

UN SISTEMA INTEGRADO DE ESTIMACIONES DEMOGRAFICAS A PARTIR DE DOS DISTRIBUCIONES POR EDAD*

*Samuel H. Preston
Population Studies Center
University of Pennsylvania*

RESUMEN

Este documento presenta un método simple para estimar la tasa de natalidad y un nivel de mortalidad para un período intercensal. La tasa de natalidad se estima por la intersección de una línea recta de ajuste de la información y el nivel de la mortalidad por la pendiente de esa línea. La fórmula que se desarrolla se basa en una generalización reciente de relaciones en poblaciones estables. Una estimación de la mortalidad al comienzo de la vida constituye una información optativa adicional relevante. Un importante subproducto del procedimiento es una estimación de la verdadera distribución por edades.

<TASA DE NATALIDAD>
<MORTALIDAD INFANTIL >
<PERIODO INTERCENSAL >
<ESTIMACION DE POBLACION >

* Traducción de: "An integrated system for demographic estimation from two ages distributions". *Demography*, Volume 20, Number 2, May, 1983.

AN INTEGRATED SYSTEM FOR DEMOGRAPHIC ESTIMATION FROM TWO AGE DISTRIBUTIONS

SUMMARY

This paper presents a simple method for estimating a birth rate and a level of mortality for an intercensal period. The birth rate is estimated from the intercept of a line fitted to data and the level of mortality from the slope of that line. The formula that is developed is based upon a recent generalization of stable population relations. An estimate of childhood mortality level is an optional but significant piece of additional input. An important by-product of the procedure is an estimate of the true age distribution.

<BIRTH RATE>

<INFANT MORTALITY>

<INTERCENSAL RATES>

<POPULATION ESTIMATE>

INTRODUCCION

Este documento presenta un método simple para estimar tasas de natalidad y, simultáneamente, un parámetro representativo del nivel de la mortalidad. Ha sido diseñado para ser usado en países que carecen de buenas estadísticas vitales, los que comprenden a la mayoría de la población del mundo (Naciones Unidas, 1980). La tasa de natalidad es inferida por la intersección en el origen de una línea recta; el nivel de la mortalidad se estima por la pendiente de esa línea. El procedimiento de estimación integra el sistema logito de un parámetro de mortalidad ideado por Brass con ecuaciones desarrolladas recientemente que generalizan la teoría de la población estable. El sistema que se propone requiere de dos distribuciones por edades sucesivas.

Una breve presentación, de un párrafo, de este método aparece en el artículo de Preston y Coale (1982, página 244). Aquí derivamos la fórmula básica, la extendemos a la situación donde se cuenta con una estimación de la mortalidad al comienzo de la vida, aplicamos el procedimiento a la información real y se comenta su sensibilidad a varias formas de error.

ANTECEDENTES

Ecuaciones estables generalizadas

Preston y Coale (1982) han mostrado que la siguiente relación se da en todo momento, en el tiempo, en una población cerrada.

$$c(a, t) = b(t)e^{-\int_0^a r(x, t)dx} p(a, t) \quad (1)$$

donde:

- $c(a, t)$ = proporción de la población de edad a en un momento t ;
- $b(t)$ = es la tasa de natalidad en el momento t ;
- $r(x, t)$ = es la tasa anual de crecimiento de personas con edad x en el momento t ;
- $p(a, t)$ = la probabilidad de sobrevivencia, de un recién nacido, hasta la edad a , de acuerdo con la tabla de vida del período correspondiente al momento t .

Relaciones estrechamente conectadas con esta ecuación se dan también para intervalos discretos de tiempo. La forma más directa de

desarrollar una analogía con ecuaciones expresadas en intervalos discretos de tiempo es suponer que la tasa anual de crecimiento, $r(x,t)$, es constante a lo largo del tiempo dentro del intervalo considerado. En este caso, $c(a,t)$ en la fórmula presentada anteriormente, es la relación entre el número de personas que cumplieron a años durante el período, y los años-persona vividos por la población, $b(t)$ son los nacimientos divididos por el total de años-persona vividos en el intervalo de tiempo y $p(a,t)$ es un promedio de las funciones $p(a,t)$ prevalecientes durante el período de tiempo ponderadas por los nacimientos. Si $r(x,t)$ es constante a través de las edades, la ecuación (1) se reduce a la familiar ecuación que caracteriza a una población estable.

Esta ecuación también se aplica a una población abierta, en tanto que $r(x,t)$ exprese la suma de tasas por edad, de crecimiento a la edad x y el tiempo t , y tasas de emigración neta (por años-persona) a la edad x en el momento t . De aquí en adelante, por razones de conveniencia, dejaremos de escribir la letra t .

Mediante transposición de términos, pasamos de la ecuación (1) a la ecuación (2), obteniendo:

$$\frac{1}{p(a)} = \frac{b e^{-\int_0^a r(x) dx}}{c(a)} \quad (2)$$

Transformación logito de Brass

Brass (1971) muestra que una relación lineal de la transformación logito de la función de sobrevivencia $p(a)$, o de su complemento $q(a)$, frecuentemente proporciona una buena representación de los cambios en las condiciones de mortalidad por edad a medida que los niveles de mortalidad se modifican. En particular, supuso que la expresión siguiente era válida para cada conjunto $q(a)$ dentro de un sistema modelo de tablas de vida.

$$\ln \frac{q(a)}{p(a)} = \alpha + B \ln \frac{q_s(a)}{p_s(a)} \quad (3)$$

donde $q_s(a)$, $p_s(a)$ son las funciones $q(a)$ y $p(a)$ en el estándar adoptado, y α y B son parámetros que relacionan la mortalidad de cualquier tabla

de vida del sistema con la tabla estándar. α es el "parámetro del nivel"; cuando ese valor cambia, se modifica el nivel de la mortalidad en la misma dirección para todas las edades. B es el parámetro de la pendiente; cuando se eleva, se aumenta la expresión $\ln [q(a)/p(a)]$ para todas las edades por arriba de la edad para la cual $q_s(a)/p_s(a)$ es igual a 1, y se disminuye para todas las edades inferiores. Valores tabulados de un sistema de tabla de vida logito se encuentran en Carrier y Hobcraft (1971). Más adelante se presenta alguna evidencia del éxito de la transformación logito.

La ecuación 3 puede ser reordenada como sigue:

$$\frac{1 - p(a)}{p(a)} = e^{\alpha} \left[\frac{q_s(a)}{p_s(a)} \right]^B \quad \text{ó}$$

$$\frac{1}{p(a)} = e^{\alpha} \left[\frac{q_s(a)}{p_s(a)} \right]^B + 1 \quad (4)$$

La derivación de ecuaciones de estimación

Las ecuaciones 2 y 4 pueden combinarse para dar:

$$\frac{be^{-\int_0^a r(x)dx}}{c(a)} = e^{\alpha} \left[\frac{q_s(a)}{p_s(a)} \right]^B + 1 \quad \text{ó}$$

$$\frac{e^{-\int_0^a r(x)dx}}{c(a)} = \frac{1}{b} + \frac{K}{b} \left[\frac{q_s(a)}{p_s(a)} \right]^B \quad (5)$$

Hacemos a $e^{\alpha} = K$ por razones de simplicidad en la exposición. Nótese que la ecuación (5) requiere procedimientos de estimación no lineales, a menos que $B = 1$. Poner $B = 1$ es un supuesto razonable, ya que es raro tener información censal de suficiente calidad que permita estimaciones precisas tanto de α como de B . B es, de los dos parámetros, el más vulnerable a malas declaraciones sistemáticas de edad (como se ilustrará más abajo). Por lo general, el mayor interés está en conocer el "nivel" de la mortalidad. Luego, suponiendo que $B = 1$, tenemos la expresión:

$$\frac{e^{-\int_0^a r(x)dx}}{c(a)} = \frac{1}{b} + \frac{K}{b} \frac{q_s(a)}{p_s(a)} \quad (6)$$

La ecuación (6) es una simple ecuación lineal. La intersección es la recíproca de la tasa de natalidad y la pendiente es K (el "nivel" de la mortalidad) dividida por la tasa de natalidad. Los elementos requeridos para estimar los valores del término de la izquierda se derivan fácilmente a partir de dos censos. La variable del lado derecho, $q_s(a)/p_s(a)$, es, desde luego, dada por una tabla de vida supuesta, la que eventualmente podría modificarse por el valor de K que se determine.

Es muy común que un país tenga información razonablemente buena sobre niveles de mortalidad al comienzo de la vida, en virtud de las preguntas sobre el número de hijos tenidos y sobrevivientes. Esta información puede ser convenientemente integrada en el procedimiento de estimación, lo que deberá hacerse cada vez que esté disponible. El índice de mortalidad de la niñez de mayor utilidad para los propósitos presentes es $p(5)$, esto es, la probabilidad de sobrevivir a la edad 5 de un recién nacido. Si la información sobre hijos tenidos y sobrevivientes es obtenida del segundo de dos censos separados por 10 años, $p(5)$ se referirá a un punto aproximadamente intermedio del período intercensal (Academia Nacional de Ciencias, 1981). La estimación de la mortalidad que se busca será, precisamente, para un punto intermedio del período intercensal.

Definimos ${}_5p_s(a)$ y ${}_5q_s(a)$ como la probabilidad de sobrevivir hasta la edad a y de morir antes de la edad a , respectivamente, de alguien que ha sobrevivido hasta la edad 5 en la tabla de vida estándar. Así ${}_5q_s(5) = 0$ y ${}_5p_s(5) = 1,00$. Suponemos que la transformación logito de un parámetro se da solamente para la mortalidad por encima de 5:

$$\frac{{}_5q(a)}{{}_5p(a)} = K \frac{{}_5q_s(a)}{{}_5p_s(a)} \quad \text{para } a \geq 5$$

Los valores que aparecen en la tabla 1 demuestran que la transformación logito de un parámetro, para mortalidad por encima de 5, dentro del sistema de tablas modelo de vida Oeste, representa variaciones de la mortalidad con mucha efectividad a lo largo de un intervalo de 30 años en la variación de las esperanzas de vida al nacer. Transformando la función ${}_5p(a)$ en una tabla de vida con $e_0 = 50$, las funciones ${}_5p(a)$ correspondientes a $e_0 = 35$ y $e_0 = 65$ pueden ser reproducidas con bastante

Cuadro 1

PROBABILIDADES DE SOBREVIVIR DESDE LA EDAD 5 HASTA
 VARIAS EDADES EN TABLAS MODELO DE VIDA FEMENINA,
 FAMILIA OESTE.
 ESPERANZAS DE VIDA AL NACER DE 35 Y 65.

Edad <i>a</i>	Esperanza de vida al nacer de 35 años ${}_5p(a)$		Esperanza de vida al nacer de 65 años ${}_{65}p(a)$	
	Real ^a	Estimada ^b	Real ^a	Estimada ^b
5	1,000	1,000	1,000	1,000
10	0,959	0,956	0,993	0,991
15	0,927	0,923	0,988	0,984
20	0,888	0,882	0,979	0,974
25	0,841	0,833	0,969	0,961
30	0,791	0,782	0,956	0,947
35	0,737	0,729	0,942	0,930
40	0,683	0,676	0,924	0,912
45	0,628	0,623	0,903	0,891
50	0,573	0,568	0,876	0,867
55	0,508	0,504	0,838	0,835
60	0,435	0,433	0,787	0,791
65	0,345	0,349	0,713	0,726
70	0,250	0,260	0,609	0,635
75	0,154	0,170	0,469	0,503
80	0,075	0,092	0,305	0,334

^a Fuente: Coale-Demeny (1966).

^b Derivada a través de una transformación de un parámetro de ${}_5p(a)$ de la tabla modelo de vida femenina, de Coale-Demeny familia Oeste, con $e_0=50$. Para derivar la estimación para $e_0=35$, el cociente ${}_5q(a)/{}_5p(a)$ en la tabla modelo de vida con $e_0=50$ fue multiplicado por una $K=2,12$ y los resultados transformados en ${}_5p(a)$. Para la estimación con $e_0=65$, se utilizó el mismo procedimiento con una $K=0,43$.

precisión. Un intervalo de 30 años excede seguramente el rango de incertidumbre en la mayoría de las aplicaciones. Una demostración gráfica es más convincente todavía, pero representar la serie de valores reales y estimados los colocaría tan próximos entre sí que difícilmente podrían ser diferenciados en el gráfico.

Designamos el valor observado de $p(5)$ como $p^*(5)$. Toda vez que

$$p(a) = p^*(5) \cdot {}_5p(a) \quad \text{para } a \geq 5,$$

podemos volver a la ecuación (2) y reescribir la expresión (6), después de introducir la nueva parametrización de la mortalidad.

$$\frac{p^*(5)e^{-\int_0^a r(x)dx}}{c(a)} = \frac{1}{b} + \frac{K}{b} \left[\frac{{}_5q_s(a)}{{}_5p_s(a)} \right] \quad a \geq 5 \quad (7)$$

La ecuación (7) continúa siendo una simple ecuación lineal cuyos parámetros pueden ser estimados rápidamente. Pero ahora incorpora la importante información adicional relativa a la mortalidad al comienzo de la vida.

IMPLEMENTANDO EL SISTEMA

Vamos a suponer que se dispone de una estimación de $p^*(5)$. La información sobre la distribución por edades de la población de dos censos, clasificada a intervalos quinquenales, se supone también disponible. El procedimiento puede ser aplicado a información dada a intervalos quinquenales o a información clasificada por edades simples. La situación primera (población clasificada por grupos quinquenales) es más simple, sin que se sacrifique precisión. Los pasos son los siguientes:

(1) Estimar ${}_s r_x$, la tasa de crecimiento por edad de la población con edades entre x y $x+5$, como:

$${}_s r_x = \frac{\ln {}_s N_x(t+h) - \ln {}_s N_x(t)}{h}$$

donde ${}_s N_x(t)$ es el número de personas enumeradas con edades entre x y $x+5$ en el momento t (el primer censo) y h es la amplitud del período intercensal expresado en un número exacto de años. El conjunto de ${}_s r_x$ puede ser calculado para $x = 0,5,10, \dots$ hasta una edad final arbitraria. Las tasas netas de emigración por edad, si están disponibles, deben ser agregadas a la serie de valores de ${}_s r_x$ (Preston-Coale, 1982).

(2) Estimar $c(a)$, el cociente entre el número de personas que alcanzan la edad a , y el total de años-persona vividos durante el período. El procedimiento más satisfactorio para hacer esto, en vista de los supuestos subyacentes a las fórmulas, es estimar los años-persona vividos entre a y $a+5$ en el período t y $t+h$ como

$${}_s \hat{N}_a(t \text{ a } t+h) = \frac{{}_s N_a(t+h) - {}_s N_a(t)}{{}_s r_a \cdot h}$$

para cada edad, y luego derivar la $c(a)$ como,

$$c(a) = \frac{{}_s \hat{N}_a(t \text{ a } t+h) + {}_s \hat{N}_{a-5}(t \text{ a } t+h)}{10 \sum_{x=0}^{\infty} {}_s \hat{N}_x(t \text{ a } t+h)} \quad (8)$$

(3) Los valores del miembro de la izquierda de la ecuación (7) pueden ahora ser calculados para todas las edades $a = 5, 10, 15, \dots$ Por ejemplo:

Edad	$\frac{p^*(5)e^{-\int_0^a r(x)dx}}{c(a)}$
5	$\frac{p^*(5)e^{-5 {}_s r_0}}{c(5)}$
10	$\frac{p^*(5)e^{-5({}_s r_0 + {}_s r_5)}}{c(10)}$
15	$\frac{p^*(5)e^{-5({}_s r_0 + {}_s r_5 + {}_s r_{10})}}{c(15)}$
⋮	
etc.	

(4) Es ahora necesario adoptar una tabla estándar de mortalidad para edades superiores a los 5 años. Desde luego, ésta debe ser lo más próxima posible al nivel y a la forma presuntos de la mortalidad adulta de la población que se está estudiando. Si no hay información disponible podría elegirse un nivel que correspondiera, en una tabla modelo de vida, al valor particular de $p^*(5)$. Nótese que no es necesario hacer ningún cálculo de logito en esta elaboración; los únicos valores que se requieren son:

$$\frac{{}_s q(a)}{{}_s p(a)} = \frac{l_s - l_a}{l_a}$$

de la tabla de vida elegida como estándar.

(5) Graficar la variable dependiente

$$y = \frac{p^*(5)e^{-\int_0^a r(x)dx}}{c(a)}$$

y la variable independiente

$$x = \frac{{}_s q_s(a)}{{}_s p_s(a)}$$

El gráfico debe mostrar una tendencia aproximadamente lineal. Si no es así, existe un error en la información, en los cálculos o en los supuestos. La *única* hipótesis que se ha adoptado, sin embargo, es que la mortalidad adulta (eso es, más allá de los 5 años) puede ser representada por una transformación logito de un estándar. El analista podrá, desde luego, experimentar con otros estándares y diferentes valores de B . Más adelante se examinan otras fuentes de error.

(6) Ajustar una línea recta a la relación entre las variables. Hay varias formas de hacer esto. Las dos alternativas más usuales son los procedimientos que utilizan el método de los mínimos cuadrados, que tienden a dar un peso mayor a las observaciones extremas, y los procedimientos de medias de agrupamientos de puntos, que normalmente se prefieren, ya que en ellos se reduce la sensibilidad de los puntos aislados. Con independencia del procedimiento de ajustamiento utilizado, la intersección es la recíproca de la tasa de natalidad. La pendiente dividida por la intersección es K , el factor por el cual ${}_s q_s(a) / {}_s p_s(a)$ debe ser multiplicado a fin de estimar la mortalidad adulta de la población que se investiga.

APLICACIONES

India

La tabla 2 muestra los valores requeridos para aplicar el procedimiento a la población femenina de la India durante el período intercensal 1961-1971. Los puntos (columnas 4 y 5, que designaremos de aquí en adelante y y x), están representados en el gráfico 1. Los puntos caen

Cuadro 2
APLICACION DEL PROCEDIMIENTO INTEGRADO A LA POBLACION FEMENINA DE LA INDIA
 1961-1971

Edad al inicio del intervalo a o x	Tasas de crecimiento por edad 1961-1971 $s^r x$ (1)	Acumulación de las tasas por edad hasta la edad a $\sum_{x=0}^a s^r x$ (2)	Proporción de personas que alcanzan la edad a en el período intercensal $c(a)/a$ (3)	$\frac{p^*(5)}{Col(3)} \cdot \exp[-5 \cdot Col(2)]^b$ (4)	$s a_s(a) / s p_s(a)$ en la tabla modelo femenina, familia Sur, Nivel 13 ($e_0 = 50$) ^c (5)
0	0,0180	0,0180	0,03016	23,515	0,0000
5	0,0232	0,0412	0,02651	23,823	0,0227
10	0,0340	0,0752	0,01982	26,883	0,0363
15	0,0255	0,1007	0,01684	27,851	0,0561
20	0,0119	0,1126	0,01666	26,527	0,0819
25	0,0127	0,1253	0,01498	27,685	0,1113
30	0,0186	0,1439	0,01263	29,921	0,1433
35	0,0280	0,1719	0,01079	30,448	0,1795
40	0,0207	0,1926	0,00897	33,029	0,2210
45	0,0226	0,2152	0,00759	34,859	0,2706
50	0,0167	0,2319	0,00585	41,605	0,3414
55	0,0271	0,2590	0,00480	44,280	0,4478
60	0,0221	0,2811	0,00380	50,081	0,6410
65	0,0347	0,3158	0,00227	70,483	1,0104
70	0,0232	0,3390	0,00148	96,267	1,8471
75	0,0235	0,3625	0,00080	158,349	4,1223
80	0,0192				
85	0,0271				

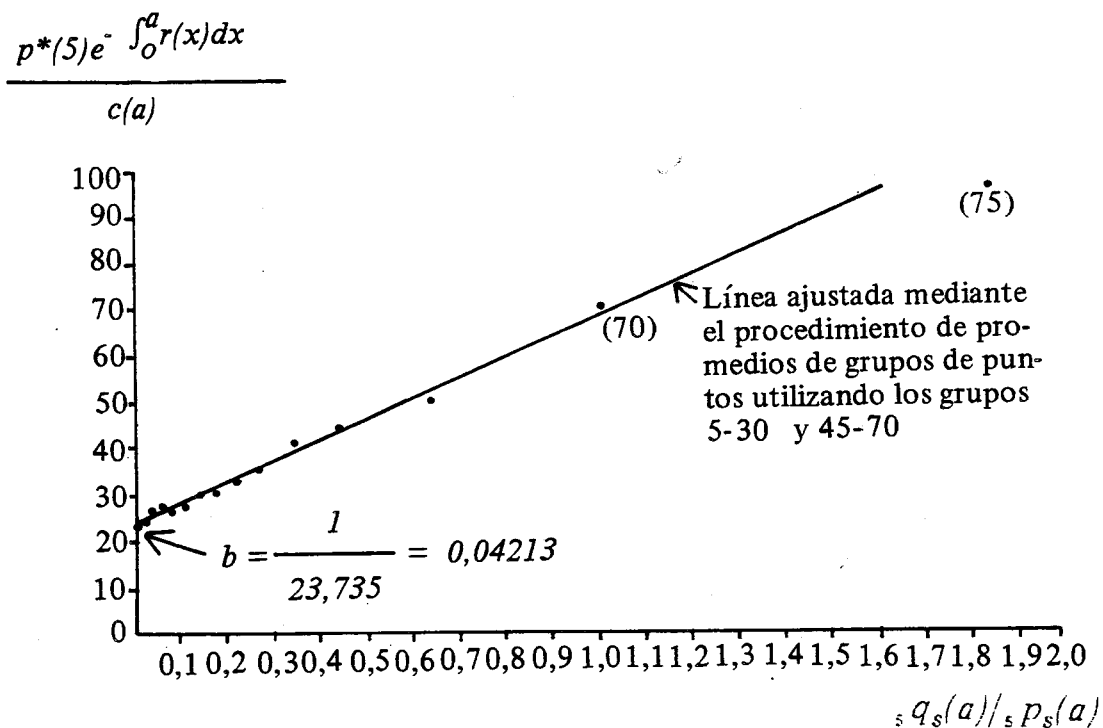
a Calculado por la fórmula (8).

b El valor estimado de $q(2)$ para el centro del período intercensal vale 0,184, de acuerdo con cálculos, no publicados, elaborados por P.N. Mary Bhat para la Academia de Ciencias de los EE.UU. Esa estimación fue utilizada para derivar la correspondiente $p(5)$, con un valor 0,776, utilizando la tabla modelo de vida, familiar Sur.

c Coale-Demeny (1966).

Gráfico 1

APLICACION DEL PROCEDIMIENTO INTEGRADO A LA POBLACION FEMENINA DE LA INDIA, 1961-1971



muy cerca de una línea recta, excepto el punto correspondiente a la edad 75. Una línea recta se ajustó a los puntos, por el procedimiento de medias agrupadas, utilizando el valor medio de los valores x e y para los grupos de edades entre 5 y 30 como una observación, y la medida de los valores de x y y para los grupos de edades entre 45 y 70, como la otra. La ecuación de la línea es:

$$y = 23,735 + 44,992x$$

implicando que

$$b = \frac{1}{23,735} = 0,0421 \quad y$$

$$K = \frac{44,992}{23,735} = 1,896$$

Cuando el valor estimado de K se aplica a ${}_5q_s(a)/{}_5p_s(a)$ y se calcula una tabla de vida a partir de la edad 5, el valor estimado de la esperanza de vida a la edad 5 es de 51 años¹. Este valor cae entre el estimado de 53,2 años (Dyson, 1979), también utilizando el modelo de tablas de vida modelo Sur e información intercensal, y el de Registrar General of India (1977, página 16) de 50,2 años. Como señala Dyson, hay razones para creer que la estimación del Registrar General es demasiado baja. Preston y Bennett (1982, p. 15), utilizando un método que no requiere de un sistema de tablas modelo de vida, derivan una estimación de la esperanza de vida a los años, e_5^0 , de 53,6.

El valor estimado de la tasa de natalidad, 42,1/1000, está muy próximo al rango de 40,5-42,0 dado por Adlakha y Kirk (1974) para el período intercensal; es posible que nuestra estimación sea algo más alta porque utiliza una mayor estimación de la mortalidad de la niñez, basada en resultados publicados recientemente de la Encuesta Nacional por Muestreo de 1965-66. Nuestra estimación es sustancialmente mayor que aquéllas derivadas del sistema de registro por muestreo de la India para años cercanos al final de la década.

Corea del Sur

La tabla 3 y el gráfico 2 muestran valores comparables utilizados para aplicar el sistema a Corea del Sur. Se eligió una tabla estándar para la mortalidad por encima de los 5 años, que se cree refleja mejor las condiciones de Corea: la tabla femenina nivel 19, de la familia Oeste ($e_0 = 65$).

La ecuación de ajuste de la línea mediante el procedimiento de promedios de grupos de puntos, aplicado a los grupos de edades 5-30 y 45-70, respectivamente es:

¹ La esperanza de vida para la edad 5 se deriva primero calculando ${}_5p(a)$ para $a = 10, 15, \dots, 80$ como ${}_5p(a) = 1/[1,896 q_s(a)/p_s(a) + 1]$. Usando el supuesto que las muertes están distribuidas uniformemente dentro de cada grupo de edades, el número de años-persona de vida entre 5 y 80 años son $5[p(10) + p(15) + \dots + p(75)] + 2.5 [p(5) + p(80)]$. A esta suma se agrega un valor $p(80) \hat{e}_{80}$ donde \hat{e}_{80} se toma de una tabla modelo de mortalidad con el mismo valor de $p(80)/p(5)$ (En este caso el modelo Sur Nivel 8.9). La suma resultante es la esperanza de vida a la edad 5.

Cuadro 3
 APLICACION DEL PROCEDIMIENTO INTEGRADO A LA POBLACION FEMENINA DE COREA DEL SUR
 1966-1975

Edad al inicio del intervalo a o x	Tasas de crecimiento por edad 1961-1971 ${}_5r_x$ (1)	Acumulación de las tasas por edad hasta la edad a $\sum_{x=0}^{a-5} {}_5r_x$ (2)	Proporción de personas que alcanzan la edad a en el período intercensal $c(a)^a$ (3)	$\frac{p^*(5)}{Col(3)} \cdot \exp \left[5 \cdot Col(2) \right]^b$ (4)	${}_5q_s(a) / {}_5p_s(a)$ en la tabla modelo femenino familia Sur Nivel 19 ($e_0=65$) ^c (5)
0	-0,0066				
5	-0,0036	-0,0066	0,02716	36,265	0,0000
10	0,0255	-0,0102	0,02619	38,284	0,0069
15	0,0484	0,0153	0,02273	38,836	0,0123
20	0,0358	0,0637	0,01858	37,298	0,0207
25	0,0101	0,0995	0,01567	36,976	0,0321
30	0,0117	0,1096	0,01406	39,181	0,0458
35	0,0306	0,1213	0,01254	41,434	0,0619
40	0,0318	0,1519	0,01101	40,496	0,0817
45	0,0330	0,1837	0,00915	41,565	0,1069
50	0,0281	0,2167	0,00758	42,542	0,1414
55	0,0192	0,2448	0,00633	44,267	0,1927
60	0,0317	0,2640	0,00507	50,208	0,2709
65	0,0229	0,2957	0,00401	54,175	0,4027
70	0,0243	0,3186	0,00294	65,899	0,6416
75	0,0242	0,3429	0,00192	89,362	1,1302
80	0,0555	0,3671	0,00110	138,202	2,2769
85	0,0693				

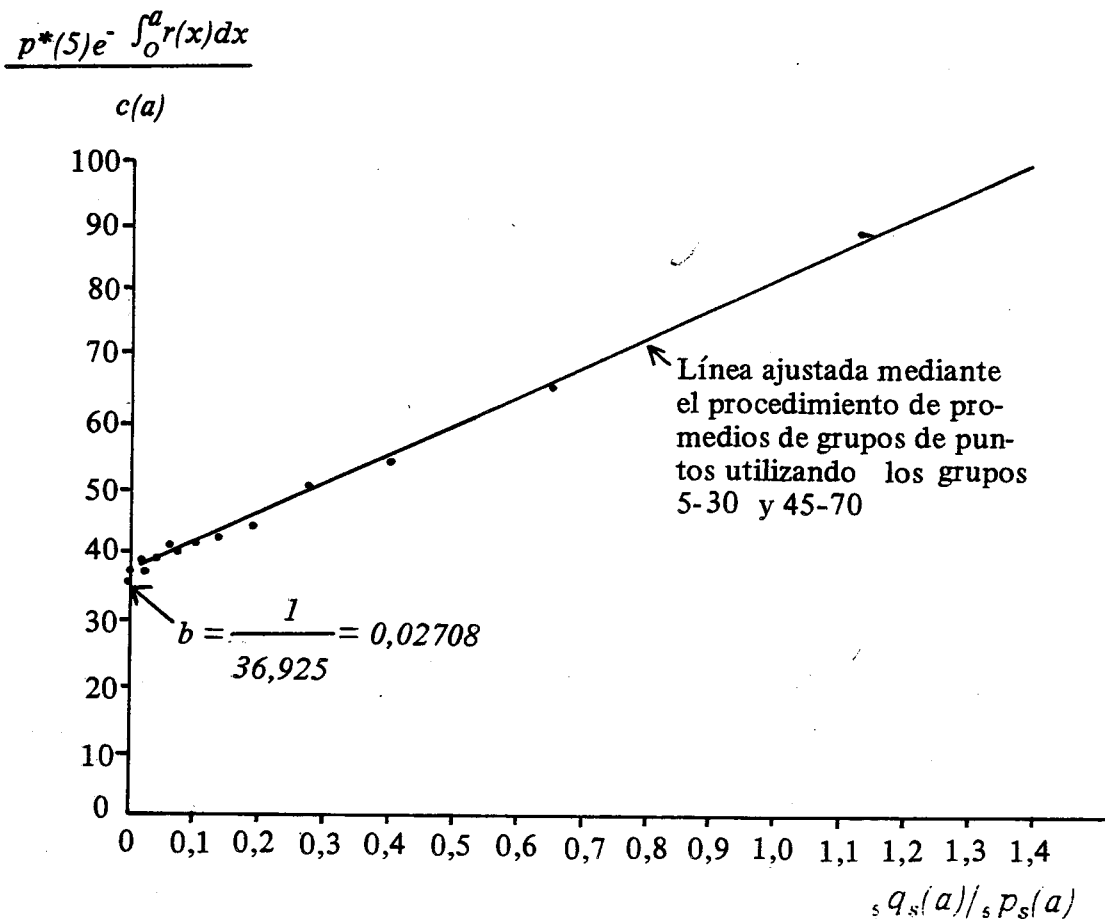
^a Calculada de la fórmula (8).

^b $p^*(5)$ estimado como 0,95292 haciendo una ponderación de las tablas de vida 1966-1970 y 1970-1975 presentadas en el libro Coale y otros (1980 p. 35).

^c Coale-Demeny (1966).

Gráfico 2

APLICACION DEL PROCEDIMIENTO INTEGRADO A LA POBLACION FEMENINA DE COREA DEL SUR, 1966-1975



$$y = 36,947 + 43,831x$$

que implica

$$b = \frac{1}{36,947} = 0,02707 \quad y$$

$$K = \frac{43,831}{36,947} = 1,186$$

Los niveles implícitos de la tasa de natalidad y de la esperanza de vida concuerdan con los presentados por Coale y otros (1980). El promedio no ponderado de las tasas de natalidad, para el período de 1967 a 1975, dado en esa fuente es de 28,6². Se deriva principalmente del uso de las técnicas de hijos propios. Algunas de las relativamente pequeñas diferencias entre las dos cifras resultan de que el procedimiento intercensal descrito aquí asigna implícitamente mayor peso a las observaciones más recientes, en una población que crece, tal como la de Corea del Sur. Durante el período, la tasa de natalidad de Corea descendió rápidamente. Un valor de $K = 1,86$ se traduce en una esperanza de vida a los 5 años de 63,07, casi idéntica al valor presentado por Coale y otros (1980, p. 35), basado en la tabla modelo de vida Oeste para 1966-70 (63,16), pero menor que aquél para el período 70-75 (65,07). Utilizando un procedimiento de estimación intercensal no paramétrico, Preston y Bennett (1982) derivaron un valor de $e_s = 62,7$ para el período intercensal.

Sensibilidad al error

El sistema de estimación propuesto aquí presenta varias ventajas en comparación con los procedimientos existentes. Quizás lo más atractivo es que parece proporcionar una estimación razonablemente robusta de la tasa de natalidad. Los métodos indirectos de estimación han sido mucho más extensamente desarrollados para la mortalidad que para la fecundidad. Más adelante, examinaremos brevemente la sensibilidad de los resultados a varias formas de error y de procedimientos usados.

Selección de los puntos para ajustar la línea

Experimentando con diferentes combinaciones de puntos para ajustar líneas en los dos casos presentados anteriormente (pero siempre incluyendo las edades más bajas y las relativamente más altas) los resultados en las tasas de natalidad variaron en menos de 1 por mil. La razón es que los puntos para edades jóvenes están muy concentrados en valores bajos, esto es, cerca de la intersección que determina el valor de b . Consecuentemente, la elección de diferentes puntos en las edades altas conduce siempre a una línea, que va a través del medio de ese conjunto de puntos. En relación con la proyección hacia atrás de la población de menos de 10 años, en la que se basan estimaciones de tasas de natalidad

² Coale y otros (1980, p. 2). Estos años fueron elegidos porque ambos censos fueron tomados en octubre.

utilizando exclusivamente el número de personas por debajo de la edad 10 en el momento del segundo censo, el procedimiento descrito aquí utiliza mucha más información sobre la distribución por edad. En principio usa información de las proporciones en todas las edades para determinar b ; en la práctica, los puntos para las edades altas, (esto es, valores altos de x y de y), tienen progresivamente menor influencia en la intersección, y sucesivamente mayor en la pendiente.

Cambios en el grado de cobertura de dos censos

Por razones similares, las estimaciones de la tasa de natalidad (y no así las estimaciones de la mortalidad adulta) son relativamente insensibles a cambios en el grado de cobertura de los dos censos. Para ilustrar este punto hemos reducido todas las tasas intercensales de crecimiento en la India en tres por mil, lo que es equivalente a un 3 por ciento de deterioro en la cobertura, en todas las edades, en el segundo censo en relación con el primero. La ecuación de la nueva línea de ajuste a los puntos 5-30 y 45-70, es:

$$y = 24,303 + 62,200x$$

lo que implica

$$b = \frac{1}{24,302} = 0,0411 \quad y$$

$$K = \frac{62,200}{24,302} = 2,559$$

La tasa de natalidad estimada se modifica solamente en uno por mil como resultado de este gran deterioro simulado en la cobertura de los censos. Sin embargo, el nivel de la mortalidad adulta aumenta sustancialmente, al haber más gente ausente en el segundo censo, que se presume ha muerto. El valor de la e_s en este caso simulado es de 47,2 años, comparado con la estimación original de 51,0.

Información inadecuada sobre migración

La no consideración de una adecuada estimación sobre migración internacional en las ecuaciones de estimación puede tener efectos diferentes, que dependerán de la incidencia por edad de la migración. Si el

error de la serie de emigración neta (que debe ser agregada al valor de $r(x)$) es invariable con la edad, entonces el efecto en las estimaciones es formalmente equivalente al efecto de una cobertura diferencial en la cabalidad de los dos censos. Como se mostró anteriormente, el efecto de las estimaciones de la tasa de natalidad será probablemente pequeño, pero podrá ser substancial en las estimaciones de la mortalidad adulta. Si la emigración es mayor de la considerada en los cálculos, la mortalidad adulta será falsamente exagerada, toda vez que los emigrantes se toman como muertos. Si la emigración es mayor de lo que se ha supuesto sólo en las edades adultas (digamos mayores de los 25-30 años), entonces únicamente la estimación de la mortalidad adulta estará afectada (resultará nuevamente exagerada). Si, en cambio, la emigración se subestima solamente en las edades más jóvenes (digamos por debajo de los 15), la tasa de natalidad estimada resultará muy baja, toda vez que la serie de valores de y estará falsamente elevada en un valor relativamente constante. En efecto, parte de la distribución por edades envejecida, que resulta de una emigración joven, aparece, en cambio, como resultado de una baja fecundidad. Un patrón de error en la emigración que se concentra en las edades 15-30, que es quizás el patrón que se encuentra comúnmente, debe tener relativamente poco efecto en la tasa de natalidad estimada, toda vez que el efecto en los valores de y se acumula y se vuelve substancial solamente en el centro del rango de puntos dibujados. Pero conducirá, de nuevo, a estimaciones exageradas de la mortalidad adulta.

Si la estimación de las tasas de emigración neta es demasiado alta (esto es, si las tasas de inmigración son demasiado bajas) entonces los resultados serán al contrario de lo que se ha descrito arriba.

Errores en la declaración de la edad y omisión diferencial de niños

Importantes fuentes de error en las estimaciones indirectas en países en desarrollo son la mala declaración de la edad y diferencias según la edad en las omisiones de los censos. El presente método es también vulnerable a estos problemas. Pero, una vez más, ellos parecen afectar más seriamente la mortalidad adulta que a las estimaciones de la tasa de natalidad. Uno de los problemas más serios para la mayoría de los métodos que estiman la tasa de natalidad —especialmente aquéllos que se apoyan en una proyección hacia atrás— pero también los procedimientos que usan la ojiva de la distribución por edades, es la diferente omisión de niños jóvenes en los censos. Para ilustrar los efectos de tal

omisión en el procedimiento que presentamos, hemos reducido arbitrariamente el número de niños en los grupos 0-4 en los dos censos de la India en un 10 por ciento y reestimado la línea (usando nuevamente los promedios de los grupos de edades 5-30 y 45-70) resultando,

$$y = 23,671 + 43,701x,$$

lo que implica

$$b = \frac{1}{23,671} = 0,0422 \quad y$$

$$K = \frac{43,701}{23,671} = 1,8462$$

Sorprendentemente, la tasa estimada de natalidad cambia solamente de 41,1/1000 a 42,2/1000. Por otra parte, en lugar de haberse reducido por la omisión de niños jóvenes, ha sufrido un ligero incremento. Lo que ha determinado que ocurriera esta ligera alza es que, cuando los niños con edades 0-4 son omitidos, las proporciones de todas las otras edades se incrementan, reduciéndose sus valores de y . En el agrupamiento de puntos cerca de la intersección uno aumenta su valor notoriamente en tanto que todos los otros lo reducen. La suma de los primeros seis valores de y de hecho declina y, como la intersección también declina, la tasa de natalidad aumenta³. Pero, en promedio, el conglomerado de puntos agrupados se mantiene más o menos en el mismo lugar. Esta insensibilidad contrasta con la muy considerable sensibilidad de la omisión de niños en las estimaciones basadas en las proyecciones hacia atrás. Si el 10 por ciento de niños en el grupo 0-4 se omitiera en forma diferencial en los censos, la proyección retrospectiva produciría un valor de b , basado en niños de 0 a 9 años, que sería mucho más bajo, en aproximadamente 5 por ciento.

Los errores en las declaraciones de edad, que son esencialmente aleatorios, agregarán dispersión a los puntos sin sesgar las estimaciones⁴.

³ La baja resulta del hecho que, de acuerdo con el procedimiento recomendado, el valor ${}_5\hat{N}_0$ aparece sólo en el cálculo de la $y(5)$ en tanto que ${}_5\hat{N}_x$ aparecerá en todas las otras edades en $y(a)$ para las dos edades x y $x+5$.

⁴ En particular, las estimaciones por el método de mínimos cuadrados no son sesgadas por errores en los valores de y , en tanto que ellos se distribuyan normalmente con una media igual a 0 y una varianza constante.

Transferencias de población a intervalos de edades adyacentes, de la misma manera, provocarán efectos menores. Sin embargo, la sistemática sobreestimación de la edad en edades avanzadas puede sesgar las estimaciones, particularmente de mortalidad adulta. Tal sobreestimación elevará los valores de y en las edades medias y maduras que estarán afectadas por evasiones, y disminuirá aquéllos en las etapas más altas, que reciban una falsa incorporación de población. Este patrón debe ser detectable por la diseminación de los puntos alrededor de la línea de ajuste con forma de loma. De hecho esto se observa en la información de la India del cuadro 1. Una estrategia de estimación puede simplemente consistir en no usar los puntos así afectados ya que, particularmente los valores muy altos de y y x , correspondientes a edades avanzadas, pesan fuertemente en los promedios por edades. Nótese que el patrón producido por la dispersión debida a la mala declaración de edad, aparece similar al que sería producido por un valor de B (el parámetro de la pendiente de la mortalidad adulta) menor a la unidad. Por esta razón, puede ser inapropiado el uso de este procedimiento de estimación para estimar la B .

Error en la mortalidad al comienzo de la vida

La sensibilidad del valor estimado de b , a errores en la estimación de $p^*(5)$ puede ser expresada muy simplemente: la estimación de b es inversamente proporcional a $p^*(5)$. Si $p^*(5)$ es exagerado en un 5 por ciento (un error muy grande en relación con una típica situación de incertidumbre acerca de $p^*(5)$), todos los valores de y aparecerán sobreestimados en un 5 por ciento. Tanto la pendiente como la intersección serán exageradas en la misma medida y la tasa estimada de natalidad subestimada en un 5 por ciento. Sin embargo, la estimación de la mortalidad adulta permanecerá inalterada, toda vez que K es el cociente entre la pendiente y la intersección, ambas distorsionadas por el mismo factor. A pesar que la estimación de b es sensible a la estimación $p^*(5)$, es afortunado que $p(5)$ sea uno de los parámetros demográficos más confiables observados en países en desarrollo, gracias al éxito de la técnica de Brass.

Error en el patrón de la mortalidad adulta estimada

Los valores de x utilizados al graficar la relación básica mostrados en los gráficos 1 y 2 son simplemente ${}_s q_s(a) / {}_s p_s(a)$. Si se elige como estándar un nivel equivocado de mortalidad, dentro del sistema logito de tablas de vida de un parámetro, ello debe tener poco efecto tanto en la

estimación de la tasa de natalidad, como en la estimación de la mortalidad adulta. Si la transformación de un parámetro es aplicable exactamente, los valores de x serán todos ellos incorrectos por un factor ϵ , el error en el supuesto nivel de mortalidad. Pero este error se reflejará exactamente en la estimación de K que se utiliza para corregir la tabla estándar inicial. La estimación final del nivel de mortalidad debe ser correcta, y la intersección no deberá ser perturbada por errores proporcionales en los valores de x .

Una dificultad más seria surge si se elige un patrón incorrecto de mortalidad adulta. A fin de ilustrar la sensibilidad de los resultados al uso de un patrón de mortalidad por edades muy diferente, Larry Heligman (de la División de Población de las Naciones Unidas) ha repetido el ejercicio de la India utilizando como tabla de vida estándar una de Naciones Unidas (1982), modelo de Asia del Sur. Este patrón es muy diferente de cualquiera de los contenidos en el conjunto de tablas de vida de Coale-Demeny. Presenta extremadamente baja mortalidad en el intervalo de edades 5-50 y muy alta mortalidad fuera de ese intervalo. Para el mismo nivel de $p(5)$ utilizado anteriormente para la India, el modelo de Asia del Sur tiene menor mortalidad que la del modelo Sur en todas las edades entre los 10 y 50 años y mayor mortalidad por arriba de los 50. La máxima diferencia entre las dos funciones ${}_s q_s(a) / {}_s p_s(a)$ alcanza a 0,515, lo que ocurre a la edad 50 (0,2823-0,2308). Utilizando el patrón de Asia del Sur como estándar resulta una estimación de la tasa cruda de natalidad de 41,4 (versus 42,1 utilizando la Familia Sur) y una estimación de $e_5^q = 52,7$ (versus 51,0 si se usa la Familia Sur); como podría haberse anticipado, la diferencia proporcional es mayor en e_5^q que en la tasa de natalidad. La estimación de la tasa cruda de natalidad se reduce utilizando el patrón de Asia del Sur porque, a un nivel dado de y , el valor de x se reduce más en las edades más bajas (excepto el correspondiente a la edad 5) que en las edades más altas. Esta distorsión tiene el efecto de elevar la intersección y bajar las tasas de natalidad estimadas. Si se experimentan con diferentes valores de B —el coeficiente de pendiente para la mortalidad adulta— también se producen cambios en la estimación de la tasa de natalidad. Con una $B = 0,841$ en el patrón de Asia del Sur, la tasa cruda de natalidad es 43,1 (comparado con 41,1, cuando $B = 1$), y e_5^q es igual a 50,6, comparado con 52,7. Aquí los cambios proporcionales en las tasas de natalidad y en e_5^q son casi los mismos.

Estimación de la verdadera distribución por edades

Es útil señalar que, si se aceptan las tasas de crecimiento intercensales por edad observadas, estamos implícitamente adoptando la “hipótesis de errores similares”, previamente utilizada por Coale, Demeny y Shorter (1968). Esto es, estamos suponiendo que han ocurrido errores proporcionales iguales en los dos censos en un grupo particular de edades. De manera que las tasas de crecimiento son correctas. Es posible estimar cuáles son esos errores mediante la reconstrucción de la “verdadera distribución por edades” y comparándola con la observada. Una estimación de la verdadera distribución por edades es probable que sea uno de los más importantes subproductos de este procedimiento, ya que no hay otro método para estimarla, a menos de hacer supuestos de estabilidad o adoptar un sistema dual de registro. Para estimar la verdadera distribución ${}_s\hat{c}_a$, podemos aceptar el valor estimado de b , las tasas observadas de crecimiento y el valor estimado de la mortalidad:

$${}_s\hat{c}_a = b e^{-\int_0^{a+2.5} r(x) dx} \frac{L_a}{l_0}$$

No es necesario que tales estimaciones de ${}_s\hat{c}_a$ sumen la unidad y, en varias aplicaciones, han diferido de ella en uno o dos por ciento. Para estimaciones finales de ${}_s\hat{c}_a$, tales estimaciones deben dividirse por la suma de las estimaciones preliminares. A pesar de que tal división parece constituir una revisión de b , debe interpretarse más propiamente como un factor requerido para adecuar las relaciones continuas a las que se han desarrollado en forma discreta.

RESUMEN

El sistema de estimación propuesto aquí se apoya en otros: en la estimación de la mortalidad adulta de Brass a través de la transformación logito de una tabla estándar; las estimaciones de la mortalidad al principio de la vida por el procedimiento ideado por Brass; y recientes generalizaciones de procedimientos relacionados con la población estable. Es nuevo sólo en la medida en que integra estos componentes en un procedimiento de un solo paso, para estimar la fecundidad y la mortalidad. Las ventajas de cálculo relativas, en comparación con la aplicación de procedimientos individuales, son posiblemente mayores en situaciones en las que los censos no están separados entre sí por 5 ó 10 años. En estas situaciones, los cálculos convencionales de la mortalidad adulta

pueden ser muy laboriosos, toda vez que las tasas de sobrevivencia no están normalmente tabuladas a intervalos de 5 a 10 años, y los ajustes o interpolaciones requeridos pueden introducir errores. El examen visual de la representación de los puntos observados, en los que se basa la estimación, constituye una ayuda para identificar los errores en la información. La simplicidad de los procedimientos facilita también el análisis de la sensibilidad de los resultados a errores de varios tipos en los datos. Los resultados de algunos análisis de sensibilidad sugieren que el procedimiento proporciona estimaciones relativamente robustas de tasas intercensales de natalidad. Comparte con otros métodos intercensales de estimación de la mortalidad adulta la sensibilidad en cuanto a diferencias en la cobertura de los censos y a problemas de mala declaración de la edad. Un subproducto importante del método es la estimación de la verdadera distribución por edades.

En tanto que el examen y resultados presentados aquí destacan la situación donde se dispone de una estimación de la mortalidad de la niñez, el procedimiento puede emplearse también cuando no se cuenta con ella. En este caso, la tabla estándar de mortalidad debe empezar a la edad 0, en lugar de empezar a la edad 5, y la ecuación (6) debe ser la que se utiliza. Hay varias razones para utilizar una estimación de mortalidad al principio de la vida si está disponible. La primera es, típicamente, una de las más confiables estimaciones demográficas derivadas en países en desarrollo. La segunda es que en ausencia de información sobre mortalidad en la niñez el valor de B adoptado, para propósitos de estimación, se aplicará a relaciones entre la mortalidad en la niñez y mortalidad adulta, y no simplemente a relaciones entre mortalidad de grupos de jóvenes y mortalidad adulta. Las relaciones entre la mortalidad en la niñez y la adulta son altamente variables (Coale-Demeny, 1966) y, por lo tanto, el procedimiento está sujeto a mayor error (apenas mitigado por la posibilidad de experimentar con valores diferentes de B). Además, errores en la información (tales como diferencias en el grado de cobertura de los censos), que previamente hubiera afectado solamente el nivel de la mortalidad adulta, ahora afectarán a niveles de la mortalidad al principio de la vida, con la cual la estimación de la tasa de natalidad está claramente ligada.

BIBLIOGRAFIA

Adlakha, Arjun, Dudley Kirk. 1974. Vital Rates in India 1961-71 Estimated from 1971 Census Data. *Population Studies* 28:381-400.

- Brass, William. 1971. On the Scale of Mortality. Pp. 69-110 in William Brass (ed.), *Biological Aspects of Demography*. London: Taylor and Francis.
- Carrier, N. H., John Hobcraft. 1971. *Demographic Estimation for Developing Societies*. London: Population Investigation Committee, London School of Economics.
- Coale, A. J., Paul Demeny. 1966. *Regional Model Life Tables and Stable Populations*. Princeton, N.J.: Princeton University Press.
- Coale, A. J., Lee-Jay Cho, Noreen Goldman. 1980. *Estimation of Recent Trends in Fertility and Mortality in the Republic of Korea*. Committee on Population and Demography Report No. I. Washington, D.C.: National Academy of Sciences.
- Demeny, P., F. C. Shorter. 1968. *Estimating Turkish Mortality, Fertility, and Age Structure*. Istanbul: Istanbul University Press.
- Dyson, Timothy. 1979. *A Working Paper on Fertility and Mortality Estimates for the States of India*. Paper presented to the Panel on India, National Academy of Sciences Committee on Population and Demography. Washington, D.C.
- India, Registrar General. 1977. *Census of India, 1971. Series 1. Paper 1 of 1977. Life Tables*. New Delhi.
- National Academy of Sciences. 1981. *Manual of Indirect Estimation Techniques*. Unpublished manuscript, Washington, D.C.
- Preston, Samuel H., Neil Bennett. 1982. *A Census-Based Method for Estimating Adult Mortality*. Population Studies Center Research Report No. 81-16. Ann Arbor: University of Michigan. (Population Studies, in press.)
- Preston, Samuel H., Ansley J. Coale, 1982. *Age Structure, Growth, Attrition, and Accession: A New Synthesis*. *Population Index* 48: 217-259.
- United Nations. 1980. *Population and Vital Statistics Report*. April edition. ST/ESA/STAT/SER.A/132. New York: U.N. Department of International Economic and Social Affairs.
- United Nations. 1982. *Model Life Tables for Developing Countries*. Population Study No. 77. New York: U.N. Department of International Economic and Social Affairs.

RECONOCIMIENTO

Los comentarios de Larry Heligman fueron importantes en el mejoramiento de este documento. Ansley Coale, Griffith Feeney, Mari Bhat, Shanta Danaraj, Duanpon Theerawanviwat, y los *referees* también hicieron comentarios útiles. Esta investigación fue apoyada por una subvención de NICHD HD 1P30 Hd10379.

NOTAS DEL EDITOR

El señor Shiro Horiuchi ha tenido la gentileza de enviarnos una lista de artículos recientemente publicados cuyo tema se vincula con el del presente número de NOTAS DE POBLACION. Es la siguiente:

“A census-based method for estimating adult mortality”, por Samuel H. Preston y Neil G. Bennett, publicado en *Population Studies*, 37 (1983);

“Mortality estimation from registered deaths in less developed countries”, por Neil G. Bennett y Shiro Horiuchi, publicado en *Demography*, 21 (1984);

“Use of direct and indirect techniques for estimating the completeness of death registration systems”, de Samuel H. Preston, publicado en *Data Bases for Mortality Measurement* (Naciones Unidas, número de venta E.83, XIII.3);

“How do age-specific growth rates reflect the impact of past history upon the current age structure?”, de Shiro Horiuchi, presentado a la reunión anual de 1984 de la Population Association of America.

El Editor de la Revista cumple con el deber de agradecer a los autores de los artículos presentados en este número por su autorización para la publicación, agradecimientos que se hacen extensivos a los editores de las revistas POPULATION INDEX y DEMOGRAPHY, que han permitido su reproducción.