \cdot 4780 = 6721$$

Al agregar a este valor las 901 viviendas existentes en los tres segmentos separados se obtiene que, para octubre de 1986, el número total estimado de viviendas para San Rafael de Heredia es de:

$$V = 6721 + 901 = 7622$$

Dicho valor se asemeja al reportado por el ICE (7636) en cuanto a cantidad de medidores residenciales para este cantón a junio de 1996.

La estimación de la variancia de V_e es:

$$Var(V_e) = \frac{N^2(1-f)}{n(n-1)} \sum_{i=1}^n (t_i - r_1 w_i)^2 = (81)^2 \left(1 - \frac{25}{81}\right) \frac{8342,46}{25 \cdot 24} = 63069,0$$

y el correspondiente error estándar es: $ee(V_e) = \sqrt{63069,0} = 251$

De este modo, los límites del intervalo de confianza del 90% para la estimación de las viviendas de San Rafael de Heredia a octubre de 1996 son:

$$\text{Límite inferior: } V - t_{n-1, \alpha} \cdot ee(V_e) = 7622 - 1,71 \cdot 251 = 7193$$

$$\text{Límite superior: } V + t_{n-1, \alpha} \cdot ee(V_e) = 7622 + 1,71 \cdot 251 = 8051$$

Con respecto al estudio de las características sexo y edad de la población, se contó con una muestra de únicamente 380 viviendas con un total de 1706 personas. En este caso, el porcentaje de "no respuesta" fue aproximadamente un 10%, correspondiente a personas que se negaron a responder o no fue posible localizarlas en sus viviendas. Del total de personas muestreadas, 837 fueron hombres, por lo que la proporción estimada de varones para los 81 segmentos es:

$$\hat{p} = \frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n Y_i} = \frac{837}{1706} = 0,491$$

Por otro lado, para los tres segmentos considerados separadamente se observaron 157 viviendas con 796 personas, de los cuales 408 fueron hombres, para una proporción de varones estimada de:

$$\hat{p} = \frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n Y_i} = \frac{408}{796} = 0,513$$

De este modo la proporción de varones estimada para el cantón de San Rafael de Heredia es:

$$\frac{0,491 \cdot 25 + 0,513 \cdot 3}{28} = 0,493$$

El Cuadro 7.2 muestra un resumen de los resultados en los cálculos de las proporciones por edad y sexo para la población de San Rafael, donde la estrategia utilizada es similar a la empleada para estimar la proporción de varones descrita anteriormente.

Cuadro 7.2
Distribución proporcional de la población de San Rafael de Heredia
por sexo y grupos de edad

Grupos de edad	Muestra de 25 segmentos		Tres segmentos separados		Total
	$\sum_{i=1}^n a_i$	Proporción Estimada p_1	$\sum_{i=1}^n a_i$	Proporción Estimada p_2	Proporción Estimada $\frac{p_1 \cdot 25 + p_2 \cdot 3}{28}$
Hombres					
0 - 9	157	0,092	92	0,115	0,094
10 - 19	179	0,105	102	0,128	0,107
20 - 29	145	0,085	79	0,099	0,087
30 - 39	141	0,083	64	0,080	0,083
40 - 49	100	0,059	45	0,056	0,059
50 - 59	46	0,027	13	0,016	0,026
60 y más	69	0,040	13	0,016	0,037
Total	836	0,491	408	0,513	0,493
Mujeres					
0 - 9	146	0,086	98	0,123	0,090
10 - 19	203	0,119	93	0,117	0,119
20 - 29	132	0,077	57	0,072	0,076
30 - 39	153	0,090	83	0,104	0,092
40 - 49	84	0,049	19	0,024	0,046
50 - 59	62	0,036	19	0,024	0,035
60 y más	89	0,052	19	0,024	0,049
Total	870	0,509	389	0,487	0,507
Ambos sexos					
0 - 9	303	0,178	190	0,238	0,184
10 - 19	381	0,223	196	0,246	0,225
20 - 29	278	0,163	137	0,172	0,164
30 - 39	294	0,172	147	0,184	0,173
40 - 49	184	0,108	64	0,080	0,105
50 - 59	108	0,063	32	0,040	0,061
60 y más	158	0,093	32	0,040	0,087
Total	1706	1,000	797	1,000	1,000

CONCLUSIONES Y RECOMENDACIONES

8. CONCLUSIONES y RECOMENDACIONES

8.1 Conclusiones

8.1.1 Aplicación de los indicadores sintomáticos

Los resultados del estudio señalan que son pocos los indicadores sintomáticos disponibles en el medio. A pesar de ello, es posible obtener estimaciones poblacionales en áreas menores con una calidad aceptable, fundamentalmente a nivel cantonal y provincial.

El número de electores inscritos en el padrón electoral y los nacimientos son los indicadores sintomáticos que mostraron la más alta correlación con respecto a la población (cercaos al 0,90). Además, el crecimiento natural, que en realidad es una combinación lineal de los nacimientos y las defunciones, mostró buenos resultados. Estas variables ofrecieron las mejores estimaciones.

Por otra parte, el análisis mostró diferente precisión para las distintas áreas político-administrativas. A nivel provincial, se pueden obtener estimaciones confiables usando el padrón electoral con los métodos que emplean tasas de ocurrencia. Los resultados reflejaron errores porcentuales inferiores al 5% y con un error promedio cercano al 1%. A nivel cantonal, la eficiencia fue menor; sin embargo, se obtuvieron buenas estimaciones, especialmente con las técnicas que utilizan regresión lineal. Utilizando el padrón electoral y los nacimientos como variables predictoras, únicamente el 3% de los cantones presenta una desviación de su valor censal superior al 10%, con un error promedio cercano al 4%. Por otro lado, a nivel distrital, la carencia de información sintomática para el período 63-84, limitó el análisis. Únicamente se pudo contar con el crecimiento natural de población y con una aproximación del número de electores. Los resultados obtenidos señalan desviaciones importantes con respecto a su valor

censal y los mejores resultados se obtuvieron de los modelos de regresión lineal. Se encontró que el 45% de los distritos mostró errores superiores al 10%, de los cuales, aproximadamente un 15% presentó errores superiores al 20%, con un error promedio cercano al 12%.

Una consideración importante con respecto a los resultados obtenidos es que las estimaciones se produjeron 11 años después del censo. Para períodos poscensales más cortos se esperarían estimaciones más exactas, debido a que las proporciones y las tasas de los indicadores sintomáticos no habrían presentado cambios tan fuertes.

Con respecto al rendimiento de las técnicas de estimación, los métodos que emplean las tasas de ocurrencia (Razón Censal y Diferencia de Tasas) proporcionaron las mejores estimaciones a nivel provincial. Esto favorece futuras estimaciones, pues si se aplican los métodos de regresión lineal a este nivel traería problemas por la cantidad reducida de provincias. Los procedimientos que utilizan regresión lineal (Razón Correlación y Tasa Correlación) producen estimaciones de mejor calidad a nivel cantonal y distrital. Sin embargo, en términos generales, las diferencias entre los cuatro métodos no son muy marcadas, por lo que se puede utilizar cualquiera de ellos sin que se tengan grandes diferencias en la calidad de las estimaciones. Esto tiene la ventaja de no limitar futuras estimaciones a una sola técnica, lo cual permite llevar a cabo una comparación de los resultados de los cuatro métodos, tal y como se hizo en el caso de provincias y cantones para 1996.

Los cálculos realizados no mostraron diferencias sustanciales en los métodos Razón Correlación y Tasa Correlación. Sin embargo, la verificación del cumplimiento de los supuestos del modelo de regresión por mínimos cuadrados, muestra que el segundo método presenta una mayor consistencia. En el análisis de los residuos realizado a nivel cantonal y distrital, el método de Razón

Correlación presentó problemas con el supuesto de homoscedasticidad, fundamentalmente a nivel distrital. Esta situación debe alertar a quienes decidan trabajar solo con esta técnica, pues podría conducir a errores estadísticos que redundarían en las estimaciones obtenidas.

Otro de los objetivos del estudio estaba dirigido a proponer un mecanismo para obtener proyecciones poblacionales a corto plazo. Los resultados indican que, con la utilización del padrón electoral y de los nacimientos, se pueden obtener proyecciones de una calidad aceptable. La experiencia mostró que es factible efectuar dichas proyecciones para los cinco años posteriores a la obtención de los indicadores sintomáticos. Como tal práctica se puede efectuar año a año, se hace posible contar con un sistema dinámico que permite actualizar las proyecciones cada nuevo año. Esto mejoraría notablemente la estrategia rígida y tradicional, que consiste en efectuar proyecciones después de un censo de población y mantenerlas fijas hasta el próximo.

Se debe resaltar que los modelos de regresión utilizados para las estimaciones y proyecciones poblacionales para el período 96-2001, fundamentalmente los referentes a la técnica Tasa Correlación, mostraron resultados de mejor calidad a la obtenida para 1984, en cuanto al coeficiente de determinación (R^2), la significancia de los coeficientes del modelo y los coeficientes de correlación con las distintas variables sintomáticas. Por este motivo, se podría suponer que dichas estimaciones serían de mayor precisión; sin embargo, esto es hipotético hasta que no se efectúe una comparación con datos censales.

Intentos por mejorar las estimaciones mediante el agrupamiento de cantones con base en ciertas características socio-demográficas, no reflejaron resultados positivos. Las estimaciones realizadas con base en cada uno de los grupos formados no fueron superiores en precisión, si se compara con las obtenidas en forma global con todas la áreas.

En la aplicación del método de viviendas no se contó con la información sintomática necesaria para realizar una prueba de exactitud similar a la empleada por los otros cuatro métodos. No obstante, los cálculos para algunos cantones en el año 1996, con los abonados del servicio eléctrico, reflejaron mucha semejanza con la alcanzada por los otros métodos para ese mismo año. Este hecho sugiere que el procedimiento puede dar valiosos aportes en el campo de las estimaciones poblacionales, tal como lo sugieren los antecedentes en países desarrollados.

Un elemento importante que los resultados del estudio dejan latente, y que debe ser discutido con detalle, lo constituye la distribución de los errores en las estimaciones de acuerdo con el tamaño poblacional. Los modelos de estimación evidencian que conforme los tamaños poblacionales van siendo más pequeños, la precisión de las estimaciones descende. Este hecho se observa claramente en la disminución que sufre el promedio de los AEP cuando se pondera por el tamaño de la población de cada área menor, así como en las diferencias, ya comentadas, entre provincias, cantones y distritos.

Este fenómeno puede ser ocasionado por muchas causas. Un primer aspecto que puede influir en que se presente esta situación lo constituye la calidad de la información sintomática, pues se hace más deficiente a niveles desagregados menores (Rincón, 1989). Otro elemento que puede afectar radica en que los errores dependen de la calidad de los conteos censales. Sin embargo, la información censal está sujeta a problemas de cobertura, los que difieren en magnitud de una área menor a otra. Lógicamente, es de esperar que esas diferencias en magnitudes sean más notorias a niveles político-administrativos más desagregados y en poblaciones más pequeñas (Wolter y Causey, 1991). Una tercera causa que podría justificar también las diferencias en el comportamiento de los errores, para diferentes tipos de áreas, que constituyó un serio problema en el manejo de información a nivel cantonal y distrital, consiste en la variación de los límites geográficos, por la creación constante de nuevas

áreas, sobre todo a nivel distrital. Esta inestabilidad geográfica obliga a llevar a cabo reagrupaciones o desagregaciones en las localidades que han sido afectadas, lo que puede provocar errores en la manipulación de los datos poblacionales y de la información sintomática.

Otro hecho que llamó la atención en el comportamiento de las estimaciones es que parece presentarse un sesgo importante en el comportamiento de los errores, ya que la mayoría de modelos presentó errores porcentuales (EP) negativos, lo cual indica que para el período en consideración existieron más subestimaciones al valor censal que sobreestimaciones. Puesto que las estimaciones fueron prorrateadas para ajustar la población total del país, la suma de los EP es cero. Por esta razón, el que se presentaran más valores de EP negativos que positivos señala que la magnitud de esos valores positivos fue superior, es decir, aunque el número de áreas sobreestimadas fue menor, los errores en sus estimaciones con respecto al valor censal fueron de mayor magnitud.

Finalmente, una de las metas implícitas del estudio consistía en desarrollar una metodología que produjera estimaciones poblacionales de mejor calidad que la obtenida suponiendo un crecimiento natural de población. Al comparar las estimaciones obtenidas por los cuatro métodos con las publicadas por la DGEC para 1984, se observaron grandes diferencias en la exactitud. Se observó que a nivel provincial el valor promedio de AEP fue de 7,1%, el cual supera en casi 6 puntos porcentuales el obtenido en el estudio. Para los cantones, este promedio fue de 12,4%, superando en alrededor de 8,5 puntos porcentuales la precisión de las estimaciones obtenidas en el estudio. Por último, para los distritos el promedio de AEP fue de 20,3%, el cual supera en aproximadamente un 9 puntos porcentuales las estimaciones hechas con indicadores sintomáticos. Estos resultados muestran que, al menos en períodos alejados del año censal, las estimaciones aquí obtenidas mejoran enormemente la calidad de aquellas que se trabajan con el supuesto de crecimiento natural.

8.1.2 Muestreo de áreas

El estudio realizado pretendió determinar la posibilidad de usar el muestro de áreas para la obtención de estimaciones poblacionales en áreas menores. El experimento mostró que la técnica es una buena alternativa para la obtención no solo de la información sobre población total, sino de otras características de las áreas, como número de viviendas, promedio de personas por vivienda, razón de crecimiento poblacional y de viviendas para el período de interés, estructura por edad y sexo, y otras que podrían considerarse. Este procedimiento permite determinar errores muestrales, por lo que es posible efectuar inferencias estadísticas utilizando las estimaciones obtenidas. Otra de las ventajas es que permite darle un trato separado a aquellos segmentos de mayor crecimiento, lo que da mayor precisión a las estimaciones obtenidas.

Los principales problemas que acarrea la aplicación de esta técnica se orientan en tres direcciones. Primero, en el proceso de muestreo se debe dar un trato aparte a aquellos segmentos que han presentado un crecimiento desproporcionado con respecto a los demás. Segundo, debe considerarse la disponibilidad de tiempo, pues el proceso requiere varias etapas para su elaboración, que hace lenta la obtención de resultados. Finalmente, deben analizarse los costos que cada una de las etapas exige para su conclusión.

Debido a las consideraciones anteriores no se justificaría su aplicación si lo que se busca es únicamente estimar el tamaño poblacional de una área administrativa, ya que se podrían obtener estimaciones de una calidad similar por medio de los indicadores sintomáticos.

La estrategia utilizada cobra especial atractivo ante la problemática del país en materia de población, donde casi trece años después del último censo se desconoce la fecha de realización del próximo. Este método podría ser aplicado

para la determinación de las características poblacionales en aquellas áreas que sean de interés para los investigadores.

8.2 Recomendaciones

Debido a la carencia de información censal, la prueba sobre la precisión de las estimaciones poblacionales tuvo que realizarse para 1984. Por esta razón, la primera recomendación va dirigida a la necesidad de repetir el estudio tan pronto como se realice un censo de población. De este modo, se podría observar la estabilidad de los indicadores sintomáticos y de los métodos de estimación.

También es importante analizar el comportamiento de otros posibles indicadores que pudieran dar su aporte al proceso de estimación. Se debe analizar más a fondo el comportamiento de los abonados del servicio eléctrico, considerándolo tanto como un indicador sintomático simple o como un estimador del número de viviendas de una localidad.

La más importante recomendación se dirige a la DGEC, como ente encargado de producir estimaciones poblacionales, consiste en revisar lo hecho hasta ahora en materia de estimaciones poblacionales en áreas pequeñas, pues la investigación demostró la existencia de fuentes importantes de datos y métodos que pueden mejorar notablemente las estimaciones, especialmente para períodos alejados de los años censales. Es necesario que el país cuente con un sistema de estimaciones y proyecciones poblacionales a nivel de áreas administrativas menores. De manera que cualquier persona interesada en conocer el tamaño poblacional de una localidad, tenga una fuente confiable a la cual recurrir. En este sentido, se puede utilizar como punto de partida la utilización de los indicadores sintomáticos con las técnicas desarrolladas en el presente estudio.

Si la DGEC, o cualquier otra institución estatal o privada, se encargara de reunir anualmente la información, simplificaría la tarea de recolectar los datos de los distintos indicadores sintomáticos cada vez que se realice una estimación poblacional para pequeñas localidades. Además debe existir una coordinación constante con las distintas instituciones que proveen la información sintomática, a fin de mejorar su calidad y llevarlas, si es posible, a niveles más desagregados. De esta forma, se dispondría de mejor y más completa información, que permitiría obtener estimaciones poblacionales más eficientes.

Por otro lado, la utilización del muestreo de áreas para obtener estimaciones de población ofrece una oportunidad de medir cambios demográficos a través de la observación directa. Esta técnica se sugiere como una alternativa ante la carencia de un censo poblacional y de vivienda. Sin embargo, debe considerarse la inversión de tiempo y costo en que se incurre. Por este motivo, la persona o institución que requiere de información poblacional, es en última instancia, quien debe decidir si los resultados que puede generar un estudio de este tipo y su utilidad, compensan el esfuerzo de llevarlo a cabo.

Debe quedar claro, con base en el presente estudio, que la exactitud de un método no puede ser generalizada, y que por más robusto que pueda ser un determinado procedimiento de estimación, no se puede asegurar que su aplicación a diferentes situaciones conduzca a resultados que reflejen la realidad. Al analizar las estimaciones deben considerarse las hipótesis correspondientes a los métodos y la calidad de la información que se maneja. Es importante entonces, tener un juicio crítico y no caer en el uso mecánico de los métodos desarrollados, ni de los resultados obtenidos.

BIBLIOGRAFÍA

9. BIBLIOGRAFÍA:

- Bogue, Donald J. (1950). **A Technique for Making Extensive Population Estimates.** Journal of the American Association, 45, 149-163.
- CELADE (1984). **Métodos para proyecciones demográficas.** CELADE- Agencia Canadiense para el Desarrollo Internacional. San José, Costa Rica.
- CELADE (1989). **Métodos para proyecciones subnacionales de población.** CELADE-DANE. San José, Costa Rica.
- CELADE (1990). **Estimaciones y proyecciones de población cantonal por sexo y grupos de edades: 1975-2000.** MIDEPLAN, CELADE y DGEC. San José, Costa Rica.
- CEPAL(1996). **Plan de Acción Regional Latinoamericano y del Caribe sobre Población y Desarrollo.** CEPAL, Naciones Unidas. Santiago de Chile.
- Cochran, Willim G. (1987). **Técnicas de muestreo.** CECSA. México D F.
- DGEC. **Población de la República de Costa Rica, por provincias, cantones y distritos.** DGEC. San José, Costa Rica.(Serie periódica de publicaciones).
- DGEC. **Censos de Población y Vivienda de 1963, 1973, 1984.** San José, Costa Rica, 1964, 1974 y 1987
- DGEC. **Anuarios Estadísticos de Costa Rica.** San José, Costa Rica.
- Dillon, William y Goldstein, Matthew (1984). **Multivariate Analysis. Methods and Applications.** John Wiley & Sons. New York.
- Duchesne, Louis (1989). **Proyecciones de población, por sexo y edad, para áreas intermedias y menores Método de "relación de cohortes".** Documento presentado en el Seminario Internacional sobre Proyecciones Subnacionales de Población en Giraldot, Colombia. CELADE. San José, Costa Rica.
- Duchesne, Louis (1991). **Método de proyecciones de población por componentes multirregionales.** CELADE. Santiago de Chile.
- Ericksen, Eugene P (1973). **A method for combining sample survey data and symptomatic indicators to obtain population estimates for local areas.** Demography, 10, 137-160.

- Ericksen, Eugene P (1974). **A regression method for estimating population changes of local areas.** *Journal of the American Statistical Association*, 69, 867-875.
- Feeney, Donald; Hibbs, James y Gillaspay, Thomas (1995). **Ratio-Correlation Method. Presentado en Rives, Norfleet y otros Basic Methods for Preparing Small-Area Population Estimate** (pag. 118-136). *Applied Population Laboratory, Department of Rural Sociology, College of Agricultural and Life Sciences. University of Wisconsin-Madison/Extension.*
- Flores, Carmen Elisa (1989). **Las proyecciones de población y la estructura económica y social.** Documento presentado en el Seminario Internacional sobre Proyecciones Subnacionales de Población en Giraldot, Colombia. CELADE. San José, Costa Rica.
- FNUAP-MIDEPLAN (1995). **Conferencia Internacional sobre Población y el Desarrollo (El Cairo, Egipto, 1994).** Selección de documentos relativos a la Posición del Gobierno Costarricense. FNUAP-MIDEPLAN. San José, Costa Rica.
- Freund, John E. y Walpole, Ronald E. (1990). **Estadística Matemática con Aplicaciones.** Cuarta Edición. Prentice-Hall Hispanoamericana, S. A. México.
- García, Agustín y Rincón, Manuel (1989). **Sistema para elaborar proyecciones subnacionales, de áreas intermedias y pequeñas, por sexo y grupos de edad.** Documento presentado en el Seminario Internacional sobre Proyecciones Subnacionales de Población en Giraldot, Colombia. CELADE. San José, Costa Rica.
- Giraldelli, Bernardette (1989). **Parámetros demográficos proporcionales: Una alternativa para aplicar el "método de componentes", para proyectar la población en áreas pequeñas.** Documento presentado en el Seminario Internacional sobre Proyecciones Subnacionales de Población en Giraldot, Colombia. CELADE. San José, Costa Rica.
- Gómez, Miguel (1985). **Elementos de Estadística Descriptiva.** EUNED. San José, Costa Rica.
- Granados, María del Pilar (1988). **Técnicas de proyecciones de población de áreas menores. Aplicación y evaluación.** CELADE-CEPAL. Santiago de Chile.

- Granados, María del Pilar (1989). **Técnicas de proyección de población en áreas menores. Aplicación y evaluación.** Documento presentado en el Seminario Internacional sobre Proyecciones Subnacionales de Población en Giraldot, Colombia. CELADE. San José, Costa Rica.
- Hernández, Oscar y Bolaños, Ramón Luis (1993). **SPSS/PC+ Básico.** EUCR. San José, Costa Rica.
- ICE. **Grado de electrificación de Costa Rica.** San José, Costa Rica
- ICE. **Estadísticas eléctricas de Costa Rica.** San José, Costa Rica
- Johnson, Richard y Wichern, Dean (1988). **Applied Multivariate Statistical Analysis.** Segunda edición. Prentice Hall, New Jersey.
- Lee, Everetts S y Goldsmith, Harold F (1982). **Population Estimates: methods for small area analysis.** SAGE Publications. Beverly Hills, California.
- Malec, D y Katzoff, J (1983). **Application of Some Common Spatial Models to Population Data.** Documento presentado en la Reunión Anual de la Asociación Estadísticos Americanos, Toronto, Ontario.
- Mandell, Marylou y Tayman, Jeffrey (1982). **Measuring Temporal Stability in Regression Models of Population Estimation.** *Demography*, 19, 135-146.
- Martin, Julia H (1995). **Basic Methods for Preparing Small-Area Population Estimate** (pag. 142-147). Applied Population Laboratory, Department of Rural Sociology, College of Agricultural and Life Sciences. University of Wisconsin-Madison/Extension.
- MEP. **Estadísticas sobre matrícula inicial en primer y segundo ciclo de la Enseñanza General Básica.** Departamento de Estadística del MEP. San José, Costa Rica.
- Murdock, Steve; Huang, Sean-Shong y Hamm, Rita (1995). **Component Method. En Rives, Norfleet y otros Basic Methods for Preparing Small-Area Population Estimate** (pag.10-53). Applied Population Laboratory, Department of Rural Sociology, College of Agricultural and Life Sciences. University of Wisconsin-Madison/Extension.
- Naciones Unidas (1975). **Métodos para hacer proyecciones de la población urbana y rural.** Manual VIII de Naciones Unidas. Nueva York.

- Namboodiri, Krishnan (1972). **On the Ratio-Correlation and Related Method of Subnational Population Estimation.** *Demography*, 9, 443-453.
- O'Hare, W (1980). **A Note on the Use of Regression Methods in Population Estimates.** *Demography*, 17, 341-343.
- Perry, Marc y Voss, Paul (1995). **Using Geo-Demographic for Improving Small-Area Population Estimates.** Applied Population Laboratory, Department of Rural Sociology. University of Wisconsin-Madison.
- Raabe, Carlos y Umaña, Rodrigo (1971). **Estimaciones de las poblaciones de los cantones de Costa Rica: 1970-1980.** Instituto de Investigaciones de la Escuela de Ciencias Económicas, Ciudad Universitaria Rodrigo Facio, Costa Rica.
- Rincón, Manuel (1986). **Métodos Cuantitativos y Modelos de Población y Desarrollo: "Elementos de las Proyecciones de Población útiles para la Planificación Socioeconómica".** CELADE San José, Costa Rica.
- Rincón, Manuel (1989). **Métodos y programas de proyecciones de población.** CELADE. San José, Costa Rica.
- Rincón, Manuel (1990). **Teoría y métodos para la preparación de estimaciones y proyecciones de población.** CELADE. San José, Costa Rica.
- Rincón, Manuel y Hernández, Harry (1989). **Método para elaborar proyecciones de población de áreas pequeñas, por sexo y grupos de edad.** CELADE. San José, Costa Rica.
- Rives, Norfleet W (1982). **Assessment of a Survey Approach.** En Lee, Everetts S y Goldsmith, Harold F. *Population Estimates: methods for small area analysis* (pag. 79-96). SAGE Publications. Beverly Hills, California.
- Serow, William y Rives, Norfleet (1995). **Small Area Analysis: Assessing the State of the Art. Presentado en Rives, Norfleet y otros Basic Methods for Preparing Small - Area Population Estimates** (pág. 1-9). Applied Population Laboratory, Department of Rural Sociology, College of Agricultural and Life Sciences. University of Wisconsin-Madison/Extension.
- Shapiro, S y Francia, R. S. (1972). **An Aproximate Analysis of Variance Test for Normality.** *Journal of the American Statistical Association* 67, 216-216
- Shryock, Henry S. y Siegel, Jacob S. (1976). **The Methods and Materials of Demography.** Academic Press. New York.

- Simpson, Stephen y otros (1996). **Updating Small Area Population Estimates in England and Wales**. *Journal of the American Statistical Association*, 159, 235-247
- Smith, Stanley (1995). **Housing Unit Method. Presentado en Rives, Norfleet y otros Basic Methods for Preparing Small-Area Population Estimate** (pág. 89-117). Applied Population Laboratory, Department of Rural Sociology, College of Agricultural and Life Sciences. University of Wisconsin- Madison/Extension.
- Smith, Stanley y Lewis, Bart (1980). **Some New Techniques for Applying the Housing Unit Method of Local Population Estimation**. *Demography*, 17, 323-339.
- Smith, Stanley y Lewis, Bart (1983). **Some New Techniques for Applying the Housing Unit Method of Local Population Estimation: Further Evidence**. *Demography*, 20, 407-413.
- Smith, Stanley y Mandell, Marylou (1984). **A Comparison of Population Estimation Methods: Housing Unit Versus Component II, Ratio-Correlation and Administrative Records**. *Journal of the American Statistical Association*, 79, 282-289.
- Stata Press (1995). **Reference Manual Stata**. College Station, Texas
- Starsinic, D y Zitter, M (1968). **Accuracy of the Housing Unit Method in Preparing Population Estimates for Cities**. *Demography*, 5, 475-484.
- Swanson, David (1980). **Improving Accuracy in Multiple Regression Estimates of Population Using Principles From Causal Modelling**. *Demography*, 17, 413-427.
- Swanson, David y Tedrow, Lucky (1984). **Improving the Measurement of Temporal Change in Regression Models Used for County Population Estimates**. *Demography*, 21, 373-381.
- Tribunal Supremo de Elecciones. **Boletín**. Departamento de Estudios y Servicios Técnicos del TSE. San José, Costa Rica
- Tribunal Supremo de Elecciones. **Computo de votos y declaratorias de elección**. Departamento de Estudios y Servicios del TSE. San José, Costa Rica.

Voss, Paul R. y otros(1995). Censal Ratio Methods. En Rives, Norfleet y otros Basic Methods for Preparing Small-Area Population Estimate (pag. 71-88). Applied Population Laboratory, Department of Rural Sociology, College Agricultural and Life Sciences. University of Wisconsin-Madison/Extension.

Wolter, Kirk y Causey, Beverley. Evaluation of Procedures For Improving Population Estimates for Small Areas. Journal of the American Statistical Association, 86, 278-284.

ANEXOS

ANEXO 1

Cuadro A 1.1
Población total e información correspondiente a ocho indicadores sintomáticos
para 22 cantones de San José y Heredia, años 1973 y 1984

Cantón	Nacim.		Crecim. Natural		Matricula Escolar		Padrón Electoral		Defunc.		Aseg. Enf. y Mat. CCSS		Permisos Constr.		Asistencia P. Salud		Población	
	1973	1984	1973	1984	1973	1984	1973	1984	1973	1984	1973	1984	1973	1984	1973	1984	1973	1984
San José	6002	7628	4758	6338	40854	35098	124926	181522	1244	1290	122431	212419	1033	1410	38130	44790	215441	241464
Escazú	730	949	601	808	4204	4088	10494	19753	129	141	2006	3325	146	173	9308	9020	25026	33101
Desamparados	2021	3313	1673	2926	15469	13436	30585	66161	348	387	4109	10134	401	608	12052	11010	74272	108824
Aserri	597	938	488	833	3878	4245	8070	16066	109	105	1330	2359	35	146	9019	8865	20091	30588
Mora	275	341	215	289	2293	1938	4632	7393	60	52	248	701	35	83	3977	7500	10733	12584
Goicoechea	1635	2337	1319	1997	8089	7483	31057	49553	316	340	7792	16941	272	420	13419	14265	61607	79931
Santa Ana	398	557	314	473	3051	2340	5884	11656	84	84	778	1368	111	110	6504	6500	14499	19605
Alajuelita	630	1062	530	947	4406	5155	9398	19702	100	115	524	972	89	105	9593	9105	23013	31390
Coronado	429	734	330	640	2723	2900	6920	13375	99	94	1177	2074	127	272	3363	6000	16336	24514
Tibás	1027	1802	832	1575	5396	5429	19852	38811	195	227	4460	6798	272	506	7072	8190	35602	57693
Moravia	476	822	396	722	4401	5374	8481	18590	80	100	2042	3450	295	270	5111	6460	19548	33038
Montes de Oca	962	1219	803	1028	4780	4608	17263	27011	159	191	7709	14170	86	127	4554	9425	33633	39065
Curridabat	340	850	275	746	2450	2540	6021	15704	65	104	3011	5487	144	272	3621	6410	15591	31954
Heredia	1420	2245	1137	1963	10209	11298	22538	42144	283	282	10445	19951	104	297	4792	6185	49105	73805
Barba	392	619	331	546	2478	2430	5567	10251	61	73	708	1473	49	172	4039	5085	12864	18933
Santo Domingo	440	662	353	555	3199	3410	8145	14948	87	107	1401	3109	101	134	3696	5595	17423	23985
Santa Bárbara	275	467	220	403	1903	2333	4851	8976	55	64	678	997	88	133	1551	2625	10738	16643
San Rafael	339	600	274	537	2397	2518	6380	12993	65	63	732	1672	110	95	3331	4405	16013	22871
San Isidro	131	239	100	202	1038	1192	2734	5149	31	37	326	468	71	88	2289	3455	5979	8528
Belén	210	310	175	262	1609	1431	3859	6922	35	48	2561	5793	75	175	3229	4210	8538	11993
Flores	168	273	133	237	1267	1403	3194	5983	35	36	885	1811	90	52	3131	795	6524	9015
San Pablo	145	244	118	204	931	1010	2646	6496	27	40	395	714	57	162	170	1200	6660	11802

Fuentes: DGEC, Departamento de Estadística del MEP, Sección de Archivo del TSE, Departamento de Estadística del Ministerio de Salud

Cuadro A 1.2
Población total e información sobre cinco indicadores sintomáticos por cantón
años 1963, 1973 y 1984

Cantón	Población			Electores			Nacimientos			Crecimiento Nat.			Matrícula Escolar			Defunciones		
	1963	1973	1984	1963	1973	1984	1963	1973	1984	1963	1973	1984	1963	1973	1984	1963	1973	1984
San José	168938	215441	241464	86694	124926	181522	8422	6002	7628	7143	4758	6338	30378	40854	35098	1279	1244	1290
Escazú	14250	25026	33101	5537	10494	19753	664	730	949	550	601	808	2303	4204	4088	114	129	141
Desamparados	33845	74272	108824	12116	30585	66161	1723	2021	3313	1482	1673	2926	6255	15469	13436	241	348	387
Puriscal	23690	24150	23123	8183	11611	15416	1103	697	677	945	570	569	4024	5454	3481	158	127	108
Tarrazú	5392	7542	8845	2249	3281	4934	288	200	269	239	173	241	1017	1707	1146	49	27	28
Aserri	13609	20091	30588	4805	8070	16066	696	597	938	582	488	833	2607	3878	4245	114	109	105
Mora	8938	10733	12584	3788	4632	7393	407	275	341	323	215	289	1648	2293	1938	84	60	52
Goicoechea	44110	61607	79931	16462	31057	49553	1947	1635	2337	1652	1319	1997	5919	8089	7483	295	316	340
Santa Ana	9026	14499	19605	3554	5884	11656	434	398	557	382	314	473	1659	3051	2340	52	84	84
Alajuelita	10848	23013	31390	3431	9398	19702	499	630	1062	413	530	947	2084	4406	5155	86	100	115
V. de Coronado	10615	16336	24514	3932	6920	1337	439	429	734	353	330	640	1699	2723	2900	86	99	94
Acosta	13092	14385	14853	4761	6102	8302	731	473	442	602	385	387	2257	3036	2343	129	88	55
Tibás	23946	35602	57693	10687	19852	38811	1162	1027	1802	1019	832	1575	4099	5396	5429	143	195	227
Moravia	11642	19548	33038	4528	8481	18590	485	476	822	411	396	722	2672	4401	5374	74	80	100
Mont. de Oca	24043	33633	39065	8648	17263	27011	978	962	1219	880	803	1028	4000	4780	4608	98	159	191
Turrubares	5496	4709	4471	1890	2138	2448	357	190	125	286	163	109	899	1053	879	71	27	16
Dota	3718	4375	4934	1352	2032	2975	138	87	116	112	71	99	697	908	614	26	16	17
Curridabat	9491	15591	31954	3124	6021	15704	346	340	850	284	275	746	1213	2450	2540	62	65	104
Peréz Zeledón	47319	67089	82370	14886	26556	43494	2453	2338	3036	2035	2018	2771	9829	15059	13325	418	320	265
León Cortes	5650	7521	8087	2023	3191	4534	305	201	229	253	163	202	1146	1830	1069	52	38	27
Alajuela	64398	96325	127472	26556	43262	78674	2998	2601	3544	2434	2021	2980	10850	17301	15841	564	580	564
San Ramón	25925	33155	39963	10848	15308	24468	1255	894	1134	1070	724	985	4964	6712	5362	185	170	149
Grecia	43923	58086	82600	11157	20557	38664	1963	2055	2782	1632	1693	2477	6984	12854	14953	331	362	305
San Mateo	3388	2969	3783	1472	1641	2192	140	41	73	123	30	62	683	587	482	17	11	11
Atenas	11018	12610	15011	4948	6621	9250	489	290	360	413	228	302	1742	2401	1832	76	62	58
Naranjo	16414	19721	23588	6857	9474	14767	779	500	693	654	401	597	3165	4298	3116	125	99	96

Cuadro A 1.2
Continuación

Cantón	Población			Electores			Nacimientos			Crecimiento Nat.			Matrícula Escolar			Defunciones		
	1963	1973	1984	1963	1973	1984	1963	1973	1984	1963	1973	1984	1963	1973	1984	1963	1973	1984
Palmares	12283	14495	17815	5026	7334	11720	635	293	505	520	227	432	2393	3387	1986	115	66	73
Poás	8179	10191	13939	3211	4795	7663	353	245	381	302	198	325	1382	1756	1577	51	47	56
Orotina	7093	8479	10494	3363	4232	6067	347	165	259	285	115	203	1444	1866	1417	62	50	56
San Carlos	36586	54952	75576	11898	21539	38066	2113	1842	2653	1805	1571	2433	6833	12575	11612	308	271	220
Alfaro Ruiz	4919	6342	7005	2045	2998	4242	205	151	197	172	126	169	905	1267	905	33	25	28
Valverde Vega	6546	8707	10716	2652	4101	6458	308	220	318	251	176	283	1162	1967	1260	57	44	35
Cartago	46722	65310	87125	18785	28992	50651	2258	1856	2534	1862	1486	2182	8990	13191	12204	396	370	352
Paraíso	18389	22281	27823	5858	9204	15696	837	672	825	657	558	714	3407	5856	4138	180	114	111
La Unión	14074	23352	41005	5205	8910	20628	604	620	1182	514	524	1036	2499	4490	5243	90	96	146
Jiménez	10439	11523	11861	4177	4965	6685	516	395	338	396	313	290	1939	2560	1979	120	82	48
Turrialba	37620	43202	50567	15060	20021	29721	2081	1364	1440	1626	1095	1238	6729	9523	7002	455	269	202
Alvarado	6465	7484	8338	2157	2892	4255	259	201	213	205	164	180	1088	1528	1117	54	37	33
Oreamuno	12035	17517	24145	4127	7215	13432	613	588	730	476	468	634	1749	3319	3216	137	120	96
El Guarco	9689	14030	20807	3571	5699	10522	550	458	649	439	371	580	1952	3262	3249	111	87	69
Heredia	30919	49105	73805	12796	22538	42144	1502	1420	2245	1270	1137	1963	5705	10209	11298	232	283	282
Barba	8504	12864	18933	3401	5567	10251	371	392	619	323	331	546	1452	2478	2430	48	61	73
Santo Domingo	11348	17423	23985	5059	8145	14948	417	440	662	349	353	555	2158	3199	3410	68	87	107
Santa Bárbara	8127	10738	16643	3096	4851	8976	328	275	467	273	220	403	1626	1903	2333	55	55	64
San Rafael	9069	16013	22871	3283	6380	12993	355	339	600	301	274	537	1597	2397	2518	54	65	63
San Isidro	4061	5979	8528	1917	2734	5149	177	131	239	146	100	202	845	1038	1192	31	31	37
Belén	4781	8538	11993	2387	3859	6922	193	210	310	156	175	262	975	1609	1431	37	35	48
Flores	4162	6524	9015	2161	3194	5983	170	168	273	140	133	237	901	1267	1403	30	35	36
San Pablo	4092	6660	11802	1426	2646	6496	168	145	244	145	118	204	635	931	1010	23	27	40
Liberia	18030	30114	38943	5714	10748	20210	1157	1223	1616	1021	1031	1461	3667	6087	7386	136	192	155
Nicoya	36276	45084	42505	11635	19200	24977	2103	1332	1211	1811	1085	1053	7270	10371	7107	292	247	158
Santa Cruz	23576	29739	31133	8633	12633	18212	1117	840	836	962	696	696	5957	6643	4936	155	144	140
Bagaces	9836	9828	10103	2586	4045	5535	633	320	304	540	263	267	1717	1992	1512	93	57	37
Carrillo	11396	14893	18475	3819	6337	9825	572	421	546	476	336	462	2678	3490	2702	96	85	84
Cañas	9117	12779	17284	3358	5473	9543	525	453	598	431	372	530	1667	2680	2676	94	81	68
Abangares	10189	11633	12575	3165	4713	6717	550	350	383	469	298	335	1903	2385	1869	81	52	48

Cuadro A 1.2
Continuación

Cantón	Población			Electores			Nacimientos			Crecimiento Nat.			Matrícula Escolar			Defunciones		
	1963	1973	1984	1963	1973	1984	1963	1973	1984	1963	1973	1984	1963	1973	1984	1963	1973	1984
Tilarán	12097	12563	14586	4356	5851	8558	769	412	500	660	356	450	2240	3132	2173	109	56	50
Nandayure	12038	12058	9604	3471	4938	5479	607	461	266	518	380	233	2562	3133	1719	89	81	33
Puntarenas	55592	65562	77279	19884	29495	42842	3012	2204	2398	2412	1724	2099	10991	14663	12384	600	480	299
Esparza	9175	12095	14998	3502	5250	8329	380	307	397	304	239	342	1825	2671	2024	76	68	55
Buenos Aires	11042	20104	27716	2151	5699	11735	485	794	1116	400	676	1036	1982	4510	5348	85	118	80
Motes de Oro	6616	6979	7444	2508	3128	4381	367	220	205	321	176	170	1402	1811	1109	46	44	35
Osa	17574	24613	26294	5693	8913	13514	1016	855	862	871	705	783	3156	5265	4627	145	150	79
Aguirre	19942	26374	23093	6660	10406	13250	1119	945	771	915	762	691	3830	5935	4207	204	183	80
Golfito	36567	62481	89059	10230	20748	43186	1656	2291	3396	1402	1961	3115	5494	14612	16314	254	330	281
Limón	40141	56750	78338	13527	24279	43538	1802	2071	3022	1281	1572	2612	7339	11714	14189	521	499	410
Pococi	16927	40260	60659	4874	13305	28619	815	1538	2191	637	1289	1993	3105	8155	10579	178	249	198
Siquirres	11317	18133	29079	3393	6380	12656	568	550	896	452	423	796	2132	3371	4987	116	127	100
Total	1336274	1871780	2416809	506308	829709	1406189	65814	56266	75429	54768	45751	65835	244076	379392	350855	11046	10515	9594

Fuentes: DGEC, Departamento de Estadística del MEP, Sección de Archivo del TSE

Cuadro A 1.3
Población total e información sobre cinco indicadores sintomáticos
por provincia, años 1963, 1973 y 1984

Provincia	Población			Electores			Nacimientos			Crecimiento Nat.			Matrícula Escolar			Defunciones		
	1963	1973	1984	1963	1973	1984	1963	1973	1984	1963	1973	1984	1963	1973	1984	1963	1973	1984
San José	487658	695163	890434	202650	338494	567400	23577	19708	27446	19946	16077	23700	86405	131041	117690	3631	3	3746
Alajuela	240672	326032	427962	90033	141862	242231	11585	9297	12899	9661	7510	11248	42507	66971	60004	1924	1787	1651
Cartago	155433	204699	271671	58940	87898	151590	7718	6154	7911	6175	4979	6854	28353	43729	38148	1543	1175	1057
Heredia	85063	133844	197575	35526	59914	113862	3681	3520	5659	3103	2841	4909	15894	25031	27025	578	679	750
Guanacaste	142555	178691	195208	46737	73938	109056	8033	5812	6260	6888	4817	5487	29661	39913	32043	1145	995	773
Puntarenas	156508	218208	265883	50628	83639	137237	8035	7616	9145	6625	6243	8236	28680	49467	46013	1410	1373	909
Limón	68385	115143	168076	21794	43964	84813	3185	4159	6109	2370	3284	5401	12576	23240	29755	815	875	708
Total	1336274	1871780	2416809	506308	829709	1406189	65814	56266	75429	54768	45751	65835	244076	379392	350678	11046	10515	9594

Fuentes: DGEC, Departamento de Estadística del MEP, Sección de Archivo del TSE

Cuadro A 1.4
Población total e información sobre dos indicadores sintomáticos
por distrito, años 1963, 1973 y 1984

CÓDIGO	Distrito	Población			Electores			Crecimiento Nat.		
		1963	1973	1984	1963	1973	1984	1963	1973	1984
01-01-01	Carmen	9221	8442	5942	5869	6414	6491	458	214	226
01-01-02	Merced	22064	19938	15363	14002	15169	15566	731	465	426
01-01-03	Hospital	38361	34689	27097	24345	26397	23873	1589	706	857
01-01-04	Catedral	31516	28590	21574	19998	21752	22906	1299	677	541
01-01-05	Zapote	12769	17311	18504	4697	8670	14139	504	373	412
01-01-06	S.Fco. dos Ríos	6422	10241	15256	2619	5321	10070	399	277	403
01-01-07	Uruca	3766	7441	8932	1795	3071	6506	169	194	352
01-01-08	Mata Redonda	3373	7666	9486	1637	3337	6772	146	146	193
01-01-09	Pavas	5530	18068	31690	2001	5814	17222	251	338	882
01-01-10	Hatillo	12750	31036	52539	3884	14039	33944	614	676	1052
01-01-11	San Sebastián	23166	32019	35081	5846	14942	24034	980	692	994
01-02-01	Escazú	6182	9858	10505	3040	5338	8920	301	334	381
01-02-02	San Antonio	4675	8405	12338	1523	2635	5362	147	159	288
01-02-03	San Rafael	3393	6763	10258	974	2521	5472	102	108	139
01-03-01	Desamparados	11511	30659	43352	4198	12460	28246	522	661	1191
01-03-02	San Miguel	4569	8171	12053	1776	3315	7328	203	200	322
01-03-03	S.Juan de Dios	2836	6465	9540	1037	2262	5676	117	160	301
01-03-04	S.Rafael Arriba	4577	11307	19105	1428	5933	11102	173	246	569
01-03-05	San Antonio	2735	5147	7519	1154	2149	4759	109	100	192
01-03-06	Frailles	1959	2491	2622	682	908	1497	89	78	46
01-03-07	Patarrá	2902	6347	10741	852	2201	5151	95	102	210
01-03-08	San Cristobal	1546	2137	2286	594	862	1354	115	91	61
01-03-09	Rosario	1210	1548	1606	395	496	1048	58	35	34
01-04-01	Santiago	5633	7178	7843	2139	3405	5438	279	146	241
01-04-02	Mercedes Sur	9324	8643	6822	2580	3906	4409	312	230	169
01-04-03	Barbacoas	2398	2326	2422	833	1174	1647	112	62	49
01-04-04	Grifo Alto	1323	1249	916	514	553	586	59	33	18
01-04-05	San Rafael	2104	2075	1426	719	857	969	61	22	21
01-04-06	Candelarita	1357	1198	1333	557	709	886	51	32	30
01-04-07	Desamparaditos	616	538	467	323	292	365	25	13	11
01-04-08	San Antonio	935	943	1894	517	715	1116	44	32	32
01-05-01	San Marcos	2546	3838	5381	1181	1807	2951	133	86	168
01-05-02	San Lorenzo	2055	2732	2391	731	1030	1374	80	53	40
01-05-03	San Carlos	791	972	1073	337	443	609	26	35	33
01-06-01	Aserri	5778	11653	20968	2330	4375	10850	236	239	598
01-06-02	Tarbaca	576	805	899	338	419	719	30	27	35
01-06-03	Vuelta de Jorco	3082	3029	3608	864	1363	1963	138	65	74
01-06-04	San Gabriel	2111	2442	2968	740	1008	1461	72	62	59
01-06-05	La Legua	2062	2162	2145	552	905	1073	107	96	66

Cuadro A 1.4
Continuación

CÓDIGO	Distrito	Población			Electores			Crecimiento Nat.		
		1963	1973	1984	1963	1973	1984	1963	1973	1984
01-07-01	Colón	3542	5122	7361	1590	2281	4302	142	101	199
01-07-02	Guayabo	1135	1288	1612	594	699	1027	37	36	29
01-07-03	Tabarcia	2442	2864	2715	866	1027	1440	97	56	50
01-07-04	Piedras Negras	735	662	222	261	207	191	21	7	4
01-07-05	Picagres	1082	797	674	478	418	433	25	14	7
01-08-01	Guadalupe	21958	27016	25506	9821	17662	21470	1008	736	1151
01-08-02	San Francisco	3900	3994	3138	1756	2451	2750	123	67	56
01-08-03	Calle Blancos	10836	12408	16155	2493	3996	7393	244	180	258
01-08-04	Mata de Plátano	1691	3569	7490	575	1222	3319	56	72	26
01-08-05	Ipis	4909	13349	26151	1426	5132	13523	190	229	495
01-08-06	Rancho Redondo	816	1231	1491	390	594	1099	32	35	12
01-09-01	Santa Ana	3137	5315	6200	1355	2248	4350	156	132	194
01-09-02	Salitral	1267	1645	2049	451	703	1219	60	35	44
01-09-03	Pozos	1415	2247	3558	559	916	1803	53	55	69
01-09-04	Uruca	1182	2097	3267	441	825	1993	47	46	76
01-09-05	Piedades	1663	2642	3614	652	991	1748	56	36	70
01-09-06	Brasil	362	553	917	96	201	543	10	11	19
01-10-01	Alajuelita	4280	6673	8280	1664	3255	5565	188	178	266
01-10-02	San Josecito	1570	3445	4872	605	1515	3779	49	81	186
01-10-03	San Antonio	530	725	1336	164	248	465	12	6	11
01-10-04	Concepción	3126	9280	12175	646	3477	7745	130	224	367
01-10-05	San Felipe	1342	2890	4727	352	903	2148	34	41	117
01-11-01	San Isidro	4607	9824	15179	1908	4283	8354	203	210	448
01-11-02	San Rafael	3132	4470	5982	1166	1853	3153	71	83	115
01-11-03	Jesús(D. Nombre)	2876	2042	3353	857	784	1868	78	37	77
01-12-01	San Ignacio	3322	4420	5036	1562	2079	3099	192	117	131
01-12-02	Guatíl	2368	2265	2196	838	1038	1286	94	51	55
01-12-03	Palmichal	2536	2270	2691	904	1066	1424	77	59	66
01-12-04	Cangrejal	1712	2181	1924	666	824	1095	90	53	53
01-12-05	Sabanillas	3154	3249	3006	792	1096	1398	149	105	82
01-13-01	San Juan	10838	18670	22415	4837	8842	17515	365	284	658
01-13-02	Cinco Esquinas	10624	13385	27140	4741	8534	15747	535	392	715
01-13-03	Anselmo LLorente	2484	3547	8138	1108	2476	5549	119	156	202
01-14-01	San Vicente	9188	15552	24661	3703	6881	14549	330	304	528
01-14-02	San Jerónimo	906	1342	2761	310	566	1161	41	41	48
01-14-03	Trinidad	1548	2654	5616	515	1033	2880	41	51	146
01-15-01	San Pedro	17342	23721	24519	6115	12305	17925	650	570	668
01-15-02	Sabanilla	2447	3475	6762	907	2173	5044	119	120	221
01-15-03	Mercedes	2800	3815	4419	808	1559	2236	49	53	39
01-15-04	San Rafael	1454	2622	3365	819	1226	1806	63	60	100
01-16-01	San Pablo	784	702	755	356	447	569	38	29	22

Cuadro A 1.4
Continuación

CÓDIGO	Distrito	Población			Electores			Crecimiento Nat.		
		1963	1973	1984	1963	1973	1984	1963	1973	1984
01-16-02	San Pedro	559	624	595	191	290	314	20	13	16
01-16-03	S.Juan de Mata	3703	2907	2648	1150	1186	1256	207	105	65
01-16-04	San Luis	450	476	473	194	215	309	21	16	6
01-17-01	Santa María	2424	2683	3324	985	1372	2131	87	42	76
01-17-02	Jardín	219	428	368	44	118	150	5	5	2
01-17-03	Copey	1075	1264	1242	324	542	694	20	25	20
01-18-01	Curridabat	5764	9581	19821	2165	4323	10476	205	188	472
01-18-02	Granadilla	2316	3685	4940	585	920	2180	47	52	138
01-18-03	Sánchez	816	872	1418	175	301	579	20	11	22
01-18-04	Tirrases	595	1453	5775	199	477	2469	11	25	113
01-19-01	S.Isidro del Gral	23831	32929	37062	7751	13213	21242	951	941	1266
01-19-02	General	2310	3250	3576	683	1326	2005	100	76	106
01-19-03	Daniel Flores	3407	5527	13069	1519	2572	5944	169	158	357
01-19-04	Rivas	2449	4885	4977	1095	1718	2496	159	139	141
01-19-05	San Pedro	15322	20498	23686	3839	7727	11806	656	704	900
01-20-01	San Pablo	1584	2103	2532	753	1017	1684	92	58	86
01-20-02	San Andrés	1048	1236	1417	351	596	699	46	26	24
01-20-03	LLano Bonito	1289	1645	1497	411	607	812	61	37	36
01-20-04	San Isidro	707	981	1138	270	431	524	38	23	18
01-20-05	Santa Cruz	1022	1556	1503	238	541	815	17	18	39
02-01-01	Alajuela	24224	33122	34556	12082	17727	27533	868	653	1007
02-01-02	San José	4866	9523	17916	1374	3065	8117	164	182	370
02-01-03	Carrizal	1988	2694	3143	895	1129	1849	114	65	85
02-01-04	San Antonio	3728	7519	11471	1475	3840	8473	199	253	361
02-01-05	Guácima	3478	4265	5740	1383	1895	3000	125	104	113
02-01-06	San Isidro	5180	7239	10118	1529	2890	5256	161	129	214
02-01-07	Sabanilla	4155	4765	4899	1169	1647	2232	134	98	112
02-01-08	San Rafael	3719	5517	8848	1173	2093	4082	124	120	144
02-01-09	Río Segundo	2156	4424	8144	1206	2767	6372	118	110	163
02-01-10	Desamparados	2440	6670	8268	703	1269	3172	69	74	119
02-01-11	Turrucares	2377	2699	3340	1133	1381	2185	104	55	79
02-01-12	Tambor	2080	3163	5148	1193	1760	3572	107	70	121
02-01-13	Garita	2109	3138	4114	795	1231	2103	73	61	80
02-01-14	Sarapiquí	1898	1587	1767	446	576	728	75	47	12
02-02-01	San Ramón	6973	9245	9624	3776	5638	8727	291	201	242
02-02-02	Santiago	1481	1653	1798	791	974	1359	79	42	56
02-02-03	San Juan	1783	3499	5757	576	971	2793	49	50	176
02-02-04	Piedades Norte	2324	2338	2373	971	1132	1462	102	53	46
02-02-05	Piedades Sur	2341	2311	2339	1013	1124	1522	119	62	62
02-02-06	San Rafael	2921	3467	4716	1033	1402	2135	112	66	120
02-02-07	San Isidro	1149	1226	1281	416	512	939	46	26	28

Cuadro A 1.4
Continuación

CÓDIGO	Distrito	Población			Electores			Crecimiento Nat.		
		1963	1973	1984	1963	1973	1984	1963	1973	1984
02-02-08	Ángeles	2024	2267	3616	546	754	1395	68	39	59
02-02-09	Alfaro	1102	1348	2030	375	490	775	36	22	31
02-02-10	Volio	798	777	1062	343	383	553	39	16	26
02-02-11	Concepción	775	911	1083	332	452	627	32	26	30
02-02-12	Zapotal	815	700	490	294	308	285	39	18	11
02-02-13	Peñas Blancas	1439	3413	3794	383	1168	1897	58	102	98
02-03-01	Grecia	7065	9984	11374	3643	5137	8297	312	218	342
02-03-02	San Isidro	2843	3271	2891	757	1258	1809	86	38	54
02-03-03	San José	2174	2674	3344	925	1458	2328	119	71	70
02-03-04	San Roque	3108	4087	5559	938	1682	2871	107	90	90
02-03-05	Tacares	3152	3439	3902	1191	1744	2375	98	106	116
02-03-06	Río Cuarto	2116	2840	3341	683	924	1619	83	70	84
02-03-07	Puente de Piedra	2260	2704	4493	1180	1593	2105	99	84	40
02-03-08	Bolivar	1982	2807	3457	776	1285	2065	85	69	60
02-04-01	San Mateo	2876	2514	3132	1306	1444	1952	107	26	53
02-04-02	Desmonte	512	455	651	166	197	240	16	5	9
02-05-01	Atenas	2962	3592	4451	1582	2175	3231	123	53	136
02-05-02	Jesús	2111	2355	2671	912	1260	1736	85	42	30
02-05-03	Mercedes	1103	1191	1506	478	584	785	34	13	13
02-05-04	San Isidro	1539	1867	1891	585	839	1118	55	47	46
02-05-05	Concepción	1288	1397	1953	600	726	1090	49	29	34
02-05-06	San José	1249	1248	1564	497	564	634	36	24	18
02-05-07	Santa Eulalia	766	960	975	294	474	656	31	19	25
02-06-01	Naranjo	9781	13959	16125	4202	6767	10251	377	277	404
02-06-02	San Miguel	2498	1357	1546	942	579	950	128	29	44
02-06-03	San José	1862	1763	2173	702	880	1318	63	33	44
02-06-04	Cirrí Sur	1389	1603	2394	595	738	1310	57	40	59
02-06-05	San Jerónimo	884	1039	1350	416	511	937	29	24	44
02-07-01	Palmares	2348	3083	3766	1330	2000	3118	102	58	105
02-07-02	Zaragoza	3182	3462	4136	1287	1895	3128	141	50	125
02-07-03	Buenos Aires	3235	4064	5270	1255	1726	2755	116	50	86
02-07-04	Santiago	1871	1946	1964	475	630	942	72	23	34
02-07-05	Candelaria	688	740	908	349	529	831	50	23	43
02-07-06	Esquipulas	959	1200	1771	329	553	946	38	22	39
02-08-01	San Pedro	3229	3995	5060	1354	1939	3118	122	69	124
02-08-02	San Juan	956	1143	1566	417	553	786	35	17	20
02-08-03	San Rafael	2026	2301	2586	818	1145	1623	74	41	65
02-08-04	Carrillos	1411	2170	3844	401	845	1741	49	58	95
02-08-05	Sabana Redonda	557	582	883	222	313	395	22	14	22
02-09-01	Orotina	3925	5038	6539	1948	2560	3768	158	60	130
02-09-02	Mastate	680	917	1019	282	378	592	15	9	24

Cuadro A 1.4
Continuación

CÓDIGO	Distrito	Población			Electores			Crecimiento Nat.		
		1963	1973	1984	1963	1973	1984	1963	1973	1984
02-09-03	Hacienda Vieja	774	766	728	334	379	481	26	9	6
02-09-04	Coyolar	971	946	1161	509	535	694	43	21	19
02-09-05	Ceiba	743	812	1047	291	381	533	43	15	23
02-10-01	Quesada	11689	17027	21696	4860	7894	13345	572	466	749
02-10-02	Florencia	7394	7237	7459	2061	2859	4152	387	286	244
02-10-03	Buena Vista	351	255	289	183	211	199	28	13	7
02-10-04	Aguas Arcas	5097	11477	20730	1323	3644	8318	200	417	681
02-10-05	Venecia	2229	3261	4381	757	1317	2280	90	85	108
02-10-06	Pital	3324	4100	6614	898	1641	3155	154	136	201
02-10-07	Fortuna	2931	3710	7126	997	1704	3356	230	197	199
02-10-08	Tigra	1653	1741	2644	461	674	1191	61	47	78
02-10-09	Palmera	1918	2432	2844	484	950	1291	87	62	54
02-11-01	Zarcelero	1499	2148	2517	839	1261	1723	59	39	77
02-11-02	Laguna	668	893	955	301	415	578	20	18	13
02-11-03	Tapezco	1007	1585	1714	373	624	897	36	37	37
02-11-04	Guadalupe	452	405	417	126	181	236	12	7	10
02-11-05	Palmira	455	580	767	179	246	442	13	10	17
02-11-06	Zapote	838	731	635	226	271	368	31	15	15
02-12-01	Sanchí Norte	3755	3525	4336	1710	1958	2955	177	106	137
02-12-02	Sarchí Sur	2443	3392	4576	691	1393	2576	69	69	67
02-12-03	Toro Amarillo	348	356	273	139	166	161	15	11	12
02-13-01	Upala	10001	15971	26061	1825	4500	8707	418	681	932
02-14-01	Los Cliles	3231	5596	11404	750	1750	3617	150	224	373
02-15-01	San Rafael	5991	9859	10098	1324	2820	3836	147	198	188
03-01-01	Parte Oriental	10344	13489	14428	5210	6089	11209	773	451	581
03-01-02	Parte Occidental	7740	8264	9500	3620	5121	6519	103	68	173
03-01-03	Carmen	4241	7392	11997	1210	2164	6037	108	86	182
03-01-04	San Nicolás	5705	9384	14675	2141	3351	7586	188	202	384
03-01-05	Agua Cliente(S.Fco)	2796	5184	8963	919	1623	3886	88	105	228
03-01-06	Guadalupe(Arenilla)	5145	7745	10476	1854	5358	5777	178	203	266
03-01-07	Corralillo	4627	5928	6538	1554	2138	3720	183	174	129
03-01-08	Tierra Balnaca	2284	2656	3527	935	1102	1799	94	72	108
03-01-09	Dulce Nombre	2052	3351	4733	793	1321	2885	93	69	65
03-01-10	Llano grande	1788	1917	2288	551	727	1233	55	56	66
03-02-01	Paraiso	7225	10211	14880	2692	4437	8568	261	229	411
03-02-02	Santiago	2402	2854	3303	712	1079	1735	88	75	64
03-02-03	Orosí	5363	5538	5705	1469	2186	3076	180	147	144
03-02-04	Cachí	3399	3678	3935	986	1501	2317	128	107	95
03-03-01	Tres Ríos	4694	6804	8661	2195	3370	6235	265	197	306
03-03-02	San Diego	2581	4584	12462	784	1412	4259	67	93	258
03-03-03	San Juan	971	1490	1941	312	475	1148	24	24	59

Cuadro A 1.4
Continuación

CÓDIGO	Distrito	Población			Electores			Crecimiento Nat.		
		1963	1973	1984	1963	1973	1984	1963	1973	1984
03-03-04	San Rafael	1775	2703	5064	588	1130	2345	36	44	70
03-03-05	Concepción	1550	3846	6638	655	1323	3843	55	89	195
03-03-06	Dulce Nombre	1312	2120	3734	476	710	1647	34	37	83
03-03-07	San Ramón	1191	1805	2505	196	490	1150	32	40	65
03-04-01	Juan Viñas	5813	6234	5543	2490	2874	3570	230	156	131
03-04-02	Tucurrique	4626	5289	6318	1687	2091	3115	166	157	159
03-05-01	Turrialba	19932	24424	27579	9103	11386	17142	984	533	653
03-05-02	La Suiza	8259	9251	11651	2828	3856	5808	330	252	267
03-05-03	Peralta	6236	7094	8678	1806	3277	4996	175	221	251
03-05-04	Santa Cruz	3193	2492	2659	1323	1502	1774	136	91	68
03-06-01	Pacayas	2795	3400	3760	1147	1451	2040	113	75	96
03-06-02	Cervantes	2377	2630	3138	665	977	1539	57	68	62
03-06-03	Capellades	1293	1454	1440	345	465	676	34	21	23
03-07-01	San Rafael	6667	10434	14824	2297	4306	8611	229	240	372
03-07-02	Cot	2489	3492	4595	831	1355	2280	110	99	141
03-07-03	Potrero Grande	888	888	959	346	416	571	46	33	29
03-07-04	Cipreces	886	1248	1748	331	509	865	47	44	46
03-07-05	Santa Rosa	1105	1455	2019	322	629	1105	45	52	46
03-08-01	El Tejar	4119	6449	11158	1757	2718	5649	174	144	325
03-08-02	San Isidro	3788	5081	6287	1089	1910	3073	169	161	177
03-08-03	Tobosi	1467	2161	2833	633	932	1615	90	58	75
03-08-04	Patio de Agua	315	339	529	91	138	186	6	8	3
04-01-01	Heredia	19249	23225	22009	9558	14014	17108	852	615	881
04-01-02	Mercedes	3381	6536	12719	936	2392	6671	64	90	196
04-01-03	San Francisco	1001	2451	12770	475	991	5368	42	50	205
04-01-04	Ulloa	2432	4275	7398	747	1251	4694	64	65	273
04-02-01	Barba	1938	3131	3911	1120	1756	3046	83	98	145
04-02-02	San Pedro	1762	2781	4882	763	1192	2652	69	56	155
04-02-03	San Pablo	1454	2377	3048	437	794	1515	52	41	74
04-02-04	San Roque	622	886	1750	190	310	537	15	11	17
04-02-05	Santa Lucía	1187	1608	2560	319	601	1000	40	71	74
04-02-06	S. José de la Mont.	1541	2081	2782	572	915	1500	65	54	81
04-03-01	Santo Domingo	3608	5148	5032	2366	3460	5457	177	171	256
04-03-02	San Vicente	995	1565	3302	180	431	1151	6	11	4
04-03-03	San Miguel	2893	4394	6735	1140	1910	3878	80	62	123
04-03-04	Paracito	448	914	1388	167	394	1144	11	21	104
04-03-05	Santo Tomás	1542	2257	3171	526	822	1233	29	28	41
04-03-06	Santa Rosa	1862	3145	4357	680	1128	2086	47	66	85
04-04-01	Sata Bárbara	1648	2496	3667	921	1325	2475	70	84	132
04-04-02	San Pedro	981	1337	1871	424	700	1155	20	22	27
04-04-03	San Juan	1352	1830	3502	422	699	1371	35	29	71

Cuadro A 1.4
Continuación

CÓDIGO	Distrito	Población			Electores			Crecimiento Nat.		
		1963	1973	1984	1963	1973	1984	1963	1973	1984
04-04-04	Jesús	1556	2144	3717	627	899	1883	59	40	89
04-04-05	Santo Domingo	2590	2931	3886	701	1229	2092	89	44	85
04-05-01	San Rafael	2835	4720	7141	1443	2189	4352	146	122	272
04-05-02	San Josecito	2395	6024	7049	597	2065	3910	32	58	112
04-05-03	Santiago	599	959	2837	203	365	1261	8	13	28
04-05-04	Los Ángeles	1662	2783	3930	685	1183	2531	71	55	93
04-05-05	Concepción	1245	1527	1914	356	578	939	44	25	32
04-06-01	San Isidro	2461	3738	5813	1193	1688	3382	90	70	155
04-06-02	San José	878	1281	1865	391	574	984	28	14	27
04-06-03	Concepción	722	960	850	334	472	782	28	16	20
04-07-01	San Antonio	2429	4334	5895	1371	2132	3569	109	113	171
04-07-02	La Ribera	1415	2592	3855	595	1004	2104	25	38	64
04-07-03	Asunción	937	1612	2243	421	723	1249	22	23	27
04-08-01	San Joaquín	2222	3624	3838	1381	1761	2956	97	93	228
04-08-02	Barrantes	1075	1457	1895	432	690	1429	24	27	23
04-08-03	Lllorente	865	1443	3282	348	743	1598	19	13	19
04-09-01	San Pablo	4092	6660	11802	1426	2646	6508	145	118	204
04-10-01	Puerto Viejo	4856	12618	18909	1080	3890	8303	250	388	576
05-01-01	Liberia	10217	18604	24498	3891	7443	14251	563	527	877
05-01-02	Cañas Dulces	1980	3177	3569	675	1039	1730	110	94	70
05-02-01	Nicoya	20556	31131	29313	6789	12630	17243	1221	768	807
05-02-02	Mansión	9170	6892	6682	2554	3563	3709	352	183	128
05-02-03	San Antonio	6550	7061	6510	2293	3007	4026	238	134	118
05-03-01	Santa Cruz	8028	10672	12866	3573	5004	7706	384	267	295
05-03-02	Bolsón	2979	3559	3917	1177	1712	2446	103	68	74
05-03-03	Veintisiete de Abril	8475	10274	9044	2388	3619	4912	298	329	215
05-03-04	Tempate	4094	5234	5306	1495	2298	3148	177	134	112
05-04-01	Bagaces	9836	9828	10103	2586	4045	5535	540	263	267
05-05-01	Filadelfia	3111	4484	5353	1180	2059	3004	158	111	153
05-05-02	Palmira	1187	1574	2529	322	588	1166	26	24	44
05-05-03	Sardinal	3712	4732	5952	1367	2050	3158	160	112	164
05-05-04	Belén	3386	4103	4641	950	1639	2497	132	89	101
05-06-01	Cañas	9117	12779	17284	3358	5473	9543	431	372	530
05-07-01	Juntas	4253	5486	6568	1950	2544	3578	221	147	186
05-07-02	Sierra	2284	1760	1562	382	648	754	78	52	22
05-07-03	San Juan	1793	1595	1475	449	684	958	95	36	36
05-07-04	Colorado	1859	2792	2970	385	837	1427	75	64	91
05-08-01	Tilarán	3426	4601	5926	1481	2189	3596	168	98	232
05-08-02	Quebrada Grande	1933	1720	1879	726	940	1114	108	61	38
05-08-03	Tronadora	2565	2623	3256	825	1100	1792	188	108	104
05-08-04	Santa Rosa	1750	1407	1309	619	595	671	101	34	26

Cuadro A 1.4
Continuación

CÓDIGO	Distrito	Población			Electores			Crecimiento Nat.		
		1963	1973	1984	1963	1973	1984	1963	1973	1984
05-08-05	Líbano	812	875	993	337	443	590	42	28	23
05-08-06	Tierras Morenas	1611	1337	1223	368	584	795	53	28	26
05-09-01	Carmona	1223	1629	1654	432	632	931	70	44	45
05-09-02	Santa Rita	1747	1858	1423	461	750	788	50	47	37
05-09-03	Zapotal	1690	1719	1508	599	800	869	80	47	38
05-09-04	San Pablo	1599	1485	1299	487	647	773	51	60	29
05-09-05	Porvenir	1881	1543	1338	587	818	819	84	57	22
05-09-06	Bejuco	3898	3824	2382	904	1291	1298	184	125	62
05-10-01	La Cruz	5833	8333	10876	1048	2266	4229	348	410	514
06-01-01	Puntarenas	26082	35782	48250	10352	17310	28121	965	770	1368
06-01-02	Pitahaya	3320	3387	3502	1393	1463	1699	103	93	89
06-01-03	Chomes	1991	2378	2686	706	1021	1469	107	70	70
06-01-04	Lepanto	10047	9928	9064	2617	3911	4796	502	345	211
06-01-05	Paquera	9964	9285	8323	3268	3712	3982	501	256	186
06-01-06	Manzanillo	1662	2163	2868	600	850	1246	85	92	85
06-01-07	Guacimal	2826	2639	2586	948	1228	1529	150	97	91
06-02-01	Espíritu Santo	4895	6686	9649	1915	3100	4914	181	144	219
06-02-02	San Juan Grande	1260	1465	1579	476	507	765	42	23	24
06-02-03	Macacona	1210	2305	1947	320	887	1676	31	42	74
06-02-04	San Rafael	770	811	1041	333	344	499	18	16	9
06-02-05	San Gerónimo	1040	828	782	458	412	475	31	15	15
06-03-01	Buenos Aires	4624	8923	11651	1046	2661	5636	141	345	464
06-03-02	Volcán	2793	3589	3551	474	1174	1558	102	113	94
06-03-03	Potrero Grande	2576	5700	9655	435	1337	3392	121	174	355
06-03-04	Buruca	1049	1892	2861	196	527	1149	36	44	123
06-04-01	Miramar	3480	3785	4191	1360	1779	2645	175	83	115
06-04-02	Unión	1696	1327	1250	797	613	732	110	46	15
06-04-03	San Isidro	1440	1867	2003	351	736	1005	36	47	40
06-05-01	Puerto Córtes	5405	7467	7368	2354	3182	4371	350	250	242
06-05-02	Palmar	8052	13660	14845	2520	4849	7209	448	440	439
06-05-03	Sierpe	4117	3486	4081	819	882	1935	73	70	102
06-06-01	Quepos	10228	11054	10853	3489	4791	6284	412	385	337
06-06-02	Savegre	2837	3419	2466	705	1197	1481	102	95	69
06-07-01	Golfito	36567	62481	89059	10230	20748	43186	1402	1961	3115
06-09-01	Parrita	6877	11901	9774	2472	4418	5484	402	343	273
07-01-01	Limón	29039	40830	52602	10745	18378	31602	873	1090	1716
07-02-01	Guapiles	7616	26026	41466	1354	7902	19959	300	1085	1485
07-02-02	Jiménez	2399	7297	9522	485	2457	4121	130	243	269
07-03-01	Siquirres	5437	9799	17522	1686	3608	7815	237	244	603
07-03-02	Pacuarito	1437	2659	4357	388	803	1842	54	36	94
07-03-03	Florida	1054	1456	1794	204	382	695	48	19	23

**Cuadro A 1.4
Continuación**

CÓDIGO	Distrito	Población			Electores			Crecimiento Nat.		
		1963	1973	1984	1963	1973	1984	1963	1973	1984
07-03-04	Germania	1453	2423	2995	601	894	1312	42	73	45
07-03-05	Cairo	1936	1796	2411	514	693	992	71	51	31
07-04-01	Bratsi	3541	5431	11013	725	2028	4788	152	189	453
07-05-01	Matina	7561	10489	14723	2057	3873	7149	256	285	430
07-06-01	Guácimo	5731	4755	6158	917	2043	3040	163	174	160
07-06-02	Mercedes	572	789	796	80	301	351	7	7	6
07-06-03	Pocora	609	1393	2717	186	603	1149	37	63	74
Total		1336274	1871780	2416809	506308	829709	1406189	54768	45751	65835

Fuentes: DGE, Sección de Archivo dei TSE.

Cuadro A 1.5
Viviendas, población en viviendas simples y colectivas y abonados del
servicio eléctrico por cantón, 1984

(Se incluye además los abonados residenciales del servicio eléctrico a junio de 1996)

Información del censo de 1984

Cantón	Viviendas simples		Población			Promedio personas por vivien. simp.	Abonados eléctricos al 30-06-96
	Total	Con servicio eléctrico	Total	En viviendas simples	En viviendas colectivas.		
San José	55019	54666	241464	239425	2039	4,352	73981
Escazú	6930	6658	33101	32977	124	4,759	11610
Desamparados	23026	22230	108824	108631	193	4,718	39250
Goicoechea	17869	17720	79931	79438	493	4,446	26541
Santa Ana	4079	3866	19605	19534	71	4,789	6783
Alajuelita	6586	6347	31390	31367	23	4,763	11478
Coronado	5050	4845	24514	24393	121	4,830	11155
Tibás	12241	12173	57693	57640	53	4,709	16824
Moravia	7233	7124	33038	32898	140	4,548	11327
Montes de Oca	9494	9419	39065	38773	292	4,084	14423
Cumidabat	6906	6576	31954	31877	77	4,616	13171
Alajuela	26827	26048	127472	125363	2109	4,673	46826
Grecia	7806	7346	38361	38346	15	4,912	14060
Palmares	3550	3377	17815	17815	0	5,018	6331
Poás	2841	2594	13939	13939	0	4,906	5180
Alfaro Ruiz	1427	1312	7005	6989	16	4,898	2298
Cartago	16822	16169	87125	86628	497	5,150	25780
Paraiso	5357	5095	27823	27823	0	5,194	10703
La Unión	8042	7655	41005	40427	578	5,027	15306
Oreamuno	4319	4139	24145	24093	52	5,578	7263
Heredia	11574	11434	54896	54520	376	4,711	23400
Barba	3782	3654	18933	18901	32	4,998	6514
Santo Domingo	5046	4886	23985	23973	12	4,751	7353
Sata Bárbara	3316	3150	16643	16610	33	5,009	5672
San Rafael	4639	4397	22871	22871	0	4,930	7636
San Isidro	1732	1647	8528	8528	0	4,924	3474
Belén	2271	2246	11993	11922	71	5,250	4202
Flores	1813	1796	9015	8980	35	4,953	2734
San Pablo	2379	2321	11802	11753	49	4,940	3816

Fuentes: DGE, Departamento de Estadística del ICE.

ANEXO 2

Ejemplos de aplicación de los métodos

Para tener una mejor comprensión con respecto a la aplicación de las técnicas de estimación poblacional, este anexo presenta un ejemplo de cálculo. Se efectuaron estimaciones de la población del cantón Central de San José, utilizando en cada caso la información del padrón electoral.

10.2.1 Método de Razón Censal

Para la aplicación de esta técnica, se debe contar con la tasa de electores para cada cantón y para Costa Rica en 1973. Además, un estimado de la tasa de electores para Costa Rica en 1984. De acuerdo con la información proporcionada por el Cuadro A 1.3 del Anexo 1, para el cantón Central de San José se tiene:

	Electores Inscritos		Población Censal	
	1973	1984	1973	1984
San José	124926	181522	215441	241464
Costa Rica	829709	1406189	1871780	2416809

Entonces, si se representa con r la tasa de electores inscritos, se tiene que:

$$r(\text{San José})_{1973} = \frac{124926}{215441} = 0,57986$$

$$r(\text{Costa Rica})_{1973} = \frac{829709}{1871780} = 0,44327$$

$$r(\text{Costa Rica})_{1984} = \frac{1406189}{2416809} = 0,58184$$

La técnica supone que la tasa de electores de San José para 1984, viene dada por:

$$r(\text{San José})_{1984} = \phi r(\text{San José})_{1973}$$

donde el valor de ϕ , puede ser estimado por:

$$\hat{\phi} = \frac{\hat{r}(\text{Costa Rica})_{1984}}{r(\text{Costa Rica})_{1973}} = \frac{0,58184}{0,44327} = 1,31261$$

Con lo cual, la estimación de $r(\text{San José})_{1984}$ es:

$$\hat{r}(\text{San José})_{1984} = 1,31261 \cdot 0,57986 = 0,76113$$

La estimación simple de la población para el cantón central de San José en 1984 viene dada por:

$$\hat{P}(\text{San José})_{1984} = \frac{\text{Elec. Insc.}_{1984}}{\hat{r}(\text{San José})_{1984}} = \frac{181522}{0,76113} \cong 238492$$

Este proceso se efectuó para los 68 cantones, por lo que la suma de las estimaciones poblacionales de todos ellos debe coincidir con la población total del país (2416809). Sin embargo, la suma de las estimaciones simples para todos los cantones es 2448305. Por esta razón es necesario prorratear las estimaciones de manera que la suma de ellas coincida con la población censal de Costa Rica para 1984. Para este fin, la estimación de cada cantón se multiplica por el cociente: $\frac{2416809}{2448305}$. Por lo tanto, la estimación de la población de San José, al 6 de junio de 1984 (fecha del Censo), viene dada por:

$$\frac{2416809}{2448305} \cdot 238492 \cong 235424$$

Entonces el error porcentual para la estimación anterior es:

$$EP = \frac{(235424 - 241464)}{241464} \cdot 100 \cong -2,5\%$$

Se puede decir que el método de Razón Censal produjo una subestimación de un 2,5% con respecto a su valor censal de la población de San José.

Método de Diferencia de Tasas

Este método es una variante del procedimiento de Razón Censal. La única modificación radica en la forma de estimar la tasa de ocurrencia del indicador sintomático. En el ejemplo se obtuvo que:

$$r(\text{San José})_{1973} = 0,57986$$

$$r(\text{Costa Rica})_{1973} = 0,44327$$

$$r(\text{Costa Rica})_{1984} = 0,58184$$

Este procedimiento supone que la tasa de electores de San José, para 1984, viene dada por:

$$r(\text{San José})_{1984} = r(\text{San José})_{1973} + [r(\text{Costa Rica})_{1984} - r(\text{Costa Rica})_{1973}]$$

Con lo cual, la estimación de $r(\text{San José})_{1984}$ es:

$$\hat{r}(\text{San José})_{1984} = 0,57986 + [0,58184 - 0,44327] = 0,71843$$

Así, la estimación simple de la población para 1984 está dada por:

$$\hat{P}(\text{San José})_{1984} = \frac{\text{Elec. Insc.}_{1984}}{\hat{r}(\text{San José})_{1984}} = \frac{181522}{0,71843} \cong 252666$$

Para este caso, la suma de las estimaciones simples para todos los cantones es 2430006 y la estimación de la población de San José es:

$$\frac{2416809}{2430006} \cdot 252666 \cong 251294$$

El error porcentual asociado es :

$$EP = \frac{(251294 - 241464)}{241464} \cdot 100 \cong 4,1\%$$

Por lo que se concluye que la estimación poblacional sobreestima el valor censal en 4,1%.

10.2.3 Método de Razón Correlación

Para la aplicación de este método se requiere nuevamente parte de la información del Cuadro A 1.3 del Anexo 1, correspondiente al cantón de San José:

	Electores Inscritos			Población Censal		
	1963	1973	1984	1963	1973	1984
San José	86694	124926	181522	168938	215441	241464
Costa Rica	506308	829709	1406189	1336274	1871780	2416809

La técnica utiliza el modelo de regresión lineal:

$$Y_u = a_0 + a_1 X_u$$

donde a_i ($i = 0,1$) corresponde a los coeficientes obtenidos por mínimos cuadrados, y para este caso:

$$Y_u = \frac{\left(\frac{P_{73}(u)}{P_{73}(M)} \right)}{\left(\frac{P_{63}(u)}{P_{63}(M)} \right)} \quad \text{y} \quad X_u = \frac{\left(\frac{S_{73}(u)}{S_{73}(M)} \right)}{\left(\frac{S_{63}(u)}{S_{63}(M)} \right)} \quad \text{para todo cantón } u.$$

$P(u)$ es la población censal correspondiente al cantón u

$P(M)$ es la población censal de Costa Rica

$S(u)$ es el número de electores inscritos en el padrón electoral para el cantón u

$S(M)$ es el número de electores inscritos en el padrón electoral para Costa Rica

De este modo, Y_u y X_u representan las razones de cambio de las proporciones que representa cada cantón del total del país, para el período 63-73. Para San José, estas razones están dadas por:

$$Y_{San\ José} = \frac{\left(\frac{215441}{1871780}\right)}{\left(\frac{168938}{1336274}\right)} = \frac{0,115}{0,126} = 0,910 \quad y \quad X_{San\ José} = \frac{\left(\frac{124926}{829709}\right)}{\left(\frac{86694}{506308}\right)} = \frac{0,151}{0,171} = 0,879$$

Una vez determinadas las razones de cambio en el período 63-73 para los 68 cantones, se procede a utilizar mínimos cuadrados para obtener los coeficientes del modelo. El Cuadro A 2.1 presenta los resultados dados por el paquete estadístico SPSS:

Cuadro A 2.1
Resultados en la aplicación procedimiento de mínimos cuadrados, a las variables X_u^* y Y_u^* , obtenidos por el paquete estadístico SPSS

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable. POB6373
Variable(s) Entered on Step Number

1. PAD6373

Multiple R .87201
R Square .76040
Adjusted R Square .75677
Standard Error .10118

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	1	2.14438	2.14438
Residual	66	.67570	.01024

F = 209.45498

Signif F = .0000

————— Variables in the Equation —————

Variable	B	SE B	Beta	T	Sig T
PAD6373	.87505	.06046	.87201	14.473	.0000
(Constant)	.12085	.06191		1.952	.0552

La prueba "F" indica que el modelo es significativo. El valor de $r^2 = 0,76$, muestra que el 76% de la variación de la razón de cambio en la proporción poblacional es explicada por la variación de la razón de cambio en la proporción del número de electores inscritos.

El modelo de regresión lineal estimado viene dado por:

$$\hat{Y}_u = 0,12085 + 0,87505 X_u$$

La razón de cambio en la proporción poblacional para San José, en el período

73-84, se estima por: $\hat{Y}_{San\ José} = 0,12085 + 0,87505 X_{San\ José}$

$$\text{Como } X_{San\ José} = \frac{\left(\frac{181522}{1406189}\right)}{\left(\frac{124926}{829709}\right)} = 0,857351, \text{ entonces } \hat{Y}_{San\ José} = 0,871075$$

$$\text{Para el período 73-84 se tiene que } \hat{Y}_{San\ José} = \frac{\hat{P}_{84}(San\ José)}{\left(\frac{215441}{1871780}\right)} = \frac{2416809}{\left(\frac{215441}{1871780}\right)}$$

$$\text{es decir: } \frac{\hat{P}_{84}(San\ José)}{\left(\frac{215441}{1871780}\right)} = 0,871075$$

Entonces, la estimación simple para la población de San José al 6 de junio de 1984 (fecha del censo), viene dada por :

$$\hat{P}_{84}(San\ José) = \frac{0,871075 \cdot 215441 \cdot 2416809}{1871780} \cong 242310$$

No obstante, la suma de las estimaciones poblacionales para los 68 cantones es 2434460, que no reproduce la población total de Costa Rica (2416809). Por esta razón es necesario proratear las estimaciones simples obtenidas, utilizando el cociente: $\frac{2416809}{2434460}$. Finalmente la estimación poblacional para San José es:

$$\frac{2416809}{2434460} \cdot 242310 \cong 240553$$

Debido a que el valor censal de la población de San José para 1984 fue 241464, el error porcentual de la estimación es:

$$EP = \frac{(240553 - 241464)}{241464} \cdot 100 \cong -0,4\%$$

Lo cual quiere decir que la estimación subestimó en un 0,4% la población censal de San José.

10.2.4 Método de Tasa Correlación

En la aplicación del método de Tasa Correlación se mantienen los principios básicos utilizados por el Método de Razón Correlación. La principal diferencia entre los dos métodos es la forma en que se aplica la razón de cambio en las proporciones. Formalmente el modelo se puede expresar por:

$$Y_u^* = a_0 + a_1 X_u^*$$

donde : $Y_u^* = \frac{\ln(Y_u)}{k}$ y $X_u^* = \frac{\ln(X_u)}{k}$

Y_u y X_u están definidos tal y como se indicó para la técnica de Razón Correlación y k corresponde al intervalo de tiempo transcurrido entre los censos.

Para San José, en el período 63-73, $Y_{San\ José} = 0,910$ y $X_{San\ José} = 0,879$. Además $k = 10,12$ y representa el tiempo (en años) transcurrido entre los censos del 63 y 73. Entonces:

$$Y_{San\ José}^* = \frac{\ln(0,910)}{10,12} = -0,009 \quad \text{y} \quad X_{San\ José}^* = \frac{\ln(0,879)}{10,12} = -0,013$$

Después de calculados estos valores en el período 63-73 para los demás cantones, se utiliza de nuevo mínimos cuadrados para obtener los coeficientes a_0 y a_1 . El Cuadro A 2.2 presenta los resultados obtenidos por medio del paquete estadístico SPSS.

La prueba "F" muestra que el modelo es significativo. El valor de $r^2 = 0,74$, indica que el 74% de la variación de Y_u^* es explicada por medio de la variación en X_u^* .

Cuadro A 2.2

Resultados en la aplicación procedimiento de mínimos cuadrados, a las variables X_u^* y Y_u^* , obtenidos por el paquete estadístico SPSS

**** MULTIPLE REGRESSION ****

Equation Number 1 Dependent Variable. POB6373
Variable(s) Entered on Step Number

1.. PAD6373

Multiple R .86173
R Square .74258
Adjusted R Square .73868
Standard Error .01005

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	1	.01922	.01922
Residual	66	.00666	.00010

F = 190.38862 Signif F = .0000

Variables in the Equation

Variable	B	SE B	Beta	T	Sig T
PAD6373	.91507	.06632	.86173	13.798	.0000
(Constant)	-7.29843E-04	1.22211E-03		-.597	.5524

El modelo de regresión lineal estimado es: $Y_u^* = -0,00073 + 0,91507 X_u^*$. Para el período intercensal 73-84, $k = 11,06$. Entonces para el cantón Central de San José, la estimación de Y_u^* viene dada por:

$$\hat{Y}_{San\ José}^* = -0,00073 + 0,91507 X_{San\ José}^*$$

Puesto que el tiempo transcurrido entre los censos de 1973 y 1984 es $k = 11,06$:

$$X^*_{San\ José} = \frac{\ln \left[\frac{\left(\frac{181522}{1406189} \right)}{\left(\frac{124926}{829709} \right)} \right]}{11,06} = -0,011392, \text{ con lo cual } \hat{Y}^*_{San\ José} = -0,01346.$$

$$\text{Se tiene entonces que } \hat{Y}^*_{San\ José} = \frac{\ln \left[\frac{P_{84}(\text{San José})}{\frac{2416809}{\left(\frac{215441}{1871780} \right)}} \right]}{11,06} = -0,01346. \text{ Por lo tanto,}$$

la estimación simple para la población de San José es :

$$\hat{P}_{84}(\text{San José}) = \frac{215441 \cdot 2416809 \cdot e^{-0,01346 \cdot 11,06}}{1871780} \cong 239687$$

La suma de las estimaciones poblacionales para los 68 cantones es 2423508.

Por lo tanto, la estimación poblacional para San José es:

$$\frac{2416809}{2423508} \cdot 239687 \cong 239025$$

El error porcentual en la estimación es:

$$EP = \frac{(239025 - 241464)}{241464} \cdot 100 \cong -1,0\%$$

Lo que da una subestimación del valor censal de un 1,0%.

ANEXO 3

Cuadro A 3.1

Matriz de correlación de las razones de cambio en las proporciones del tamaño poblacional y de cinco indicadores sintomáticos, a nivel provincial, Período 63-73

Variable	Pob	Nac	P Elec	C Nat	Def	M Esc
Población	1,00					
Nacimientos	0,91	1,00				
Padr Electoral	0,88	0,94	1,00			
Crec Natural	0,87	0,99	0,93	1,00		
Defunciones	0,82	0,61	0,66	0,52	1,00	
Matr Escolar	0,75	0,89	0,74	0,89	0,44	1,00

Cuadro A 3.2

Matriz de correlación de las razones de cambio en las proporciones del tamaño poblacional y de cinco indicadores sintomáticos, a nivel provincial, Período 73-84

Variable	Pob	P Elec	Nac	C Nat	M Esc	Def
Población	1,00					
Padr Electoral	0,99	1,00				
Nacimientos	0,96	0,95	1,00			
Crec Natural	0,94	0,90	0,99	1,00		
Matr Escolar	0,81	0,88	0,78	0,69	1,00	
Defunciones	0,53	0,44	0,62	0,73	0,09	1,00

Cuadro A 3.3

Matriz de correlación de las razones de cambio en las proporciones del tamaño poblacional y de cinco indicadores sintomáticos, a nivel cantonal, Período 63-73

Variable	Pob	P Elec	Nac	C Nat	M Esc	Def
Población	1,00					
Padr Electoral	0,87	1,00				
Nacimientos	0,86	0,85	1,00			
Crec Natural	0,83	0,83	0,99	1,00		
Matr Escolar	0,79	0,77	0,81	0,81	1,00	
Defunciones	0,73	0,69	0,72	0,63	0,50	1,00

Cuadro A 3.4

Matriz de correlación de las razones de cambio en las proporciones del tamaño poblacional y de cinco indicadores sintomáticos, a nivel cantonal, Período 73-84

Variable	Pob	P Elec	Nac	C Nat	Def	M Esc
Población	1,00					
Padr Electoral	0,93	1,00				
Nacimientos	0,89	0,79	1,00			
Crec Natural	0,87	0,76	0,99	1,00		
Matr Escolar	0,74	0,75	0,57	0,58	1,00	
Defunciones	0,69	0,60	0,78	0,70	0,31	1,00

Cuadro A 3.5

Matriz de correlación de las razones de cambio en las proporciones del tamaño poblacional y de dos indicadores sintomáticos, a nivel distrital, Período 63-73

Variable	Pob	P Elec	C Nat
Población	1,00		
Padr Electoral	0,79	1,00	
Crec Natural	0,71	0,71	1,00

Cuadro A 3.6

Matriz de correlación de las razones de cambio en las proporciones del tamaño poblacional y de tres indicadores sintomáticos, a nivel distrital, Período 73-84

Variable	Pob	P Elec	C Nat	M Esc
Población	1,00			
Padr Electoral	0,80	1,00		
Crec Natural	0,59	0,65	1,00	
Matr Escolar	0,52	0,54	0,31	1,00

ANEXO 4

Cuadro A 4.1
Resultados en los AEP y en los EP de las estimaciones poblacionales
a nivel provincial, por el método de Razón Censal,
según indicadores sintomáticos utilizados, 1984

Variable Sintomática	prom. de AEP	prom de EP	Desv est de EP	Valor mín de EP	Valor máx de EP	Porc de AEP sup a 10%	Porc de AEP inf. a 5%
P Elec	1,2	0,3	1,8	-1,9	2,7	0	100
P Elec - C Nat	1,3	-0,5	1,9	-1,6	3,7	0	100
P Elec - Nac	1,5	0,2	2,1	-1,8	3,6	0	100
Nac - P.Elec - M.Esc	2,4	0,4	3,6	-5,5	6,5	0	71
P Elec - M Esc - Def	2,7	-0,8	3,3	-4,9	3,4	0	100
C.Nat - P.Elec - M.Esc	2,7	0,6	4,1	-5,6	7,8	0	71
Todos	2,8	-1,0	3,4	-5,8	3,4	0	86
P Elec - M Esc - Def - Nac	2,7	-1,0	3,4	-5,4	3,6	0	71
C. Nat	3,9	-1,1	4,8	-7,4	4,5	0	71
P Elec - M ESC	3,9	1,4	5,4	-4,7	11,5	14	86
Nac	4,5	-1,7	4,9	-7,2	4,4	0	57
M Esc	6,9	2,5	10,0	-8,4	22,3	14	57
Def	10,1	-5,1	13,6	-24,0	12,7	43	43

Cuadro A 4.2
Resultados en los AEP y en los EP de las estimaciones poblacionales
a nivel provincial, por el método de Diferencia de Tasas,
según indicadores sintomáticos utilizados, 1984

Variable Sintomática	prom. de AEP	prom de EP	Desv est de EP	Valor mín de EP	Valor máx de EP	Porc de AEP sup a 10%	Porc de AEP inf. a 5%
P Elec	1,4	-0,7	1,9	-2,7	2,0	0	100
P Elec - C Nat	1,7	-0,6	2,2	-1,6	4,2	0	100
P Elec - Nac	2,4	-0,5	2,9	-1,8	4,1	0	100
Nac - P.Elec - M.Esc	3,0	-0,2	3,9	-6,4	5,4	0	71
C.Nat - P.Elec - M.Esc	3,0	-0,0	4,2	-6,5	7,1	0	71
P Elec - M ESC	3,2	0,5	4,9	-6,1	9,6	0	71
C. Nat	3,4	-0,9	4,2	-7,3	3,0	0	71
Todos	3,7	-1,4	4,2	-6,5	4,5	0	71
P Elec - M Esc - Def	4,0	-1,6	4,8	-8,2	5,4	0	57
P Elec - M Esc - Def - Nac	4,0	-1,5	4,6	-7,4	5,0	0	57
Nac	4,2	-1,6	4,6	-7,2	4,0	0	57
M Esc	6,7	1,7	10,3	-10,5	22,0	29	43
Def	11,3	-5,6	15,0	-25,8	12,7	43	29

Cuadro A 4.3
Resultados en los AEP y en los EP de las estimaciones poblacionales
a nivel provincial, por el método de Razón Correlación,
según indicadores sintomáticos utilizados, 1984

Variable Sintomática	prom. de AEP	prom de EP	Desv est de EP	Valor mín de EP	Valor máx de EP	Porc de AEP sup a 10%	Porc de AEP inf. a 5%
P Elec	1,8	0,2	2,4	-2,8	4,1	0	100
C.Nat - P.Elec - M.Esc	2,4	0,5	2,8	-2,8	4,0	0	100
P Elec - C Nat	2,8	-0,2	3,6	-4,2	6,0	0	86
P Elec - M Esc - Def	2,7	-0,9	3,3	-5,4	3,5	0	86
P Elec - M ESc	2,7	0,8	3,2	-2,8	5,6	0	86
P Elec - Nac	3,6	-0,8	4,6	-6,4	6,8	0	71
Nac	4,2	-1,0	5,4	-8,0	7,8	0	57
C. Nat	4,6	-0,7	5,9	-7,3	9,4	0	57
P Elec - M Esc - Def - Nac	4,9	-2,4	6,2	-11,4	6,0	14	43
M Esc	5,8	2,8	7,7	-6,3	16,1	14	43
P.Elec - Nac - M.Esc	6,5	-2,2	9,1	-18,1	10,0	29	57
Def	7,6	-3,4	9,6	-18,3	8,2	29	43
Todos	8,7	-3,6	11,0	-20,8	9,1	29	29

Cuadro A 4.4
Resultados en los AEP y en los EP de las estimaciones poblacionales
a nivel provincial, por el método de Tasa Correlación,
según indicadores sintomáticos utilizados, 1984

Variable Sintomática	prom. de AEP	prom de EP	Desv est de EP	Valor mín de EP	Valor máx de EP	Porc de AEP sup a 10%	Porc de AEP inf. a 5%
P Elec	1,6	0,2	2,3	-2,6	3,7	0	100
P Elec - C Nat	2,7	-0,2	3,4	-3,8	5,7	0	86
C.Nat - P.Elec - M.Esc	2,9	0,9	3,7	-3,2	7,1	0	100
P Elec - M Esc - Def	3,0	1,1	3,9	-3,2	8,2	0	86
P Elec - M ESc	3,0	1,1	4,0	-3,2	8,3	0	86
P Elec - Nac	3,2	-0,9	4,0	-6,2	5,4	0	71
Nac	3,7	-1,2	4,8	-7,6	6,1	0	71
C. Nat	4,6	-0,7	5,9	-7,2	9,5	0	57
M Esc	6,7	2,4	9,5	-7,8	20,9	14	43
Def	7,6	-3,5	9,7	-18,9	8,6	29	43
Todos	7,9	-3,4	10,3	-19,7	12,7	29	29
P Elec - M Esc - Def - Nac	8,5	-4,0	11,5	-24,6	8,1	14	14
P.Elec - Nac - M.Esc	9,4	-3,9	12,8	-28,3	9,6	29	43

Cuadro A 4.5
Resultados en los AEP y en los EP de las estimaciones poblacionales
a nivel cantonal, por el método de Razón Censal
según indicadores sintomáticos utilizados, 1984

Variable Sintomática	prom. de AEP	prom de EP	Desv est de EP	Valor mín de EP	Valor máx de EP	Porc de AEP sup a 10%	Porc de AEP inf. a 5%
P Elec - Nac	4,7	-0,5	6,3	-21,6	16,7	10	66
P.Elec - Nac - M.Esc	4,8	-1,3	5,9	-12,8	17,5	9	56
P Elec	5,0	-0,9	6,5	-21,2	15,6	13	62
P Elec - C Nat	5,0	-0,5	6,8	-23,3	17,7	13	63
P Elec - M Esc - Def	5,1	-0,9	6,7	-14,0	19,9	13	57
P Elec - M Esc - Def - Nac	5,1	-0,7	6,5	-14,9	17,3	12	56
P.Elec - M.Esc - C.Nat	5,2	-1,2	6,2	-13,7	17,0	12	56
Todos	5,4	-0,6	7,0	-18,3	17,1	18	57
P Elec - M ESc	6,9	-1,8	8,5	-18,1	19,7	28	43
Nac	8,8	-0,1	12,2	-34,0	33,7	34	43
C. Nat	9,5	-0,1	13,6	-37,4	44,2	32	37
M Esc	11,3	-2,8	14,0	-33,8	30,1	51	31
Def	15,9	1,0	19,1	32,0	43,3	68	15

Cuadro A 4.6
Resultados en los AEP y en los EP de las estimaciones poblacionales
a nivel cantonal, por el método de Diferencia de Tasas,
según indicadores sintomáticos utilizados, 1984

Variable Sintomática	prom. de AEP	prom de EP	Desv est de EP	Valor mín de EP	Valor máx de EP	Porc de AEP sup a 10%	Porc de AEP inf. a 5%
P Elec	4,3	-1,0	5,5	-16,6	14,1	6	66
P Elec - C Nat	4,6	-0,9	6,1	-21,7	16,0	10	68
P Elec - Nac	4,7	-0,7	6,2	-20,7	16,4	10	65
P.Elec - Nac - M.Esc	5,4	-1,6	6,6	-14,9	17,2	15	57
P.Elec - M.Esc - C.Nat	5,8	-1,7	6,8	-15,7	16,7	13	53
Todos	6,1	-0,7	7,8	-20,3	18,1	21	53
P Elec - M Esc - Def - Nac	6,4	-0,7	7,9	-17,9	23,1	21	51
P Elec - M Esc - Def	6,8	-0,9	8,5	-16,1	29,4	24	40
P Elec - M ESc	7,3	-2,2	8,7	-18,3	20,9	25	37
Nac	8,3	-0,4	11,1	31,5	33,7	32	44
C. Nat	8,3	-0,8	11,4	-35,2	30,6	31	43
M Esc	13,1	-3,4	16,0	-38,1	36,4	59	22
Def	17,9	1,9	21,5	-34,6	49,5	72	16

Cuadro A 4.7
Resultados en los AEP y en los EP de las estimaciones poblacionales
a nivel cantonal, por el método de Razón Correlación
según indicadores sintomáticos utilizados, 1984

Variable Sintomática	prom. de AEP	prom de EP	Desv est de EP	Valor mín de EP	Valor máx de EP	Porc de AEP sup a 10%	Porc de AEP inf. a 5%
P.Elec - Nac - M.Esc	3,8	-0,4	4,7	-8,8	13,5	3	72
P Elec - M Esc - Def - Nac	3,9	-0,5	5,0	-10,7	14,7	4	72
P Elec - Nac	4,1	0,1	4,9	-8,4	12,1	3	72
P Elec - C Nat	4,1	0,0	5,0	-8,6	13,0	3	66
P.Elec - M.Esc - C.Nat	4,2	-0,6	5,2	-10,5	13,5	6	63
Todos	4,5	-1,0	5,6	-12,3	17,8	6	65
P Elec - M Esc - Def	4,6	-0,9	5,8	-12,0	15,7	10	66
P Elec	5,0	-0,4	6,4	-18,5	13,9	13	59
P Elec - M ESC	5,6	-1,2	6,8	-16,6	15,1	18	54
Nac	6,7	1,1	7,9	-16,7	21,9	21	51
C. Nat	6,8	1,4	8,6	-19,9	23,6	15	43
M Esc	9,5	-1,5	11,5	-31,7	25,2	41	32
Def	10,5	1,9	12,8	-26,8	18,8	37	29

Cuadro A 4.8
Resultados en los AEP y en los EP de las estimaciones poblacionales
a nivel cantonal, por el método de Tasa Correlación
según indicadores sintomáticos utilizados, 1984

Variable Sintomática	prom. de AEP	prom de EP	Desv est de EP	Valor mín de EP	Valor máx de EP	Porc de AEP sup a 10%	Porc de AEP inf. a 5%
P.Elec - Nac - M.Esc	3,8	-0,7	4,6	-9,2	14,2	1	72
P Elec - Nac	4,0	-0,1	4,8	-11,4	12,5	3	72
P Elec - C Nat	4,0	-0,2	4,8	-10,6	12,3	4	63
P Elec - M Esc - Def - Nac	4,2	-0,8	5,3	-10,9	15,7	6	68
P Elec - M Esc - Def	4,7	-1,1	6,1	-13,7	17,0	10	62
P.Elec - M.Esc - C.Nat	4,9	-0,6	6,1	-15,9	16,2	10	62
P Elec	4,9	-0,6	6,3	-19,4	13,8	12	56
Todos	5,7	-1,3	7,2	-16,4	19,8	13	54
P Elec - M ESC	5,7	-1,4	7,1	-17,3	16,4	19	51
Nac	6,4	0,9	7,9	-18,0	21,3	24	51
C. Nat	6,7	1,0	8,4	-20,0	22,0	24	44
M Esc	9,6	-1,7	11,6	-31,3	25,3	43	37
Def	10,4	1,6	13,0	-26,2	30,2	44	28

ANEXO 5

Cuadro A 5.1

Resultados en los AEP de las estimaciones poblacionales a nivel cantonal,
por indicador sintomático, según método utilizado, 1984

(Para los 81 cantones existentes en 1984)

Indicadores sintomáticos utilizados	Método	
	Razón Censal	Diferencia de Tasas
Padrón Electoral		
Promedio de AEP	5,9	4,8
Valor Máximo de AEP	21,5	19,7
Porcentaje de AEP inf a 5%	53,1	63,0
Porcentaje de AEP sup a 10%	17,3	8,6
Padrón electoral y Crecimiento Natural		
Promedio de AEP	5,1	5,3
Valor Máximo de AEP	23,6	21,7
Porcentaje de AEP inf a 5%	63,0	60,5
Porcentaje de AEP sup a 10%	16,1	17,3
Padrón electoral y Nacimientos		
Promedio de AEP	4,8	5,3
Valor Máximo de AEP	21,8	20,7
Porcentaje de AEP inf a 5%	65,4	59,3
Porcentaje de AEP sup a 10%	12,3	12,3
Padrón electoral , Nacimientos y Matrícula Escolar		
Promedio de AEP	5,3	5,4
Valor Máximo de AEP	17,2	17,1
Porcentaje de AEP inf a 5%	50,6	59,3
Porcentaje de AEP sup a 10%	11,1	13,6

Cuadro A 5.2

Resultados en los AEP de las estimaciones poblacionales a nivel distrital,
por indicador sintomático, según método utilizado, 1984

(Para los 404 distritos existentes en 1984)

Indicadores sintomáticos utilizados	Método	
	Razón Censal	Diferencia de Tasas
Padrón Electoral		
Promedio de AEP	13,8	11,5
Valor Máximo de AEP	82,1	44,2
Porcentaje de AEP inf a 5%	23,0	28,0
Porcentaje de AEP sup a 10%	51,0	46,0
Crecimiento Natural		
Promedio de AEP	26,5	23,4
Valor Máximo de AEP	181,1	111,5
Porcentaje de AEP inf a 5%	19,1	13,1
Porcentaje de AEP sup a 10%	69,8	70,0
Matricula Escolar		
Promedio de AEP	25,0	35,3
Valor Máximo de AEP	117	588
Porcentaje de AEP inf a 5%	16,1	11,0
Porcentaje de AEP sup a 10%	65,8	76,9
Padrón electoral y Crecimiento Natural		
Promedio de AEP	16,9	15,6
Valor Máximo de AEP	85,1	69,3
Porcentaje de AEP inf a 5%	19,1	19,1
Porcentaje de AEP sup a 10%	65,8	65,8

10.6 Anexo 6 Método de viviendas

Este ejemplo va dirigido a obtener la estimación de la población del cantón Central de San José para el año 1996. La información para aplicar el procedimiento la suministra el Cuadro A 1.5 del Anexo 1. Para este cantón se tiene lo siguiente (1984):

Número de viviendas simples (V_{1984}): 55019

Número de viviendas con servicio eléctrico (CE_{1984}): 54666

Promedio de personas por vivienda (PPV_{1984}): 4,352

Número de personas en viviendas colectivas (RC_{1984}): 2039

Además, se tiene que el número de clientes residenciales del servicio eléctrico al 30 de junio de 1996 (CE_{1996}) 73981.

Suponiendo que la información del censo sobre las viviendas con servicio eléctrico, representa a los clientes residenciales de dicho servicio en 1984, para estimar el número de viviendas se prefiere utilizar la fórmula:

$$\hat{V}_{1996} = \left(\frac{V_{1984}}{CE_{1984}} \right) CE_{1996}$$

La estimación del número de viviendas del cantón Central de San José al 30 de junio de 1996 es: $\hat{V}_{1996} = \left(\frac{55019}{54666} \right) \cdot 73981 \cong 74459$.

Para estimar el valor de PPV_{1996} se utiliza la fórmula: $\hat{PPV}_{1996} = D_{1996} PPV_{1984}$ donde D_t corresponde a la razón de cambio en el PPV desde el último censo. Para ello, se utiliza la información de la Encuesta de Hogares de 1995, donde se estima, para la Región Central, el número de personas por vivienda en 4,189. Dicho valor en 1984 se calculó en 4,689. Por lo tanto, el número promedio de personas por vivienda para San José en 1996, se estima en:

$$\hat{PPV}_{1996} = \frac{4,189}{4,689} \cdot 4,352 \cong 3,888.$$

Por otro lado, se puede suponer que el número de personas en residencias colectivas no ha sufrido cambios importantes en este período, por lo que

$$RC_{1996} \cong 2039.$$

Finalmente la población estimada para el cantón Central de San José, para 1996, por el método de viviendas viene dada por:

$$P(\text{San José})_{1996} = V_{1996} \cdot PPV_{1996} + RC_{1996} = 74459 \cdot 3,888 + 2039 \cong 291535$$