

5 OCT. 1977

Griffith Feeney

East-West Population Institute
The East-West Center
Honolulu, Hawaii

ESTIMACION DE PARAMETROS DEMOGRAFICOS
A PARTIR DE INFORMACION CENSAL
Y DE REGISTROS

Este artículo, del original inglés *Estimation of Demographic Parameters from Census and Vital Registration Data*, fue presentado a la XVIII Conferencia General de la Unión Internacional para el Estudio Científico de la Población, realizada entre el 8 y el 13 de agosto de 1977 en México, y ha sido traducido y publicado en CELADE con la autorización del autor.



Serie D, N° 93

Santiago de Chile

Agosto de 1977

BIBLIOTECA "GIORGIO MORTARA"
CENTRO LATINOAMERICANO
DE DEMOGRAFIA

Las opiniones y datos que figuran en este trabajo son responsabilidad del autor, sin que el Centro Latinoamericano de Demografía (CELADE) sea necesariamente partícipe de ellos.

I N D I C E

	<u>Página</u>
1. Introducción	1
2. Modelos para representar la fecundidad, por edad	2
3. Información censal sobre fecundidad acumulada.....	4
4. Fecundidad constante	6
5. Estimaciones a partir de dos censos	9
6. Utilización de información del registro de nacimientos	11
7. Método de los hijos sobrevivientes para la mortalidad constante.	12
8. Tendencias de la mortalidad a partir de información de hijos sobrevivientes	16
9. Conclusiones	22
REFERENCIAS	25

Indice de cuadros y gráficos

Cuadros

1	Información de hijos sobrevivientes para Borneo del Norte: censo del 11 de agosto de 1960	18
---	---	----

Gráficos

1	Tendencias lineales de la mortalidad infantil coherentes con la proporción de hijos sobrevivientes entre los hijos nacidos vivos de mujeres con 15-19 años, según el censo de 1960 de Borneo del Norte	20
---	--	----

REPORT

The following information was obtained from the records of the Department of the Interior, Bureau of Land Management, regarding the land parcels described herein. The information is presented in the form of a table.

Parcel No.	Section	Township	Range	County	State
1	36	10N	10E	Lincoln	Nebraska
2	36	10N	10E	Lincoln	Nebraska
3	36	10N	10E	Lincoln	Nebraska
4	36	10N	10E	Lincoln	Nebraska
5	36	10N	10E	Lincoln	Nebraska
6	36	10N	10E	Lincoln	Nebraska
7	36	10N	10E	Lincoln	Nebraska
8	36	10N	10E	Lincoln	Nebraska
9	36	10N	10E	Lincoln	Nebraska
10	36	10N	10E	Lincoln	Nebraska

APPENDIX

The following information was obtained from the records of the Department of the Interior, Bureau of Land Management, regarding the land parcels described herein. The information is presented in the form of a table.

Parcel No.	Section	Township	Range	County	State
1	36	10N	10E	Lincoln	Nebraska
2	36	10N	10E	Lincoln	Nebraska
3	36	10N	10E	Lincoln	Nebraska
4	36	10N	10E	Lincoln	Nebraska
5	36	10N	10E	Lincoln	Nebraska
6	36	10N	10E	Lincoln	Nebraska
7	36	10N	10E	Lincoln	Nebraska
8	36	10N	10E	Lincoln	Nebraska
9	36	10N	10E	Lincoln	Nebraska
10	36	10N	10E	Lincoln	Nebraska

RESUMEN

Se presenta en este documento una breve reseña de los métodos para estimar tasas por edad de fecundidad y de mortalidad a partir de información de un censo o de una encuesta relativos al número de hijos nacidos vivos tenidos por las mujeres y al número de hijos sobrevivientes a la fecha del censo o encuesta. Se señala que la aparente diversidad de estos métodos esconde una notable similitud de enfoques, a tal punto que cada uno de estos métodos puede considerarse como la realización, dentro de un contexto particular, de un enfoque estadístico singular a los problemas de estimación. Se sugiere que el reconocimiento explícito de este enfoque puede resultar útil en los desarrollos adicionales que puedan hacerse de los métodos existentes o en los de nuevos métodos.

SUMMARY

This paper presents a brief survey of methods for estimating age-specific birth and death rates utilizing census or survey data on numbers of children born to women and number of children surviving at the time of the census or survey. It is pointed out that the apparent diversity of these methods conceals a remarkable similarity of approach, so much so that each of the methods may be regarded as the realization in a particular context of a single strategic approach to estimation problems. It is suggested that explicit recognition of this approach may prove useful in further development of existing methods and in the development of new methods.

Faint, illegible text, possibly bleed-through from the reverse side of the page.

Faint, illegible text, possibly bleed-through from the reverse side of the page.

1. Introducción

En un documento presentado a la conferencia de población de esta asociación en Londres en 1969, el profesor W. Brass observó que irónicamente, una de las causas de los resultados poco satisfactorios que se tienen a veces al utilizar las técnicas de análisis aplicables a información limitada y defectuosa es el rápido mejoramiento en cuanto a cobertura y cabalidad de la información disponible. "Las técnicas necesariamente rígidas para imponer orden en información muy restringida y de dudosa calidad, son demasiado inflexibles cuando las informaciones son mejores. Mayores progresos tendientes a mejorar la precisión de las medidas pueden hacerse sólo mediante análisis que tomen en cuenta, en mayor grado, las peculiaridades de una población en particular ...". (1969:185). En este documento se examinan estrategias y tácticas para alcanzar esa flexibilidad analítica, refiriéndose, en particular, a los desarrollos recientes de los análisis de la información censal relativa a la fecundidad y a la sobrevivencia de los hijos. La información censal consiste en las tabulaciones de las respuestas de la población femenina a las preguntas sobre el número de hijos nacidos vivos que han tenido y sobre el número de esos niños que están vivos al momento del censo. El propósito es convertir esta información en tasas de fecundidad por edad o en estadísticas resúmenes de ellas, y en tablas de vida. Tanto las tasas de fecundidad como las tablas de vida pueden calcularse directamente a partir de informaciones de los registros de nacimientos por edad de la madre y de defunciones por edad respectivamente, de manera que lo que se presenta en este documento es de interés para poblaciones donde las estadísticas de registro no existen o son muy incompletas. En las secciones 2 a la 7 se da una visión de los procedimientos desarrollados en la década pasada. En la sección 8 se presenta el trabajo que personalmente he desarrollado en relación a estimaciones de tendencias de la mortalidad a partir de información de sobrevivencia de hijos.

2. Modelos para representar la fecundidad, por edad

Un modelo sencillo para describir la fecundidad por edad es:

$$m_x = TGF \cdot m'_x \quad (2.1)$$

donde m'_x son constantes dadas con $\sum m'_x = 1$ y TGF es un parámetro de la fecundidad total $\sum m_x$.

Este modelo es la base para el cálculo de lo que se llama tasa de fecundidad estandarizada; indirectamente puede utilizarse también cuando se cuenta con información que provee de indicios sobre la forma de la fecundidad, pero no sobre su nivel. (Véanse los ejemplos que se dan en la sección siguiente). Otro modelo, presentado por Brass, se llama "Fecundidad polinomial"

$$m(a) = \begin{cases} c(a-s)(s+33-a)^2 & \text{para } s < a < s+33 \\ 0 & \text{cualquiera} \end{cases} \quad (2.2)$$

(Brass 1975:18-23).

La tasa global de fecundidad correspondiente a este polinomio es:

TGF = 98826,75c hijos por mujer y la edad media de la fecundidad es:

$$MF = s + 13,2 \text{ años,}$$

de aquí que el modelo puede expresarse fácilmente con parámetros representativos de la tasa global de fecundidad y de la edad media de la fecundidad. Los cálculos numéricos con este modelo se facilitan porque $0,25(s+33-a)^4 - 11(s+33-a)^3$ es una antiderivación del polinomio en (2.2), de donde:

$$\int_0^x m(a) da = c \cdot \left[\frac{(s+33-x)^4}{4} - 11(s+33-x)^3 + 98826,75 \right], s < x < s+33 \quad (2.3)$$

En particular, si las mujeres se distribuyen uniformemente por edad dentro del grupo de edades x a $x+n$,

$${}_n m_x = \int_x^{x+n} m(a) da = \int_0^{x+n} m(a) da - \int_0^x m(a) da, \quad (2.4)$$

donde si se hacen los reemplazos correspondientes utilizando (2.3), ${}_n m_x$ queda expresada en términos de los parámetros \underline{c} y \underline{s} . El modelo (2.4), por lo tanto, conduce a la parametrización

$$m_x = TGF \cdot \phi_x(MF), \quad (2.5)$$

de las tasas de fecundidad por edad m_x donde $\phi_x = {}_1 m_x$ definida por (2.4) y (2.3), esta última con $\underline{c} = 98826,75^{-1} x$ y con $\underline{s} = MF - 13,2$.

Puede observarse que aunque (2.3) es válida solamente para x comprendida en el rango que se indica, sustituyendo $\min\{\max\{x, s\}, s+33\}$ por x en el lado derecho del signo igual, la fórmula queda válida para todo x .

Coale y Trussell han propuesto recientemente un modelo más elaborado de fecundidad:

$$m(a) = TGF \cdot G(a; a_0, k) \cdot r(a; m) \quad (2.6)$$

donde:

$$G(a; a_0, k) =$$

$$\frac{0,19465}{K} \int_0^x \exp \left[\frac{-0,174(a-a_0-6,6k)}{K} - \exp \left[\frac{-0,2881(a-a_0-6,06k)}{K} \right] \right] da \quad (2.7)$$

y

$$r(a; m) = M n(a) \exp[mv(a)], \quad a \geq 0 \quad (2.8)$$

donde $n(a)$ y $v(a)$ son patrones que se definen empíricamente y M es una constante que se elige de forma tal que

$$\int_0^{\infty} m(a) da = TGF$$

(Coale y Trussell 1974). El patrón $n(a)$ representa la fecundidad matrimonial y el patrón $v(a)$ conjuntamente con el parámetro m representa el aminoramiento de la fecundidad debido al control de la natalidad dentro del matrimonio. La función $G(a; a_0, k)$ es un modelo de la distribución por edad de las mujeres al primer matrimonio (Coale y McNeil 1972). Los parámetros a_0 y K son esencialmente parámetros escalares y de ubicación, con a_0 que representa la edad más joven al matrimonio y K representa la intensidad con que ocurren los matrimonios dentro de un intervalo mayor o menor de edades.

3. Información censal sobre fecundidad acumulada

En los censos de población frecuentemente se pregunta a las mujeres acerca del número de hijos nacidos vivos que han tenido durante toda su vida. Esta información se tabula generalmente por edad de la madre, mostrando ya sea el número total de hijos nacidos vivos de mujeres por grupos de edades, o la distribución de mujeres en cada grupo de edades según el número de hijos nacidos vivos. El número total de hijos nacidos vivos puede calcularse como suma del número de mujeres que han tenido exactamente un hijo, más dos veces el número de mujeres que han tenido exactamente dos hijos y así sucesivamente, pero cabe hacer notar que este procedimiento es aplicable solamente si la distribución se presenta en forma completa. En la práctica las tabulaciones aparecen frecuentemente truncadas para economizar espacio en las publicaciones censales. Véase Feeney 1967a para mayores detalles. Son más habituales las tabulaciones por grupos quinquenales de edades, pero los desarrollos recientes en materia de computación y análisis conducen cada vez más a la preparación de tabulaciones por edades simples.

Sea N_x el número de mujeres con edad x en años cumplidos al momento del censo y supongamos que todas estas mujeres tienen exactamente la edad $x+1/2$. El número de nacimientos ocurridos a estas mujeres en el t ésimo año anterior al censo puede expresarse entonces mediante

$$N_x m_{x-t+1}(t)$$

donde $m_x(t)$ simboliza la tasa de fecundidad para el intervalo anual de edades centrado en torno a la edad x para el t ésimo año anterior al censo. El total de nacimientos de estas mujeres equivale a la suma de estos términos para todo valor relevante de t .

$$B_x = \sum_{t=1}^{x-9} N_x m_{x-5+1}(t), \quad x=10, \dots, 59 \quad (3.1)$$

donde B_x simboliza a todos los hijos nacidos vivos de mujeres con edad cumplida x al momento del censo; la edad límite menor y mayor del período reproductivo se fija en los 10 y 60 años, respectivamente.

Las fórmulas que siguen son más fáciles de escribir que de leer. La clave es imaginar los términos de la suma en (3.1) escritos en una forma en que las filas correspondan a las edades y las columnas a los años anteriores al censo. Todas las fórmulas se derivan multiplicando todos los términos de una fila por una constante o por las primeras filas combinadas por suma, lo que corresponde a agrupar las edades y multiplicando las filas resultantes por una constante.

Dividiendo ambos lados de la relación (3.1) por N_x , se tiene:

$$b_x = \sum_{t=1}^{x-9} m_{x-t+1}(t), \quad x=10, \dots, 59, \quad (3.2)$$

donde b_x simboliza el número medio de hijos nacidos vivos tenidos por mujeres con edad cumplida x . Estas ecuaciones son tautológicas en el entendimiento que las tasas de fecundidad por edad se aplican solamente a las mujeres sobrevivientes a la fecha del censo; de otra manera ellas implicarían el supuesto que las mujeres no sobrevivientes a la fecha del censo, tienen las mismas tasas de fecundidad por edad que las mujeres sobrevivientes.

Las ecuaciones para grupos de edades se obtienen al introducir en (3.1) las sumas de edades simples que componen los grupos de edades y dividiendo ambos lados del resultado por el número de mujeres en esos grupos de edades.

$${}_n b_x = \frac{\sum_{y=x}^{x+n-1} \sum_{t=1}^{y-9} N_y m_{y-t+1}(t)}{\sum_{y=x}^{x+n-1} N_y}, \quad n=1, \dots, 60-x \quad (3.3)$$

$$x=10, \dots, 59$$

donde ${}_n b_x$ simboliza el promedio de hijos nacidos vivos de mujeres con edad cumplida x a $x+n-1$ años. En el supuesto que las mujeres se distribuyen uniformemente en este intervalo de edad:

$$N_y = N, \quad y=x, x+1, \dots, x+n-1 \quad (3.4)$$

lo que si se sustituye en (3.3) resulta:

$${}_n b_x = \frac{1}{n} \sum_{y=x}^{x+n-1} \sum_{t=1}^{y-9} m_{y-t+1}(t) \quad (3.5)$$

4. Fecundidad constante

Si se cuenta con información censal de b_x , la relación (3.2) define un sistema de 50 ecuaciones en $1+2+\dots+50 = 820 m_x(t)$ incógnitas. Si la fecundidad ha sido constante, sin embargo,

$$m_x(t) = m_x(1), \quad t=1, \dots, 60-x, \quad x=10, \dots, 59 \quad (4.1)$$

y la relación (3.2) se convierte en:

$$b_x = \sum_{t=1}^{x-9} m_{x-t+1}, \quad x=10, \dots, 59 \quad (4.2)$$

Esta relación define un sistema de 50 ecuaciones con 50 incógnitas m_x , $x=10, 11, \dots, 59$, y tiene la siguiente solución simple:

$$m_x = b_x - b_{x-1}, \quad x=10, \dots, 59 \quad (4.3)$$

Este método fue desarrollado por Mortara (1949:40-50).

Supongamos a continuación que se cuenta con tasas de fecundidad por edad calculadas con datos de registros y que, aunque los nacimientos están subregistrados, el grado de subregistro es independiente de la edad de la madre. Supongamos, alternativamente, que las tasas de fecundidad por edad se han calculado a partir de información obtenida de un censo o encuesta donde se ha preguntado a las mujeres si han tenido o no hijos nacidos vivos durante el año -o algún período determinado- anterior al censo, y si es así, cuántos nacimientos; y supongamos que los errores en la declaración del número de nacimientos, debidos a falta de memoria, o a inexactitud del período de referencia, o a otra causa, son independientes de la edad de las mujeres. En cualquier caso, los datos proporcionan la forma de la fecundidad pero no el nivel, esto es, es aplicable la parametrización (2.1),

$$m_x = TGF \cdot m'_x, \quad x=10, \dots, 59.$$

Incorporando estas expresiones de m_x en la relación (4.2) se obtiene:

$$b_x = TGF \sum_{t=1}^{x-9} m'_{x-t+1}, \quad x=10, \dots, 59 \quad (4.4)$$

Este sistema de 40 ecuaciones con la sola incógnita de la TGF no tendrá una solución, en general, pero las ecuaciones pueden resolverse individualmente para obtener:

$$TGF_x = b_x \div \left[\sum_{t=1}^{x-9} m'_{x-t+1} \right], \quad x = 10, \dots, 59 \quad (4.5)$$

La dispersión y estructura de estas soluciones se pueden examinar representándolas gráficamente en relación a x y determinar una TGF si las circunstancias lo permiten. La idea de este método se debe a Brass que lo ha desarrollado para aplicarlo a información dispuesta en grupos de edades, como se indica más adelante.

En el supuesto de fecundidad constante (4.1) la doble suma en la relación (3.5) se reduce a:

$${}_n b_x = \sum_{i=10}^{x-1} m_i + \sum_{i=0}^{n-1} \frac{n-i}{n} m_{x+i}, \quad n=1, \dots, 60-x \quad (4.6)$$

$$x=10, \dots, 59$$

Como puede verse, simplemente escribiendo los términos de la suma en una forma ordenada. Esta fórmula se aplica para valores de x y n más allá de los rangos que se indican, haciendo $m_x = 0$ para $x < 10$ y $x > 60$. Sustituyendo: $m_x = \text{TGF} \cdot m'_x$ por m_x , se tiene:

$${}_n b_x = \text{TGF} \cdot \left\{ \sum_{i=10}^{x-1} m'_i + \sum_{i=0}^{n-1} \frac{n-i}{n} m'_{x+i} \right\}, \quad n=1, \dots, 60-x \quad (4.7)$$

$$x=10, \dots, 59$$

que puede tratarse en forma similar a la relación (4.4).

Puede utilizarse la ecuación (4.6) cuando la información censal relativa a hijos nacidos vivos no está disponible por edades simples, pero no se requiere que m_x esté disponible por edades simples. Si no es éste el caso, pueden interpolarse valores para las edades simples a partir de la información agrupada y aplicarse luego las ecuaciones a los valores interpolados. Puede observarse que, alternativamente, la suma del lado izquierdo de la ecuación (4.7) puede calcularse mediante:

$$\sum_{(x,n) \in I} m_x \quad (4.8)$$

donde I es un conjunto indicador que define los grupos que cubren todas las edades comprendidas entre las edades exactas 10 y x años. La suma de la derecha en la relación (4.7) es indeterminada, pero puede expresarse como $k m_x$ donde

$$k = n - n \left[\sum_{i=0}^{n-1} i \cdot \left\{ m_{x+i} \div \sum_{i=0}^{n-1} m_{x+i} \right\} \right] \quad (4.9)$$

Si se supone que los valores de m_x se conforman al modelo de fecundidad poli-
nomial (2.5), el término entre llaves ({}), depende solamente del valor de la
MF, y el valor de este "multiplicador" puede tabularse para varios grupos de
edades y varios valores de la MF (Brass 1975:18-23). Los valores de la MF
pueden estimarse incorporando la parametrización (2.5) en la relación (4.7)
que da:

$$\frac{b}{n^x} = TGF \cdot \phi_x(MF), \quad \begin{array}{l} n=1, \dots, 60-x \\ x=10, \dots, 59 \end{array} \quad (4.10)$$

donde:

$$\phi_x(MF) = \sum_{i=0}^{x-1} \phi_x(MF) + \sum_{i=0}^{n-1} \frac{n-1}{n} \phi_{x+i}(MF), \quad (4.11)$$

$$\begin{array}{l} n=1, \dots, 60-x \\ x=10, \dots, 59 \end{array}$$

La razón de estas dos ecuaciones tiene la forma:

$$\frac{\frac{b}{n^x}}{\frac{b}{n^{x-n}}} = \frac{\phi_x(MF)}{\phi_{x-n}(MF)} \quad (4.12)$$

que puede resolverse rápidamente mediante interpolación entre los valores ta-
bulados de la expresión de la derecha, que depende solamente de la MF y no de
las informaciones. El valor obtenido de la MF dependerá hasta cierto punto
de la agrupación de edades que se escoja, como antes; en la práctica es útil
examinar la dispersión de las estimaciones que se obtienen de diferentes a-
grupaciones.

5. Estimaciones a partir de dos censos

Supongamos ahora que la información referente a hijos nacidos vivos es-
tá disponible para dos censos sucesivos con un intervalo intercensal de años

completos. Supongamos que la fórmula (3.2) se refiera al último censo y que la suma en el lado derecho se divida en dos grupos de términos: unos referentes a la fecundidad anterior al primer censo y los otros a la fecundidad del período intercensal.

$$b_x = \begin{cases} \sum_{t=1}^{x-9} m_{x-t+1}(t) & x = 10, \dots, T+9 \\ \sum_{t=1}^T m_{x-t+1}(t) + \sum_{t=T+1}^{x-9} m_{x-t+1}(t) & x = T+10, \dots, 59 \end{cases} \quad (5.5)$$

Si ninguna mujer de las que tuvieron edades dentro del período reproductivo en el último censo muriera en el período intercensal, el último término de la derecha representa el promedio de hijos nacidos vivos a mujeres con edad cumplida $x-t$ años al primer censo. Esto nos lleva a:

$$b_x = \sum_{t=1}^T m_{x-t+1}(t) + b'_{x-T}, \quad x=10, \dots, 59+T \quad (5.6)$$

donde b'_x simboliza al promedio de hijos nacidos vivos en el primer censo. Haciendo $b'_x = 0$ para $x < 10$ y $m_x(t) = 0$ para $x > 60$ para todo t , la relación (5.6) define un sistema de $60 \cdot T$ ecuaciones con $50 \cdot T$ incógnitas $m_x(t)$. Esta ecuación, en general, no puede resolverse directamente a menos que $T=1$, lo que no tiene interés ya que prácticamente no existe caso alguno en que el período intercensal sea 1. Sin embargo, se abren dos opciones. Los valores intercensales de b_x , a intervalos anuales pueden interpolarse a partir de los pares de valores censales, y los diferentes sistemas anuales que se generan pueden resolverse directamente, procedimiento que propuso Ansley Coale siguiendo una estrategia similar a la que desarrolló Norman Ryder en relación con la nupcialidad (Coale: información personal). Se puede ver, alternativamente, que las sumas en (5.6) son tasas de fecundidad para grupos de edades de T años y que la totalidad de estas tasas generadas como rangos de x para el intervalo 10 a $59+T$ años se divida en T grupos que cubren la totalidad y

de forma excluyente el período reproductivo. Mientras que las tasas individuales no pueden compararse rápidamente, se puede calcular para cada subconjunto una tasa global de fecundidad. Este procedimiento es obvio y si no ha sido propuesto en la literatura correspondiente es porque probablemente es muy raro que los censos presenten tabulaciones de hijos nacidos vivos por edades simples de las mujeres. La ventaja, y es realmente una ventaja, de las tabulaciones por edad simple es que el promedio de hijos nacidos vivos a mujeres con x años cumplidos al momento del censo se aproxima estrechamente a la paridez media de la edad exacta $x+1/2$, la edad central del intervalo de edad. La desventaja de los grupos quinquenales convencionales, y es realmente una desventaja, es que la fecundidad de la edad central no corresponde a la fecundidad media del grupo. Arretx presentó un procedimiento ingenioso para determinar la edad exacta a que corresponde la fecundidad dada para un grupo quinquenal (1973).

6. Utilización de información del registro de nacimientos

Supongamos por el momento que está disponible información de los registros de nacimientos, aunque afectada por subregistros, que provee los números de las tasas de fecundidad por edad para un número importante de años anteriores al censo. Los denominadores correspondientes se pueden obtener rejuveneciendo la población femenina por edad. Se necesita para ello estimaciones de la mortalidad adulta, que puede obtenerse de diferentes formas, y cuando la mortalidad es baja, los valores de las relaciones de sobrevivencia inversas son relativamente insensibles a los errores de las estimaciones. La fórmula (3.2) puede escribirse:

$$b_x = \sum_{t=1}^{x-9} m_{x-t+1}(t) \psi_x(t), \quad x=10, \dots, 59 \quad (6.1)$$

donde $m_x(t)$ simboliza la tasa de fecundidad para la edad cumplida x para el t ésimo año anterior al censo calculada con los nacimientos registrados y $\psi_x(t)$ simboliza la razón entre los nacimientos totales y registrados para esa

edad y período de tiempo considerado. Supongamos ahora que el subregistro, medido por $\psi_x(t)$ es independiente de la edad de la madre y ha disminuido linealmente en los años anteriores al censo, esto es:

$$\psi_x(t) = u + r(t-1), t=1, 2, \dots, \quad (6.2)$$

donde u representa el nivel de subregistro durante el año precedente al censo y r simboliza la tasa de descenso de este nivel en los años anteriores. Introduciendo estas expresiones de $\psi_x(t)$ en (6.1) se tiene:

$$b_x = \sum_{t=1}^{x-9} m_{x+t-1}(t) [u+r(t-1)], x=10, \dots, 59 \quad (6.3)$$

o, reordenando los términos

$$b_x = u \left[\sum_{t=1}^{x-9} m_{x+t-1}(t) \right] + r \left[\sum_{t=1}^{x-9} (t-1)m_{x+t-1}(t) \right], x=10, \dots, 59 \quad (6.4)$$

Esta expresión constituye un sistema de 40 ecuaciones con 2 incógnitas u y r y cualquier par de ellas puede resolverse para obtener u y r . Este método fue desarrollado y aplicado a Yugoslavia por Macura (1972:9-10 y continuación).

7. Método de los hijos sobrevivientes para la mortalidad constante

En los censos se pregunta a las mujeres, frecuentemente, sobre el número de hijos nacidos vivos que han tenido y sobre los que están vivos al momento del censo, lo que permite calcular las proporciones de hijos fallecidos entre el total de hijos nacidos vivos tenidos por las mujeres en cada grupo de edades. Desagregando el total de hijos nacidos vivos en cohortes de nacimientos y calculando los sobrevivientes de cada cohorte al momento del censo, se tiene la siguiente relación:

$$Q_i = 1 - \sum_{t=1}^{n(i)} p(t)c_i(t), \quad i=1, 2, \dots, N \quad (7.1)$$

donde el subíndice i especifica el grupo de edades, Q_i simboliza la proporción de hijos fallecidos entre el total de hijos tenidos al momento del censo, $c_i(t)$ la proporción de los nacimientos ocurridos el t ésimo año anterior al censo, entre el total de hijos nacidos vivos, y $p(t)$ la proporción de estos hijos que están vivos al momento del censo.

La proporción $c_i(t)$ puede calcularse directamente a partir de los términos de la suma en (3.1) si se conocen las tasas de fecundidad por edad $m_x(t)$, y estas tasas pueden estimarse suponiendo que la fecundidad es constante y ajustando un modelo de fecundidad a la información disponible relativa a la paridez media de las mujeres por edad, como por ejemplo mediante (4.12). Los valores $p(t)$ pueden expresarse como:

$$p(1) = L_0(1)$$

$$p(2) = L_0(2) \cdot \frac{L_1(1)}{L_0(1)} \quad (7.2)$$

$$p(3) = L_0(3) \cdot \frac{L_1(2)}{L_0(2)} \cdot \frac{L_2(1)}{L_1(1)}$$

donde $L_x(t)$ simboliza los años vividos entre las edades exactas x y $x+1$ de acuerdo a la tabla de vida que expresa la mortalidad en el t ésimo año anterior al censo. Introduciendo estas expresiones en (7.1) se llega a una ecuación para cada grupo de edades con un total de $n(n+1):2$ incógnitas

$$L_x(t), \quad t=1, 2, \dots, N-x, \quad x=0, 1, \dots, N-1.$$

Si se supone que la mortalidad es constante,

$$L_x(t) = L_x, \quad t=1, \dots, N-x \quad (7.3)$$

$$x=0, 1, \dots, N-1$$

los términos de la izquierda en (7.2) se anulan si hacemos $p(t) = L_{t-1}$, $t=1, 2, \dots, N$, lo que reduce el número de incógnitas a N . Sin embargo, esto está lejos de ser una solución ya que N es el límite superior del grupo de edades más viejo menos la edad al comienzo del período reproductivo, por ejemplo 35, si los grupos de edades son: 15-19, ..., 45-49 y el período reproductivo comienza a los 15 años. Introduciendo además el supuesto que el patrón de mortalidad se ajusta a una familia de tablas modelo de vida de un parámetro, la ecuación (7.1) se reduce a:

$$Q_i = 1 - \frac{n(i)}{\sum_{t=1} L_{t-1}(\omega) c_i(t)}, \quad i=1, 2, \dots, N \quad (7.4)$$

donde $L_x(\omega)$ simboliza los años vividos entre las edades exactas x y $x+1$ en la tabla modelo de vida identificada por el parámetro de valor (ω) .

Cada una de estas ecuaciones puede resolverse en forma independiente para ω , produciendo valores, digamos de $\omega_1, \dots, \omega_n$. Cualquier tabla de vida que se desee, puede calcularse a partir de estos valores ω utilizando las ecuaciones de definición de la familia de tablas modelo.

Brass desarrolló un procedimiento ingenioso para derivar estimaciones de la mortalidad que evita resolver en forma directa la relación (7.4). Debido al supuesto de mortalidad constante, todos los hijos nacidos vivos experimentaron la mortalidad representada por una sola tabla de vida, y por el supuesto de la familia de tablas modelo, esta tabla de vida corresponde a la tabla modelo con parámetro ω para un ω dado. Se desprende que $q(x) = q(x;\omega)$, donde $q(x) = 1 - l_x$ simboliza la probabilidad de morir entre el nacimiento y la edad x de acuerdo a la tabla de vida que refleja la mortalidad de la población, y $q(x;\omega)$ simboliza el valor correspondiente en la tabla de vida identificada con el valor del parámetro ω . Se deduce fácilmente de la relación (7.4) que $q(x) = Q \cdot M(x, i, \omega)$ donde el multiplicador M es un conjunto igual a

$$\frac{q(x;\omega)}{1 - \frac{n(i)}{\sum_{t=1} L_{t-1}(\omega) c_i(t)}} \quad (7.5)$$

Esta cantidad depende del grupo de edades, representado por i , de los valores de $c_i(t)$, y del valor de x , todos los cuales se conocen o se estiman. Sin embargo, también depende de w , esto es, del nivel de la mortalidad. De aquí que introduciendo el multiplicador no se avanza obviamente en la estimación de la mortalidad. Pero Brass descubrió que si x se elige adecuadamente en relación a i , M es aproximadamente constante en relación a los cambios de w . De aquí entonces que el valor del multiplicador puede calcularse aproximadamente sin conocer el nivel de la mortalidad. Procedió luego a tabular estos multiplicadores para valores de $c_i(t)$ obtenidos ajustando el modelo (2.2) a la fecundidad observada, que se supuso constante. (Brass y otros 1968:105-120).

Para ajustar el modelo (2.2) Brass utilizó la razón entre la paridez media de las mujeres de 15-19 años y la de las mujeres de 20-24 años mediante la ecuación (4.12); los multiplicadores estimados fueron entonces funciones de esta razón, llamados corrientemente como "razón de paridez media". Sullivan estudió esta relación calculando valores exactos tanto del multiplicador (7.4) como de la razón de paridez media utilizando (3.5), para todas las combinaciones posibles de 65 patrones observados de fecundidad y 40 tablas modelo de vida, generando un total de $65 \cdot 40 = 2\ 600$ valores para ambas cantidades. Luego calculó regresiones entre los multiplicadores y las razones de paridez media para varios subconjuntos de estas observaciones (Sullivan 1972:82-83). Desarrolló también regresiones similares para aplicarlas a las proporciones de hijos fallecidos entre el total de hijos tenidos por mujeres clasificadas por duración del matrimonio. Se las ha utilizado relativamente poco por falta de informaciones adecuadas (véase, sin embargo, Sishawy 1975). Los resultados de las regresiones por edad de Sullivan mejoran el procedimiento original de Brass al sustituir el polinomio modelo de Brass (2.2) por patrones observados de fecundidad. Trussell refinó aún más el procedimiento incluyendo otras variables independientes en la regresión y sustituyendo los patrones modelo derivados de los modelos de Coale-Trussell (2.6-8) por los modelos observados de Sullivan (Trussell 1975). El significado de este último punto es que el patrón de fecundidad, por edades simples, está disponible, con raras excepciones, solamente para las poblaciones donde el período reproductivo se inicia tarde; de aquí que los resultados de Sullivan son de aplicabilidad dudosa en las poblaciones donde la fecundidad se inicia a edades

tempranas. El modelo de Coale-Trussell genera estructuras por edad de la fecundidad con edades iniciales tempranas, de modo que parcialmente suple la falta de información. Sería preferible utilizar patrones observados, aunque el excelente ajuste de los modelos de Coale-Trussell a diferentes patrones observados sugieren que las estructuras derivadas para fecundidades tempranas son de hecho buenas aproximaciones a las estructuras observadas.

8. Tendencias de la mortalidad a partir de información de hijos sobrevivientes

Supongamos ahora que la mortalidad ha estado cambiando en los años precedentes al censo, pero que la tabla de vida que refleja la mortalidad de cualquier momento anterior al censo corresponde a una tabla de vida de las familias de tablas de un parámetro, y supongamos además que la tendencia en el tiempo de ese parámetro es lineal. Aunque puede considerarse que el parámetro varía en forma continua a través del tiempo, es suficiente suponer que la tabla de vida aplicable a mitad de cada año lo es también a lo largo del año. Entonces el valor del parámetro para el t ésimo año anterior al censo es $\omega_t = \omega + (t-1/2)r$, $t=1, 2, \dots, k$, y se pone $L_x(t)$ en (7.2) equivalente a L_x de la tabla modelo de vida identificada por el valor del parámetro ω_t . Esto define cada $p(t)$ en la ecuación básica (7.1) como una función de ω y r , por lo que se obtiene

$$Q_i = 1 - \sum_{t=1}^{n(i)} p(t; \omega, r) c_i(t), \quad i=1, \dots, N \quad (8.1)$$

La expresión $p(t; \omega, r)$ simboliza aquí la proporción entre las personas nacidas durante el t ésimo año anterior al censo que sobreviven al momento del censo si la mortalidad estuviera cambiando a una tasa constante r durante los años anteriores al censo y hubiera tenido el nivel ω al momento del censo.

Se podría intentar encontrar valores para ω y r para los cuales los valores de la izquierda de la relación (8.1) fueron "próximos" a las proporciones observadas de Q_i o resolver subsistemas de pares de ecuaciones para ω y r .

Otra posibilidad es considerar el "conjunto de solución" de cada ecuación en (8.1), esto es, el conjunto de todas las combinaciones de valores de \underline{r} y ω que satisfacen la ecuación. Los puntos del conjunto de solución pueden simbolizarse $(r, \Omega(r))$, donde \underline{r} toma todos los valores posibles de la tasa de cambio y $\Omega(r)$ simboliza la solución para ω de la ecuación que se plantea manteniendo \underline{r} constante en los valores indicados. El conjunto de solución puede visualizarse mediante el gráfico de la función Ω . Dada una serie finita de valores de \underline{r} , r_1, \dots, r_m , se pueden calcular los correspondientes valores de ω -valores $\omega_1, \dots, \omega_m$, y representando los puntos $(r_1, \omega_1), \dots, (r_m, \omega_m)$ se obtiene una aproximación finita de este gráfico. Los valores intermedios pueden obtenerse por interpolación.

El conjunto de solución de la ecuación iésima en (8.1) define una familia de tendencias lineales en la mortalidad -de un parámetro- coherente con los hijos sobrevivientes en el iésimo grupo de edades. Los hijos sobrevivientes observados, Q_i , para un grupo de edades dado, puede llevar a concluir, por ejemplo, que si la mortalidad hubiera sido constante antes del censo, la mortalidad infantil debería haber sido, en consecuencia, de 197,2 muertes infantiles por mil nacimientos; que si la tasa de mortalidad infantil hubiera venido decreciendo a una tasa anual de $r=2$ muertes infantiles por mil nacimientos, entonces al momento del censo la mortalidad infantil habría sido de 195,0; que si la tasa anual de descenso hubiera sido $r=4$ defunciones infantiles por mil nacimientos, entonces la tasa de mortalidad infantil al momento del censo debió haber sido de 192,7; y así sucesivamente para cualquier serie específica de valores de \underline{r} . Estos valores fueron obtenidos de hecho a partir de información relativa a hijos sobrevivientes del censo de 1960 de Borneo del Norte, en el que para el grupo de edades 15-19, $Q = 0,189$ y los valores de $c(t)$ se estimaron en 0,501, 0,302, 0,148, 0,046 y 0,003, para $t=1, \dots, 5$ y cero para $t>5$. En el cuadro 1 se presenta la información. La familia de tablas modelo que se utilizó para definir la función $p(t; \omega, r), t=1, \dots, 5$, a través de (7.2) se definió mediante la transformación logito de Brass, reparametrizada con la tasa de mortalidad infantil y con β constante e igual a la unidad, aplicada a los valores λ_x del estándar general de Brass (Brass 1971:60-110).

Cuadro 1

INFORMACION DE HIJOS SOBREVIVIENTES PARA BORNEO DEL NORTE: CENSO
DEL 11 DE AGOSTO DE 1960

Grupos de edades	Mujeres total	Hijos nacidos vivos	Hijos sobrevivientes
15-19	19 945	4 349	3 527
10-24	18 175	26 121	21 285
25-29	19 507	59 454	46 733
30-34	14 273	59 733	45 929
35-39	13 809	68 212	50 083
40-44	10 404	54 479	38 505
45-49	8 071	42 840	29 282
50-54	6 207	31 673	20 571
55-59	3 483	17 898	11 395
60-64	3 624	17 750	10 780
65-69	1 798	9 024	5 393
70-74	3 131	15 171	8 041

Fuente: Borneo del Norte, Report on the Census of Population taken on 10 August, 1960, (Kuching, Sarawak: Government Printing Office, marzo 1962). Tabla 5, página 140, para el total de mujeres y tabla 16, página 234, para los hijos tenidos e hijos sobrevivientes.

Específicamente:

$$L_x(\omega) = \begin{cases} 0,3 + 0,7l_1(\omega) & \text{Si } x=0 \\ 0,5[l_x(\omega) + l_{x+1}(\omega)] & \text{Si } x=1, 2, \dots \end{cases} \quad (8.2)$$

donde:

$$l_x(\omega) = \left[1 + \exp \{ [A(\omega) + 0,5 \ln(1 - l_x)] \div l_x \} \right]^{-1} \quad (8.3)$$

y

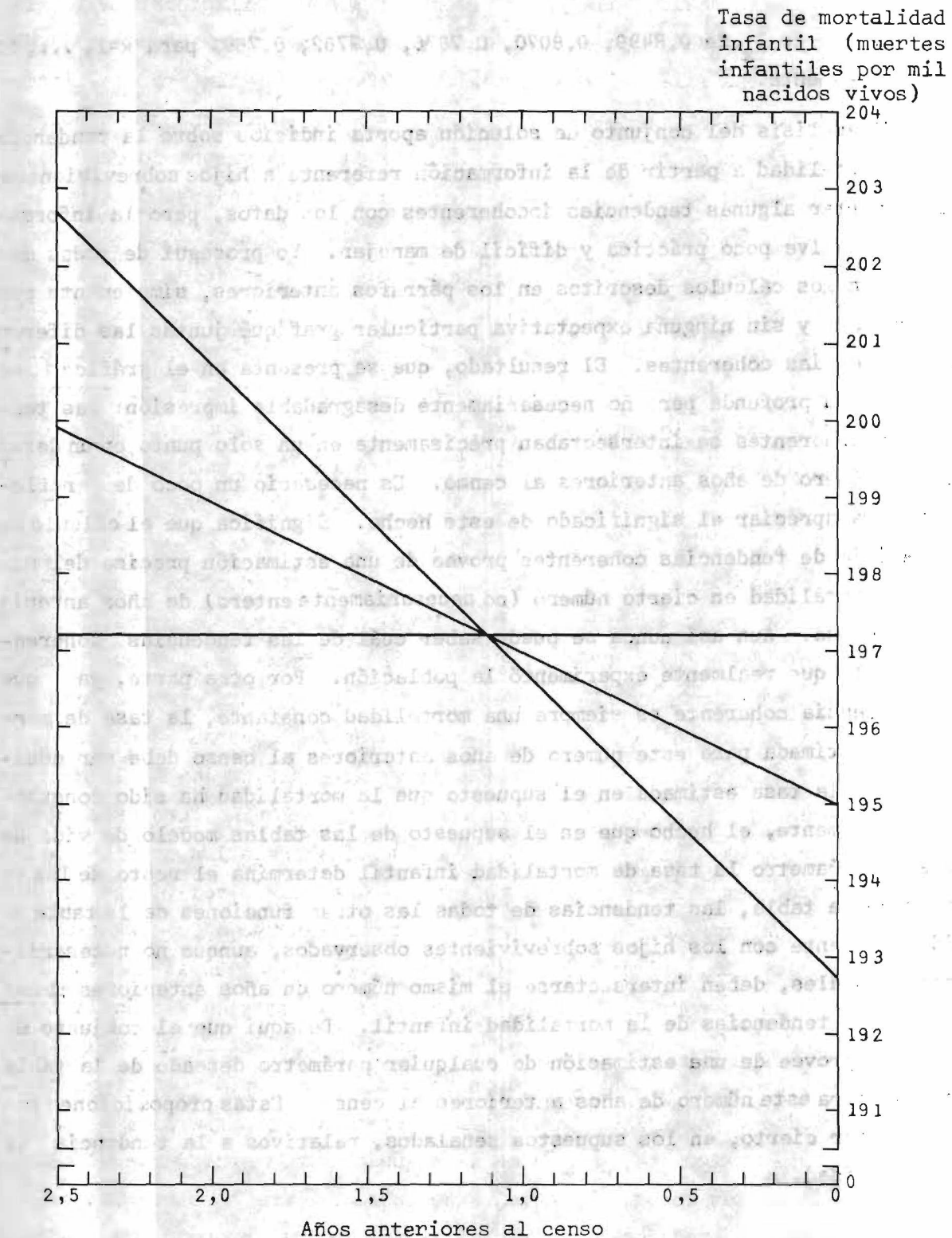
$$A(\omega) = 0,5 \ln \{ [1 \div (1-\omega)]^{-1} \} - 0,5 \ln (1-l_1) \div l_1 \quad (8.4)$$

con valores de l_x de 0,8499, 0,8070, 0,7876, 0,7762, 0,7691 para $x=1, \dots, 5$, respectivamente.

El análisis del conjunto de solución aporta indicios sobre la tendencia de la mortalidad a partir de la información referente a hijos sobrevivientes, al descartar algunas tendencias incoherentes con los datos, pero la información se vuelve poco práctica y difícil de manejar. Yo proseguí de todas maneras con los cálculos descritos en los párrafos anteriores, simplemente por curiosidad, y sin ninguna expectativa particular grafiqué juntas las diferentes tendencias coherentes. El resultado, que se presenta en el gráfico 1, me produjo una profunda pero no necesariamente desagradable impresión: las tendencias coherentes se intersectaban precisamente en un solo punto en un determinado número de años anteriores al censo. Es necesario un poco de reflexión para apreciar el significado de este hecho. Significa que el cálculo de la familia de tendencias coherentes provee de una estimación precisa del nivel de mortalidad en cierto número (no necesariamente entero) de años anteriores al censo. Aun así nunca se puede saber cuál de las tendencias coherentes fue la que realmente experimentó la población. Por otra parte, ya que una tendencia coherente es siempre una mortalidad constante, la tasa de mortalidad estimada para este número de años anteriores al censo debe ser equivalente a la tasa estimada en el supuesto que la mortalidad ha sido constante. Finalmente, el hecho que en el supuesto de las tablas modelo de vida de un solo parámetro la tasa de mortalidad infantil determina el resto de los valores de la tabla, las tendencias de todas las otras funciones de la tabla de vida coherente con los hijos sobrevivientes observados, aunque no necesariamente lineales, deben intersectarse el mismo número de años anteriores al censo que las tendencias de la mortalidad infantil. De aquí que el conjunto de solución provee de una estimación de cualquier parámetro deseado de la tabla de vida para este número de años anteriores al censo. Estas proposiciones son válidas por cierto, en los supuestos señalados, relativos a la tendencia de la mortalidad.

Gráfico 1

TENDENCIAS LINEALES DE LA MORTALIDAD INFANTIL COHERENTES CON LA PROPORCION DE HIJOS SOBREVIVIENTES ENTRE LOS HIJOS NACIDOS VIVOS DE MUJERES CON 15-19 AÑOS, SEGUN EL CENSO DE 1960 DE BORNEO DEL NORTE



Fuente: Cuadro 1 y Sección 2 del texto.

A continuación calculé tendencias lineales coherentes para todos los grupos de edades disponibles en veinte censos diferentes que correspondían a Camboya, Fiji, Indonesia, Malasia, Papua Nueva Guinea, Filipinas, Tailandia, Samoa Americana, Brunei, Borneo del Norte y Sabah (que ahora es parte de Malasia), Corea y las islas Gilbert y Ellice. Ocurre generalmente una intersección común de las tendencias lineales coherentes, y con una gran precisión numérica (del orden de medio a uno por ciento) para los grupos de edades jóvenes y se eleva, a lo sumo, a un 5 por ciento para las mujeres de edad 70-74 en el censo de 1960 de Borneo del Norte. Yo no sé si las intersecciones constituyen sólo buenas aproximaciones o son exactas con discrepancias aparentes debidas a errores de cálculo. La falta de respuesta a una pregunta especulativa y teórica como es ésta no invalida por cierto el uso de los resultados para estimar la mortalidad. Los valores de los años anteriores al censo aumentan con la edad de las mujeres, en alrededor de dos años y medio con cada grupo quinquenal sucesivo. Los valores para las informaciones de Borneo del Norte en el cuadro 1, por ejemplo, son: 1,12, 2,64, 5,28, 7,50, 10,31, 13,55, 17,00, 20,27, 27,07, 25,61 y 27,49 años para los grupos de edades 15-19, ..., 70-74, respectivamente. En consecuencia, aunque las tendencias coherentes para cualquier grupo de edades proveen solamente del nivel de mortalidad a un solo momento, sin indicación acerca de cuál tendencia fue la experimentada, los datos para los grupos de edades tomados en conjunto proporcionan una serie de estimaciones a lo largo de un período de 2-3 décadas anteriores al censo. Esto evidentemente constituye de paso una verificación de la coherencia interna del supuesto de linealidad, ya que los cálculos para cada grupo de edades son completamente independientes. El examen de los gráficos de las estimaciones de las tasas de mortalidad infantil a través del tiempo sugiere con claridad que ni todos los resultados ni todas las veces pueden aceptarse a primera vista. La estimación basada en el grupo de edades 15-19 años indica un aumento de la mortalidad infantil anterior al censo en prácticamente cada caso, incluso en casos de dos o tres censos sucesivos de una misma población. Hay pocas dudas que este aumento sea falso. Otra característica, menos pronunciada, pero claramente visible, es el aumento aparente de la mortalidad durante la tercera, y a veces la segunda década precedente al censo. En la mayoría de los casos, probablemente, puede ser una consecuencia del deterioro de las informaciones correspondientes a los grupos de edades

avanzados. Los resultados, sin embargo, aparentemente proporcionan un panorama útil de la mortalidad para dos, y en algunos casos cerca de tres décadas antes del censo.

9. Conclusiones

En las técnicas de estimación que se presentan aquí, se distinguen tres etapas principales: la formulación de ecuaciones tautológicas que relacionan las estimaciones desconocidas que se desea determinar, con los datos disponibles; la introducción de la parametrización que resulta de derivar ecuaciones con menos incógnitas que las ecuaciones tautológicas originales, y la solución final de estas últimas ecuaciones. Las ecuaciones tautológicas originales, tales como la (3.2), (6.1) o la (7.1), están severamente subdeterminadas, esto es, muchas combinaciones de valores de las cantidades desconocidas que se desea determinar son igualmente coherentes con la información disponible. Cabía esperar esto y podría decirse, ciertamente, que constituye la característica típica de los problemas de la información insuficiente. Si las ecuaciones tautológicas originales tienen una solución única para los valores del parámetro desconocido que quiere determinarse, en términos de la información disponible, casi no puede decirse que la información es "insuficiente", aunque uno podría enfrentarse con el "problema de la información inconveniente" si la solución es difícil de obtener.

Se puede decir que el conjunto de solución de las ecuaciones tautológicas originales agota los indicios contenidos en los datos relevantes para la determinación de las cantidades desconocidas que se desean determinar. Valores específicos de estas cantidades (o un conjunto "menor" de solución) pueden obtenerse solamente introduciendo más información al problema. Ya que puede suponerse que la información disponible ha sido agotada en la formulación de las ecuaciones tautológicas, cabe preguntarse si, en base de una evidencia circunstancial o de regularidades empíricas observadas en otras poblaciones, algunos puntos del conjunto de solución son estimaciones más probables que otros de las cantidades desconocidas que se desea determinar. Esta información se introduce en el problema mediante una serie de ecuaciones de

parametrización, que expresan varias de las cantidades desconocidas en el problema, como funciones de un menor número de parámetros formales. Las familias de los patrones modelo de fecundidad definidos mediante (2.1), (2.2) y (2.3-5) y las familias de tablas modelo de vida utilizadas en las secciones (7) y (8) son ejemplos de parametrización basada en regularidades empíricas, mientras que las relaciones (4.1), (6.2) y (8.1) representan supuestos sobre tendencias en el tiempo de varias cantidades que podrían justificarse por evidencias circunstanciales.

La sustitución de las expresiones paramétricas en las ecuaciones paramétricas por las incógnitas correspondientes en las ecuaciones tautológicas originales da como resultado un sistema de ecuaciones derivadas con un menor número de incógnitas. Son ejemplos de ecuaciones derivadas las relaciones (4.2), (4.7), (6.4), (7.4) y (8.1). Debido a que la parametrización reduce el número de incógnitas en el problema, las ecuaciones derivadas pueden tener una única solución o pueden estar sobredeterminadas, pero en ningún caso estarán menos subdeterminadas que las ecuaciones tautológicas originales. Desde el punto de vista matemático, el único propósito de la parametrización es reducir el número de incógnitas y obtener así un conjunto más manejable de ecuaciones. Desde el punto de vista demográfico la parametrización representa el conocimiento que se introduce en el problema más allá del que está contenido en los datos disponibles. Por la naturaleza de la situación, es imposible verificar directamente el conocimiento así introducido, y debe más bien considerarse a la parametrización como una fuente de error en las estimaciones, como lo son los errores derivados de los datos originales.

La etapa final consiste en resolver las ecuaciones derivadas para valores del parámetro desconocido y calcular los valores correspondientes de las estimaciones desconocidas deseadas originalmente a partir de las ecuaciones paramétricas. Así, por ejemplo, se resuelve (7.4) para w y luego se calcula, digamos $q(5)$ a partir de w utilizando las fórmulas que definen la familia de la tabla modelo de vida. Sin embargo, el último paso puede, en algunos casos, degenerarse como en (4.2-3). Cuando las ecuaciones derivadas están sobredeterminadas se busca las soluciones del "mejor ajuste".

Prácticamente en todas las técnicas asociadas con estimaciones a partir de datos "limitados o insuficientes" pueden observarse estas etapas, con varias particularidades individuales.

Con relación al desarrollo de técnicas en el futuro, es evidente que las posibilidades de la parametrización "débil", esto es, parametrizaciones que implican mayor número de parámetros desconocidos (el punto de vista matemático) y menores supuestos demográficos restrictivos (el punto de vista demográfico) apenas ha comenzado a explorarse. Es típico de los métodos corrientemente utilizados, reducir los problemas a determinar una sola incógnita. Prácticamente no se utilizan los métodos de evaluación, en contraste con los métodos derivativos para los sistemas simultáneos de ecuaciones no lineales y para la minimización numérica de las soluciones por "mejor ajuste", y mientras que es verdad que las dificultades tienden a crecer con el cuadrado o el cubo del número de incógnitas, las facilidades de la computación moderna presentan problemas bi o tridimensionales que pueden resolverse.

Con respecto a la tabulación, hay un punto demasiado obvio para dejar de mencionarse, y se refiere a la gran conveniencia de producir tabulaciones por años simples. El recurrir a rodeos elaborados, cuando se trata de estimar la fecundidad con información agrupada, puede ser inevitable en el caso en que los números son excesivamente pequeños, pero esto no sucede nunca en el caso de información proveniente de censos o de estadísticas vitales y a menudo en el caso de informaciones de encuestas grandes. Agrupar la información es, por cierto, una herramienta analítica útil, pero sólo si se tiene la opción de elegir el agrupamiento y esto implica una tabulación básica por años simples.

REFERENCIAS

- Arretx, Carmen, "Fertility Estimates Derived from Information on Children Ever-born Using Data from Successive Censuses", en International Population Conference: Liege 1973. Volumen 2, pp. 247-261. Lieja: Unión Internacional para el Estudio Científico de la Población, 1973.
- Brass, William, "Disciplining Demographic Data", en International Population Conference: London 1969. Volumen 1, pp. 183-204. Lieja: Unión Internacional para el Estudio Científico de la Población, 1971a.
- Brass, William, "On the Scale of Mortality", en W. Brass, Ed., Biological Aspects of Demography. Nueva York, Barner and Noble, 1971b.
- Brass, William, Methods of Estimating Fertility and Mortality from Limited and Defective Data. Basado en seminarios efectuados el 16-24 septiembre 1971, en el Centro Latinoamericano de Demografía (CELADE), San José, Costa Rica. Una publicación ocasional de Laboratories for Population Statistics, Universidad de Carolina del Norte en Chapel Hill. Chapel Hill, N.C., 1975.
- Brass, William, y otros, The Demography of Tropical Africa. Princeton, New Jersey: Princeton University Press, 1968.
- Coale, A.J. y D.R. McNeil, "The Distribution by Age of the Frequency of First Marriage in a Female Cohort", en Journal of the American Statistical Association 67 (340, diciembre), 743-749, 1972.
- Coale, Ansley J. y T. James Trussell, "Model Fertility Schedules: Variations in the Age Structure of Child-bearing in Human Populations", en Population Index 40 (2), 185-258, 1974.
- Feeney, Griffith, Tabulation of Census and Survey Data on Child Survivorship. Asian and Pacific Census Newsletter 3(1), 5-6. Disponible en el East-West Population Institute, East-West Center, Honolulu, Hawaii, 1976a.
- Feeney, Griffith, "Estimating Infant Mortality Rates from Child Survivorship Data by Age of Mother", en Asian and Pacific Census Newsletter 3(2). 12-16, 1976b.
- Mortara, Giorgio, Methods of Using Census Statistics for the Calculation of Life Tables and Other Demographic Measures. (Con aplicaciones a la población del Brasil) Population Studies, N° 7, Department of Social Affairs, United Nations. Lake Success, Nueva York, 1949.
- Macura, Miroslav, Estimates of the Completeness of Registration of Births and Infant Deaths in Yugoslavia and its Main Provinces from the Late 1940's to 1961. Ph.D. Dissertation, Department of Economics, Princeton University. (Disponible en Xerox University Microfilms, Ann Arbor, Michigan), 1972.

McNeil, D.R. y J.W. Tukey, "Higher-order Diagnosis of Two-way Tables, Illustrated on Two Sets of Demographic Empirical Distributions", en Biometrics 31 (junio), 487-510, 1975.

Sishawy, Nazek Nosseir el., Measures of Fertility and Mortality in Governance of Egypt 1947 and 1960. Ph. D. Dissertation, Office of Population Research, Princeton University, 1975.

Sullivan, Jeremiah, "Models for the Estimation of the Probability of Dying between Birth and Exact Ages of Childhood", en Population Studies 26(1), 116-133, 1972.

Trussell, T. James, "A Re-estimation of the Multiplying Factor for Determining Childhood Survival", en Population Studies 29(1), 97-107, 1975.

Tukey, John, "The Future of Data Analysis", en Annals of Mathematical Statistics 33, 1-67, 1962.