

permiten a los matemáticos interesados seguir los hilos y sin duda, mejorar y desarrollar la idea más adelante.

II. ERRORES EN LOS DATOS

En la práctica los datos están sujetos a errores, y cuando la mayor parte de los datos consiste en defunciones por edad, los errores pueden ser grandes. Debido a que los datos pueden ser obtenidos en múltiples formas, no será posible detallar todas las fuentes de error que deben o pueden ser tenidas en cuenta. Algunas indicaciones generales, pueden sin embargo, servir como advertencia.

- Defunciones Infantiles. Una proporción sustancial de los niños mueren muy cerca del nacimiento. Por una variedad de razones y en una variedad de formas, esto lleva a que se trate las muertes infantiles de modo diferente a las muertes de edades mayores, tanto en lo que respecta a la "distribución de las muertes" y al registro del hecho.

De ahí que los datos que dan adecuada representatividad a las defunciones de las edades viejas, o al menos iguales deficiencias en todas esas edades, pueden verse afectadas por las excesivas deficiencias en las muertes infantiles. Será entonces necesario desconocer todo evento anterior al primer cumpleaños, y tratar la población como si comprendiera sólo aquellos de más de un año y por lo tanto definir población que "entra" como sigue: No un nacimiento, sin la celebración del primer cumpleaños.

- Edad: el significado de "edad" no debe necesariamente seguir la definición: "número de años completos desde el nacimiento". Aún cuando ésta sea la intención básica, el registro puede ser incorrecto por una variedad de razones como la ignorancia del informante o algún incentivo para declarar mal la edad. La sobre declaración de la edad en los adultos es muy común, pero puede ser solucionada combinando los datos para los grupos más viejos, por ejemplo en un grupo de 65 y más. La tendencia a redondear la edad, especialmente hacia la edad más próxima terminada en 0, se encuentra también con frecuencia. Su influencia pueda ser grandemente reducida combinando los datos en grupos desde una edad terminada en 5 a una terminada en 4; ejemplo 45 - 54 o 45 - 64. Grupos entre edades terminadas en 0 y 9 deben ser evitados.

Los datos para nuestro problema, pueden derivarse de muy variadas y curiosas fuentes. La importancia de tener cuidado y sentido común para evitar errores excesivos en esas circunstancias, no puede ser suficientemente subrayada.

III. TEORIA GENERAL

Como la teoría matemática de la población estable está descrita adecuadamente en otra parte, un breve resumen es suficiente aquí. Una población estable es aquella que se deriva de tasas de mortalidad y fecundidad específicas por edad constantes. Debido a que debemos condensar nuestros datos en amplios grupos de edades, aceptamos fluctuaciones de año a año con facilidad, que proporciona en promedio, es decir, para décadas se mantenga razonablemente constantes. En verdad, si nuestros datos se refieren a un período largo, es posible aceptar aún casi cualquier fluctuación periódica, con tal de que no haya una tendencia a largo plazo subyacente. Puede demostrarse que bajo estas condiciones, existe un número r (tasa intrínseca de crecimiento natural) tal que el número de nacimientos de cada año excede al del año anterior en $100 r \%$, es decir, tal que los nacimientos de un año son numéricamente iguales a los del año anterior multiplicados por $(1 + r)$. De otro modo, los nacimientos de un año son aproximadamente iguales a los del año siguiente multiplicados por $(1 - r)$. Para explicar la teoría será conveniente suponer que las defunciones dadas, se refieren a un solo año calendario. Denotemos a los nacimientos de ese año por B . Los del año anterior serán $B(1-r)$, del año anterior $B(1-r)^2$, etc.

La mortalidad de esta población es constante, por hipótesis, y puede ser definida por una serie de las funciones de la tabla de vida. De ahí que, para cada una y toda generación que comprende la población en el año cuyas defunciones son dadas, una fracción l_x/l_0 ha sobrevivido desde el nacimiento hasta su x cumpleaños, l_{x+1}/l_0 hasta su $x + 1$ cumpleaños y así $\frac{l_x}{l_0} - \frac{l_{x+1}}{l_0}$ ($= \frac{d_x}{l_0}$) morirán entre el x y $x + 1$ cumpleaños.

Los que mueren a la edad x (último cumpleaños) en un año específico, pueden ser tomados como pertenecientes a la generación que nació x años antes, cuando los nacimientos eran $B(1 - r)^x$. Se ha demostrado que una fracción d_x/l_0 morirá a la edad x , luego el Número de muertos a la edad x será $B(1-r)^x d_x/l_0$. Este número, que se simbolizará con D_x es dado. (En la práctica los datos, estarán, por supuesto, en grupos de edades no en edades simples). Luego nuestra ecuación fundamental es
$$\frac{B(1 - r)^x d_x}{l_0} = D_x$$

Debido al hecho, que las muertes dadas no se refieren a un año calendario individual, sino a un número de años no especificado, en nuestra ecuación B , debe referirse también a un número no especificado de años. De ahí que pierda toda

interpretación práctica y se convierta simplemente en una constante desconocida que debe ser eliminada.

Supongamos que conocemos el valor de r . Luego para cada edad x , podremos calcular $(1-r)^x$, dividir las D_x y luego evaluar el lado izquierdo de la ecuación:

$$\frac{D_x}{(1-r)^x} = \frac{B d_x}{l_0}$$

A una edad, digamos 90, la vida cesa. Luego $l_{90}=0$, $l_{89}=d_{89}$, $l_{88}=d_{88}+d_{89}$, etc., es decir l_x se obtiene si d_x , d_{x+1} , etc., se suman hasta la edad más vieja. En particular l_0 se obtiene de sumar todas las d_x .

Luego

$$\sum_{x=0} \frac{D_x}{(1-r)^x} = \frac{B}{l_0} (d_0 + d_1 + d_2 + \dots) = \frac{B l_0}{l_0} = B$$

l_0 puede ser elegido arbitrariamente, digamos sea 10,000.

Entonces

$$\sum_{x=0} \frac{D_x}{(1-r)^x} = \frac{B}{l_0} (d_x + d_{x+1} + \dots) = \frac{B l_x}{l_0} = \frac{\sum_{x=0} \frac{D_x}{(1-r)^x} l_x}{l_0}$$

y así

$$l_x = \frac{l_0 \sum_x \frac{D_x}{(1-r)^x}}{\sum_0 \frac{D_x}{(1-r)^x}}$$

Cualquier otra función puede ser derivada, dados todos los valores de l_x .

IV. APLICACION PRACTICA

En la práctica es esencial agrupar los datos en grandes grupos de edades aunque sea para tener números suficientemente grandes para obviar fluctuaciones debidas al azar. Más importante aún, algunos de los errores en la declaración de la edad, mencionados arriba, involucran transferencias de defunciones de una edad a otra y son eliminadas si las defunciones se agrupan para incluir en el mismo grupo de edades las edades entre las cuales se transfieren las defunciones.

Supongamos un grupo típico que comprende las defunciones entre las edades exactas "a" y "b". Es decir:

$$\sum_x^{b-1} D_x = D_{a,b}$$

Luego correctamente podemos dividir algunas muertes del grupo por $(1-r)^a$, y otras por $(1-r)^{b-1}$ y otras por los otros valores intermedios. En la práctica tendremos que dividir el grupo entero por un valor promedio que se denotará por $(1-r)^{a,b}$.

El problema de estimar el apropiado valor medio se analiza en el Capítulo I. Se encuentra que, para muchos propósitos prácticos, una estimación suficientemente correcta, se obtiene por $(1-r)^{a+b-1/2}$ lo que equivale a decir que si las $D(a,b)$ se refieren a muertes de personas entre, digamos, el 20 y 45 cumpleaños (es decir, $a = 20, b = 45$), luego $(1-r)^{20,45}$ puede estimarse como $(1-r)^{20+45-1/2} = (1-r)^{32}$.

Este procedimiento no es enteramente satisfactorio para las edades más jóvenes y para las más viejas.

Para el grupo de edades 1-4, el error de tomar $(1-r)^{1+4/2} = (1-r)^{2\frac{1}{2}}$ no es muy grande, pero un mejor resultado se obtiene si $(1-r)^2$ es usado. Pero para grupos de edades más amplios que comienzan a la edad 1, se introduce errores sustanciales de 5 por ciento o más, si se usa la fórmula anterior. Como una aproximación burda y sencilla, la siguiente regla conducirá a obtener errores sustancialmente menores.

Grupo de edad	1-4	1-9	1-14	1-19
Estimación de $(1-r)^{a,b}$	$(1-r)^2$	$(1-r)^3$	$(1-r)^4$	$(1-r)^5$

Los datos para los grupos de edades 1-4, así como de muertes infantiles, pueden ser considerados demasiado deficientes para ser usados, y el estudio puede comenzar a la edad 5. La fórmula $(1-r)^{a+b-1/2}$ puede ser usada para los grupos de edad 5-9, pero un mejor resultado se obtiene si el grupo se extiende a 5-19.

Para los últimos grupos de edades, en la fórmula $(1-r)^{a+b-1/2}$ puede tomarse arbitrariamente b como 90. Esto lleva a errores de 2 por ciento en algunos casos, pero es satisfactorio si el último grupo de edades se tomó como 45 y más u 80 y más.

Para grupos de 50 y más, y 75 y más es mejor tomar $b = 88$ y para los restantes grupos posibles, 55 y más, 60 y más etc. tomar $b = 86$. Otro método más complejo de estimar $(1-x)^{a,b}$ se describe en el Apéndice I.

Un ejemplo calculado con datos hipotéticos puede aclarar el procedimiento. Suponemos que los datos son:

Muertes de	1 - 14	1 573	(Estas muertes pertenecen a una población estable construida para años individuales de edades en la Tabla de Vida Inglesa N° 3 (E.L.T ₃), Hombres, con $r = 0,01$).
	15 - 44	1 304	
	45 y más	2 543	

Se supone también que $r = 0,01$ es conocida. Los primeros cálculos son como sigue:

Edad	D _{a,b}	Estimación de $(1-x)^{a,b}$ con $r=0,01$	$\frac{B}{l_0}(1-x)^{a,b}$	Suma col(4) acumulada desde abajo	$l_x(1-x) = 10\ 000$ Col(5) · 10 000/8 378
(1)	(2)	(3)	(4) = (2)/(3)	(5)	(6)
1-14	1 573	$(1-x)^4 = 0,9606$	1 638	8 378	10 000
15-44	1 304	$(1-x)^{\frac{15+44}{2}} = 0,7434$	1 754	6 740	8 045
45 y más	2 543	$(1-x)^{\frac{45+90-1}{2}} = 0,5100$	4 986	4 986	5 951

Si de hecho las defunciones están disponibles subdivididas en grupos de edades más pequeños, la última columna de la tabla anterior mostraría valores de l_x para valores de x menos espaciados y otras funciones de la Tabla de Vida podrían haber sido deducidos directamente de estos valores de l_x . En la práctica, por las razones discutidas anteriormente, aún estuvieran disponibles las defunciones en subdivisiones grandes, sería necesario combinar los datos en grupos

ejemplos para reducir el error. De ahí que en todo caso, los valores que probablemente se utilicen son sólo unos pocos valores de x . Las tablas modelo de las Naciones Unidas proporcionan el medio de solucionar esta dificultad. Toda una serie de Tablas de Vida se da con un rango de $e_0^0 = 20$ a $e_0^0 = 73,9$. Los valores de l_x ($l_0 = 100\ 000$) se presentan para "niveles" con e_0^0 de intervalos $2\frac{1}{2}$ años. Por ensayos y errores se encuentra que los niveles 30, 35, 40 son apropiados para los resultados anteriores. Luego los cálculos se hacen como sigue:

Nivel	$l_0 = 100\ 000$			$l_1 = 10\ 000$		
	e_1^0	l_1	l_{15}	l_{45}	l_{15}	l_{45}
30	44,1	77 535	62 778	42 154	8097	5437
35	46,4	79 075	65 641	45 918	8301	5807
40	48,7	80 427	68 147	49 425	8473	6145

Los valores de l_x calculados a partir de las defunciones dadas pueden variarse y contrastar con los modelos de Naciones Unidas. $l_{15} = 8045$ está cerca del nivel 30, con una $e_1^0 = 44,1$ y $l_{45} = 5951$ está alrededor de la mitad entre los niveles 35 y 40, correspondiente a un valor de 47,5 para e_1^0 . El verdadero valor para $E.L.T_3$, en los que se basaron los datos es 46,65.

La razón por la cual diferentes niveles se igualan o dispersan con l_{15} y l_{45} es que la forma de la mortalidad en $E.L.T_3$ es diferente de los modelos de Naciones Unidas. La $E.L.T_3$ muestra mortalidad relativamente alta en las edades jóvenes y relativamente baja en las edades viejas o adultas. Sin embargo, los datos en los que se basan $E.L.T_3$ es conocido que son en cierta medida erróneos y por lo tanto, no es claro hasta qué punto la aparente desviación del modelo de Naciones Unidas es real. Probablemente todos los datos de mortalidad alta, son hasta cierto punto sospechosos de modo que las verdaderas variaciones en esa mortalidad no son conocidas.

A pesar de estas dificultades, los valores de e_1^0 por ejemplo, sugeridos por esta confrontación con los modelos de Naciones Unidas, están bastante cerca, y sin duda, otras funciones de la Tabla de Vida pueden ser igualmente obtenidas por este método.

La técnica empleada hasta ahora ha requerido que el valor de r sea conocido. En algunas circunstancias, el problema de determinar r , por alguna otra fuente puede no ser difícil. Por ejemplo, dado un período estable en el crecimiento de la población que se extiende por varias décadas, índices relativamente erróneos del tamaño de la población cerca del comienzo y del fin del período, llevaron a un menor error en el valor de r que si los dos índices se refirieran a períodos más cercanos en el tiempo.

Índices adecuados del tamaño de la población pueden ser: fuerza militar, número de contribuyentes, área cultivada o número de ciudades, aunque la posibilidad de ser inducido a error por estos índices es evidente. Una aproximación alternativa puede ser observada en otras poblaciones que viven bajo circunstancias similares para las cuales " r " es conocida.

Antes de seguir adelante, sin embargo, debemos determinar cuán sensible es nuestra selección de una Tabla de Vida apropiada a los errores en el valor de r estimado o supuesto. Este estudio, al mismo tiempo, dará una idea acerca del error que se puede introducir por la consideración arbitraria de que la población es estacionaria, es decir, que $r = 0$.

Usaremos los datos del ejemplo anterior, en el cual el valor de r es actualmente conocido e igual 0,01. Tomaremos, los supuestos alternativos de que $r = -0,01$ ó $r = +0,01$.

Los resultados pueden compararse con los valores correctos de l_{xx} para E.L.T.₃.

Es evidente que un error sustancial aparece en l_{xx} a menos que " r " sea tomado cercano de su valor correcto. En particular, si la población se supone arbitrariamente estacionaria ($r=0$) el error sería sustancial a menos que el valor de r sea bastante pequeño.

Función l_x de la Tabla de Vida

Porcentaje de error en la estimación de l_x , con $r =$

Edad exacta	Valor correcto (E.L.T. ₃)	Valor estimado del r supuesto			Porcentaje de error		
		-0,01	0	0,01 (valor correcto)	-0,01	0	0,01 (valor correcto)
10	10 000	10 000	10 000	10 000			
15	8 044	6 011	7 098	8 045	-25,3	-11,8	-0,01
45	5 927	3 446	4 962	5 951	-41,9	-20,8	0,4

Los valores anteriores de l_x pueden convertirse en valores de e_x , por conversión con tablas de vida modelo de Naciones Unidas como se indica en los siguientes resultados:

	Valor correcto	Valores de e_x		
		Deducidos cuando r vale		
		-0,01	0	0,01 (valor correcto)
deducido de l_{15}	46,65	Alrededor de 27*		
deducido de l_{45}	46,65	32	39	47,5

* Estimado. Fuera de la serie los modelos de Naciones Unidas.

LA ELIMINACION DE LA DEPENDENCIA DE r

Se ha mostrado arriba que la técnica que aquí se discute, la cual requiere un valor estimado de r , dará resultados razonablemente correctos sólo si r se estima dentro de 1 ó 2 por ciento. Algunos aspectos del problema de estimar r han sido discutidos, pero parecería ser que se obtiene alguna ventaja si se puede obtener el resultado sin tener que estimar r , a partir de datos extremos.

En el Apéndice II se presenta la teoría matemática de una técnica para encontrar el apropiado nivel de las tablas modelo de Naciones Unidas que concuerdan con los datos sin requerir previamente una estimación de x , más aún el método incluye la evaluación de r . Esencialmente la técnica equivale a tratar la técnica discutida anteriormente con varios valores de r , seleccionando el valor de r que lleve al mismo modelo de Tablas de Vida que concuerde con ambos valores de l_x calculados y en consecuencia seleccionando la Tabla de Vida modelo contrastada.

Si toda experiencia en mortalidad se adaptara al patrón de la Tabla de Vida de Naciones Unidas, por ejemplo, si cada experiencia se adecuara exactamente a uno de los modelos de Naciones Unidas, esta técnica daría siempre el resultado correcto. Se ha mostrado que los E.L.T.₃ aparentemente tienen un patrón algo diferente. Esta discrepancia puede derivarse de errores en los datos básicos, pero por otro lado esta particular desviación de los patrones de Naciones Unidas concuerdan con el patrón común. Las tablas de Naciones Unidas representan el promedio de un gran número de Tablas de Vida nacionales, y todas se refieren al presente siglo, tres cuartos del siglo a partir de 1920. Por eso, predominantemente es moderno, cuando por ejemplo: el control de las muertes por infección había sido superado. En contraste E.L.T.₃ se basa en la mortalidad inglesa de un siglo atrás y puede esperarse que tenga una mortalidad infantil y juvenil más alta, debido a la alta incidencia de las muertes por infección.

Algún trabajo experimental ha mostrado que cuando los E.L.T.₃ se compararen con los modelos de Naciones Unidas por medio de las técnicas del Apéndice II, muy curiosos e incorrectos resultados se obtuvieron. El punto es que las técnicas del Apéndice II son críticas con respecto a la forma de la mortalidad involucrada.

Si el factor importante en los cambios de los patrones de mortalidad es la extensión de los controles de muertes por infección, entonces tal vez las técnicas del Apéndice II puedan aplicarse satisfactoriamente a experiencias corrientes medidas a través de los modelos de Naciones Unidas y también a experiencias pasadas, si esto se mide con respecto a modelos modificados que se conforman más a los E.L.T.₃; pero esta hipótesis no puede ser aceptada a ojos cerrados.

Un usuario no matemático necesitará tener en cuenta si sus datos requieren el empleo de grupos de edades particulares desde el momento que los grupos de edades determinan la fórmula a usar. Algunos ejemplos con una serie particular de grupos de edades pueden dar alguna luz.

Si los grupos de edades son 1-14, 15-44, 45 y más, los datos (D_{1-15} , D_{15-44} y D_{45-90}) se trabajan como sigue:

- 1) Calcular $A = D_{1-15} / D_{15-44}$; $B = D_{45-90} / D_{15-44}$ y $A^3 B^2$
- 2) Determinar por ensayo y error la tabla de Naciones Unidas tal que:

$$K = (1 - 1_{15}) / (1_{15} - 1_{45}) \quad \text{y} \quad L = 1_{45} / 1_{15} - 1_{45}$$

tal que $K^3 L^2 = A^3 B^2$. Esta es la Tabla de Vida apropiada.

- 3) Por logaritmos determinar r a partir de la ecuación

$$(1-r)^{25,5} = K/A$$

Para defunciones masculinas en Ceilán 1937, el trabajo es:

$$D_{1-15} = 13,002; \quad D_{15-44} = 13,054; \quad D_{45-90} = 18,879$$

$$A = 0,9660; \quad B = 1,446; \quad A^3 B^2 = 2,07$$

Para la Tabla de Vida de Naciones Unidas tenemos:

Nivel	40	45
$K^3 L^2$	1,97	2,25

Interpolando 2,07 entre 1,97 y 2,25 da el nivel 42 ($e_0 = 41$ años).

$$(1-r)^{25,5} = 0,6485$$

$$1-r = 0,9831$$

$$r = 0,017 \quad (17 \text{ por mil})$$

Las tasas brutas de mortalidad y natalidad de Ceilán en 1930 eran del orden de 21,22 y 37,38 respectivamente para un valor $r = 0,016$. Esto no es garantía que r haya mantenido su valor en la centuria pasada que es el período relevante para establecer una población estable hacia 1937.

Otro ejemplo "satisfactorio" lo dan las defunciones de hombres en Chile en 1936, como sigue: $A^3B^2 = 1\ 366$, K^3L^2 es 1 547 para el nivel 25 ($e_0^0 = 32,5$) y 1 530 para el nivel 30 ($e_0^0 = 35,0$), $r = 0,0064$. Las tasas brutas de mortalidad y natalidad de 30-35 por ciento y 25 por ciento, sugieren un r algo mayor. Las Tablas de Vida chilenas dan una $e_0^0 = 35$.

En muchas pruebas -algunas se dan en el Apéndice II- los valores de A^3B^2 no se igualaron con ninguna de las Tablas de Vida modelo, es decir, ello era francamente imposible. El error se puede deber a:

- 1) A que la población no sea estable (o porque la mortalidad y fecundidad cambiaron o porque hubo migración)
- 2) Al número falso de defunciones
- 3) El nivel de la mortalidad distinto de aquel de los modelos de Naciones Unidas

DETERMINACION DE OTRAS CARACTERISTICAS DE LA POBLACION

Si las técnicas ampliamente discutidas aquí, en un caso particular, establecen satisfactoriamente la tabla de vida que define la mortalidad de este caso y el valor de r , otras características pueden obtenerse rápidamente. Por técnicas ya conocidas se puede determinar la estructura por edad de la población, por lo tanto, las muertes anuales en cada edad por mil en todas las edades, y por suma, la tasa bruta de mortalidad. La tasa bruta de natalidad excede a ésta en r .

Si tenemos la información adicional del período, es decir para n años calendario, a los cuales se refieren los datos originales, la división por n de las muertes actuales para todas las edades (o para una edad particular) dará el número anual de muertes.

Conociendo éste por la tasa bruta de mortalidad (o por una tasa específica por edad) se obtendrá la población de todas las edades (o por una edad específica). Conociendo la estructura por edad de la población, se puede determinar todos los elementos de la población y, a partir de la tasa bruta de natalidad, los nacimientos anuales.

La extensión de estos cálculos para determinar una u otra de estas características, serán de un valor relevante, si las características pueden estimarse también a partir de otra fuente, puesto que esto provocará una manera de controlar esencialmente el manejo de datos demográficos más finamente.

16 de septiembre, 1971.

APÉNDICE I

El Valor Promedio de $(1-r)^x$

La ecuación fundamental es $D_x = (1-r)^x d_x$ en que x es la edad en el último cumpleaños, r la tasa verdadera de crecimiento natural aproximada por la tasa anual de crecimiento, D_x el número dado de defunciones observadas y d_x las defunciones de la tabla de vida $(l_x - l_{x+1})$. Dados D_x y r , la ecuación precedente se resuelve como: $d_x = D_x / (1-r)^x$. En la práctica los datos estarán agrupados. Simbolizando por $D_{a,b}$ las defunciones en un grupo de edades típico entre las edades exactas a y b , se tiene: $D_{a,b} = \sum_{x=a}^{b-1} (1-r)^x d_x$. Esto se resuelve para $\sum_{x=a}^{b-1} d_x$ dividiendo $D_{a,b}$ por el valor promedio ponderado apropiado de $(1-r)^x$. Designando este promedio por $(1-r)^{a,b}$ se tiene:

$$\frac{D_{a,b}}{(1-r)^{a,b}} = \sum_{x=a}^{b-1} d_x$$

Entonces $D_{a,b}$ entonces $(1-r)^{a,b} = \sum_{x=a}^{b-1} (1-r)^x d_x / \sum_{x=a}^{b-1} d_x$.

Esto da el valor exacto de $(1-r)^{a,b}$, pero no puede usarse para evaluar $(1-r)^{a,b}$, puesto que las d_x son desconocidas. Lo que se requiere es una estimación aproximada de $(1-r)^{a,b}$ en términos de cantidades conocidas.

Para evitar dificultades posteriores en la interpretación de las inconsistencias, resulta necesario explicar desde un comienzo la precisión de diversas estimaciones de $(1-r)^{a,b}$ y, si es posible, idear por lo menos una técnica capaz de producir estimaciones de gran exactitud. Se dispondrá luego de ésta como último recurso, aunque en todas las situaciones prácticas es que resulte adecuada una exactitud menos estricta de precisiones estimaciones menos complejas de exactitud suficiente.

En la construcción de los modelos para examinar el tópico principal de la investigación no se considerará un detalle más refinado que las defunciones por años individuales, pues en teoría este tratamiento ya introduce un error. Sin embargo, el doble problema de la magnitud general del error en el tratamiento por años individuales y por grupos quinquenales de edades puede resolverse en una sola investigación. Parece evidente que el agrupamiento de las defunciones en grandes grupos de edades involucra un error más grande que lo que sucede con los grupos de menor amplitud. Por lo tanto, si el error en el tratamiento por grupos quinquenales de edades se examina partiendo de un modelo construido en años individuales, no se introducirá una distorsión significativa, y el error medido indicará en debida forma el error de los grupos quinquenales de edades y a un mismo tiempo, como límite

mediana bruta, el error de los años individuales.

3. Para el estudio del error de los grupos quinquenales de edades, se requiere una tabla de vida de mortalidad alta en años individuales de edad, y la Tabla de Vida Inglesa 3 - Hombres (I.V.I. 3) resulta apropiada.^{2/} Se excluirá el primer año de vida, porque su tratamiento resultaría complejo, y además un derroche de esfuerzo, ya que es probable que los datos del tipo que estamos considerando serán tan deficientes en defunciones infantiles que será necesario excluirlas de todos modos. Se estudia el resto de los primeros cinco años de vida, reducido ahora a un grupo de cuatro años, aunque la deficiencia de datos puede también exigir su exclusión. Los grupos quinquenales de edades estudiados han sido seleccionados entre aquellos en que d_x presentaba la inclinación más pronunciada o la mayor curvatura. Para finalizar el rango de edades se ha incluido el grupo de 85 y más. La tasa de crecimiento r supuesta fue del uno por ciento anual, cerca de su límite superior a lo largo de un período prolongado. Para simplificar el cálculo, se omitió, el factor común $(1 - r)^a$ para el rango de edades (a, b) .

Se han puesto a prueba dos procedimientos de estimación, la estimación (1) $(1 - r) \frac{a + b - 1}{2}$, es decir el valor de $(1 - r)^x$ a la edad media en el intervalo,^{3/} y $(1 - r)^c$ en que $(c + \frac{1}{2})$ es la edad media de los fallecidos en el intervalo.^{4/} Se verá que la estimación (2) es extremadamente precisa a lo largo de todo el rango, lo que sugiere que si se requiere un alto nivel de precisión en cualquier comprobación, puede ser utilizada con confianza. Ya se ha sugerido que el error en estas fórmulas varía según el cuadrado del ancho de la franja de edades. En casi todas las edades el error es tan pequeño que incluso con un grupo de edades con un ancho de 20 o 30 años, el error en la fórmula será cuando más de un uno por ciento. Sin embargo, para su aplicación práctica, en oposición a la comprobación, se producirá un error adicional del uso de esta fórmula, especialmente por desconocimiento de la edad media correcta de los fallecidos. Por lo tanto será necesario examinar más tarde la variación de la edad media de los fallecidos en una serie de tablas de vida.

Los errores en la estimación (1) son mucho más considerables en las edades menores de 10 o mayores de 80, pero de otro modo son escasamente más grandes que los de la estimación (2). La mayor facilidad para usar esta fórmula, y el hecho que no requiere el uso de una tabla de vida de reemplazo, implican que en general puede preferirse esta fórmula.

Grupo de edades (a, b-1)	$b-1$ $\sum_a^b d_x$	$b-1$ $\sum_a^b d_x (1-r)^x$	Valor [†] correcto $(1-r)^{a_2} b$ [(Col. (3))-(Col. (2))]
(1)	(2)	(3)	(4)
1-4	11,269	11,166.40	0.990,895
5-9	3,386	3,331.85	0.984,083
10-14	7,708	7,675.63	0.981,048
15-19	2,688	2,042.97	0.978,434
60-64	6,174	6,047.95	0.979,866
65-69	7,110	6,965.97	0.979,763
70-74	7,541	7,391.95	0.980,265
75-79	6,774	6,645.74	0.981,066
80-84	4,736	4,651.86	0.982,234
85 y más	3,298	3,193.24	0.968,235

* $(C + \frac{1}{2})$ es la edad media exacta de los fallecidos en (a, b).
† Todos los valores divididos por $(1-r)^a$.
‡ Suponiendo que 90 es la edad última.

Estimación(1) $\frac{a+b-1}{2}$ (1 - r)	Estimación(2) ^h (1 - r) ²	Error Porcentual	
		Estimación (1)	Estimación (2)
(5)	(6)	(7)	(8)
0.905,038	0.991,892	-0.59	0.10
0.980,100	0.984,642	-0.40	0.06
"	0.979,706	-0.10	-0.14
"	0.979,214	0.17	0.08
"	0.979,116	0.05	-0.05
"	0.979,608	0.04	-0.01
"	0.980,297	-0.01	0.01
"	0.990,987	-0.10	-0.01
"	0.982,191	-0.22	0
" $\frac{1}{2}$	0.986,051	1.21	-0.02

9. En la práctica, es probable que se empleen grupos de edades más amplios que los de 5 años: la importancia de los grupos más limitados radica principalmente en la construcción de modelos para fines experimentales.

Al combinar los datos que aparecen en el cuadro de la Sección 3 con grupos de edades más amplios, emerge lo siguiente:

a) Edades jóvenes. El error en la estimación (1), que para el grupo de edades (1 - 4) fue de más de $\frac{1}{2}$ por ciento, empeora constantemente si el grupo se amplía en forma progresiva a 1 - 9, 1 - 14, y 1 - 19.

Edad...	1-4	1-9	1-14	1-19
Error porcentual en la estimación (1)...				-0.59	-2.06	-3.72	-4.85

Para todos estos grupos la estimación (2) presenta un error menor al uno por mil. Si, no obstante, se omite el grupo 1 - 4, el error en un grupo que comienza a la edad de 5 puede controlarse ampliando el grupo hasta la edad de 19.

Edad...	5-9	5-14	5-19
Error porcentual en la estimación (1)...				-0.40	-1.16	-0.17

b) Edades avanzadas. Al final de la vida, el error considerable en la estimación (1) para cada uno de los grupos 30-34 y 35 y más se reduce a -0,13 por ciento si se combinan los grupos. Con una combinación adicional de los grupos, el error de la estimación (1) crece de -1,13 por ciento para el grupo de 75 y más a -1,94 por ciento para el de 65 y más, pero de allí en adelante disminuye, a -1,04 por ciento para 50 y más, y a -0,45 para 45 y más. (Para todos estos cálculos el cumpleaños 90 se tomó arbitrariamente como el final de la vida). La necesidad de compresión en las edades más avanzadas puede perfectamente manifestarse de todos modos a causa de una cierta exageración de la edad en esta región. Puede disminuirse el error mediante el supuesto de una edad final levemente más baja.

c) Edades intermedias. Para la región intermedia, es decir entre 20 y 44, el procedimiento de la estimación (1) es satisfactorio para cualquier grupo de edades. Incluso si se amalgama todo el grupo, el error es de sólo un 0,39 por ciento.

Resumen

Denótese $\frac{b-a}{2} d_x (1-r)^{\frac{a+b}{2}}$ por $(1-r)^c$, en que d_x es la función de la tabla de vida $l_x - l_{x+1}$ y r es la tasa real de crecimiento natural, o una aproximación a ella, la tasa anual de crecimiento. Se consideraran dos fórmulas para estimar $(1-r)^{a,b}$. La estimación (1) = $(1-r)^{\frac{a+b-1}{2}}$ y la estimación (2) = $(1-r)^c$ en que $(c+\frac{1}{2})$ es la edad media de los fallecidos en el intervalo (a,b) . Se encuentra que la estimación (2) es sumamente exacta en todo el tramo de edades entre el primer cumpleaños hasta el fin de la vida (el primer año no se considera). Entre los casos considerados el error rara vez alcanza al uno por mil; suponiendo que se emplea la edad media correcta de los fallecidos.^{5/}

La estimación (1), de más fácil cálculo puesto que no involucra ningún supuesto respecto a una edad media particular de los fallecidos, está expuesta a errores considerables a menos que se tenga cuidado en la selección de los grupos apropiados de edades. Las tablas siguientes muestran aquellos grupos de edades en que los errores son considerables, y los agrupamientos claves que mantienen bajo el error.

ERRORES EN LA ESTIMACION (1)

Edades jóvenes	Grupos de edades...	1-4 [*]	5-9	5-14	5-19
	Error porcentual...	-0.59	-0.40	-1.16	-0.17

* Al elevar el límite superior de las edades en forma rápida, este error aumenta.

Edades avanzadas	Grupo de edades	85 y más	80-84	80 y más	75 y más	65 y más	50 y más	45 y más
	Error porcentual	1.21	-0.22	-0.13	-1.13	-1.94	-1.04	-0.45

Edades Centrales. Cualquier agrupación en el rango 20-44 llevará a un error más pequeño, siendo el máximo 0.30 para todo el grupo.

APENDICE II

HALLAZGO DE UNA TABLA MODELO DE VIDA PARA IGUALAR LA TABLA DE VIDA ESTIMADA EN DOS EDADES

I. Teoría general

Supóngase que los datos comprenden las defunciones en tres grupos de edades consecutivos, extendiéndose el último hasta el fin de la vida. (Estos supuestos no son esenciales, pero simplifican el álgebra). Denótese estas

definiciones por $D_{a,b}$, $D_{b,c}$ y $D_{c,90}$. Tomamos luego:

$$\frac{D_{a,b}}{a+b-1} = \frac{B(1 - 1_b)}{1_0} \dots \dots \dots (1)$$

$$\frac{D_{b,c}}{b+c-1} = \frac{B(1_b - 1_c)}{1_0} \dots \dots \dots (2)$$

$$\frac{D_{c,90}}{c+90-1} = \frac{B(1_c)}{1_0} \dots \dots \dots (3)$$

Al dividir (1) por (2) y (2) por (3) se eliminan B y 1_0 de modo que:

$$\frac{D_{a,b}}{D_{b,c}} (1 - r)^{\frac{c-a}{2}} = \frac{1_a - 1_b}{1_b - 1_c} \dots \dots \dots (4)$$

$$\frac{D_{b,c}}{D_{c,90}} (1 - r)^{\frac{90-b}{2}} = \frac{1_b - 1_c}{1_c} \dots \dots \dots (5)$$

Al dividir la $(90 - b)$ ava potencia de (4) por la $(c - a)$ ava de (5) se elimina r de modo que:

$$\frac{(D_{a,b})^{90-b} (D_{c,90})^{c-a}}{(D_{b,c})^{c+90-a-b}} = \frac{(1_a - 1_b)^{90-b} (1_c)^{c-a}}{(1_b - 1_c)^{c+90-a-b}} \dots \dots \dots (6)$$

Las potencias involucradas son grandes y difíciles de manejar. Sin embargo, todas son o $(90 - b)$, o $(c - a)$ o su suma. De este modo, si puede disponerse que ambas tengan un factor grande en común, se produce una simplificación considerable. Por ejemplo, tomando un ejemplo no acorde con la realidad (entre otras cosas introduce la mortalidad infantil) supongamos que $a = 0$, $b = 30$, y $c = 60$. Luego $(90 - b) = 60$, $(c - a) = 60$ y su suma es 120.

Podemos tomar entonces la 60ava raíz de todos los términos en nuestra ecuación, para reducirla a:

$$\frac{(D_{a,b}) (D_{c,90})}{(D_{b,c})^2} = \frac{(1_a - 1_b) 1_c}{(1_b - 1_c)^2}$$

Puede evaluarse el lado izquierdo. Por tanto, podemos encontrar la tabla modelo de vida de Naciones Unidas que produce este valor también en el lado derecho.

Una elección de a, b y c más acorde con la realidad implicaría evitar los primeros años de vida, cada una de las edades que termina en cinco (para reducir el error a causa de preferencias por ciertos dígitos) y asegurar que los grupos elegidos incluyesen suficientes defunciones. También habría que mantener bajo el valor de c, para evitar la división de la región donde se exagera la edad con fines de prestigio; a = 25, b = 45 y c = 55 podrían resultar satisfactorias. En este caso la ecuación básica se reduce a:

$$\frac{(D_{a,b})^3 (D_{c,90})^2}{(D_{b,c})^5} = \frac{(1 - r_a)^3 (1 - r_c)^2}{(1 - r_b)^5}$$

Si hay interés en que el valor de r logre la igualdad, puede evaluarse substituyéndolo nuevamente en la ecuación (4) o (5).

2. Caso especial 1

Cabría señalar que en las ecuaciones precitadas se supone que para todos los grupos de edades $(1 - r)^{a,b}$ puede estimarse como $(1 - r)^{\frac{a+b-1}{2}}$. Se discutieron fórmulas modificadas para algunos grupos de edades en el texto y, si están implicados estos grupos de edades, hay que realizar los ajustes necesarios. Si se eligen los grupos de edades 1 - 14, 15 - 44 y 45 y más, se encontrará que cuando se emplea la ecuación especial $(1 - r)$, el resultado final es casi exactamente:

$$\frac{(D_{1,15})^3 (D_{45,90})^2}{(D_{15,45})^5} = \frac{(1 - r_{1,15})^3 (1 - r_{45})^2}{(1 - r_{15,45})^5}$$

$$y, \quad \frac{D_{1,15}}{D_{15,45}} (1 - r)^{25\frac{1}{2}} = \frac{(1 - r_{1,15})}{(1 - r_{15,45})}$$

Al denotar $\frac{D_{1,15}}{D_{15,45}}$ por A, $\frac{D_{45,90}}{D_{15,45}}$ por B, $\frac{1 - r_{1,15}}{1 - r_{15,45}}$ por K y $\frac{1 - r_{45}}{1 - r_{15,45}}$ por L, la primera ecuación es $A^3 B^2 = K^3 L^2$.

Los valores de $K^{\frac{3}{2}}$ para algunas de las tablas modelo de vida de Naciones Unidas son como sigue:

Edad	0	5	10	20	40	60	80	95	115
K^3	20	22 $\frac{1}{2}$	25	30	40	50	60.4	68.2	73.9
K^3	1.05	1.04	1.07	1.23	1.97	3.4	6.4	8.1	8.2

Los valores de K^3 calculados a partir de datos extraídos del United Nations Demographic Yearbook son los siguientes:

País y año	Singapur 1936	Havaii 1936	Islandia 1936	Hungría 1936	Formosa 1930	Musulmanes de Palestina 1936	Egipto 1945	Egipto 1936
K^3	0.12	0.21	0.25	1.77 ^a	4.6 ^b	13	27	52

- ^a Correspondiente a un nivel tolerable de mortalidad, pero $r = -0,008$.
- ^b Mortalidad bastante leve.

Estos fueron seleccionados como casos extremos entre los resultados poco satisfactorios que pueden obtenerse. Por otros resultados que son altamente satisfactorios aparecen en el texto. El lector quizás especule acerca de si las migraciones o alguna otra causa que hace que la población no sea estable, o alguna otra causa, llevó a los fracasos precisados.

Para indicar el posible alcance de las dificultades debidas a variaciones en el patrón de mortalidad puede mencionarse que para la T.V.I. 3 (reconociendo que es una tabla de vida errónea) $K^3 = 6,2$ correspondiendo al nivel del modelo de Naciones Unidas, alrededor de 80 con $\frac{0}{0} = 80$, pero para la T.V.I. 3 $\frac{0}{0} = 40$.

3. Caso especial II

El caso especial considerado más arriba presenta el mérito relevante de que todas las subdivisiones de las defunciones ocurren en edades que terminan en dígito cinco, minimizando de esta modo el error debido a la preferencia por ciertos dígitos, y, segundo, que es probable que el último grupo de edades, 45 años y más, incluya todas las edades de las cuales han sido transferidas las defunciones dentro la práctica común de exagerar la edad en las edades más avanzadas. Presenta, sin embargo, una notoria desventaja: todas las comprobaciones se realizaron con datos de poblaciones relativamente bien conocidas, de modo que si la técnica produjera un resultado erróneo, se

no detectar esto evitándose así llegar a conclusiones erróneas. En las circunstancias en que la técnica podría ser de verdadero valor, la verificación del resultado podría no hallarse disponible, y las comprobaciones anteriores han demostrado claramente que resulta esencial algún tipo de control.

Mediante la elección de otros grupos de edades, vale decir 1-14, 15-39, 40-59 y 60 y más, podría desarrollarse una técnica que incorporase algún tipo de control.

Al denotar $D_{1,15}/D_{15,40}$ por m , $D_{15,40}/D_{40,60}$ por n , y $D_{40,60}/D_{60,90}$ por o , y los valores correspondientes de la tabla de vida por M , N y O , la aplicación de la teoría general anterior demuestra que (aproximadamente la tabla modelo de vida de Naciones Unidas correcta es aquella para la cual (i) $\frac{M}{N} = \frac{m}{n}$, (ii) $\frac{N}{O} = \frac{n}{o}$ y (iii) $\frac{M}{O} = \frac{m \cdot o}{n}$. En teoría, un modelo que satisficiera una de estas ecuaciones, satisfaría a todas. En la práctica el hecho que todas fueran satisfechas proporcionaría un control sumamente importante. El valor de r es dado aproximadamente por:

$$(1 - r)^{23} = \frac{M}{m} = \frac{N}{n} = \frac{O}{o}$$

El rango de valores de estos criterios en las tablas modelo de vida de Naciones Unidas es el siguiente:

$\frac{D}{O}$	20	30	35	40	50	60.4	70.2	73.9
Nivel	0	20	30	40	60	80	100	115
M/N	1.265	1.125	1.107	1.101	1.126	1.239	1.39	1.353
N/O	0.531	0.865	1.065	1.281	1.730	2.067	1.87	1.675
M/O	0.672	0.973	1.179	1.410	1.948	2.561	2.62	2.266

Se han hecho algunos ajustes al realizar comprobaciones con datos extraídos del Demographic Yearbook. Para algunos países no se mostró la división de las defunciones en el punto de las edades 40 ó 60 y, en esas circunstancias sólo se estimó una sub-división aproximada. En otros casos el alza y baja del número de defunciones en grupos quinquenales alternados de edades claramente indicaba preferencia por ciertos dígitos. Una vez más sólo se realizó una corrección muy aproximada.

La siguiente selección de resultados incluye casos en que los tres criterios resultaron inaceptablemente bajos, en que todos éstos resultaron inaceptablemente altos, y en que dos resultan inaceptables pero el tercero resultó razonable.

Caso	Singapur 1936	Egipto 1936		Islandia 1936	Costa Rica 1937	Chile 1936	Hungría 1936	Formosa 1936
		Datos brutos	Con un ajuste considerable con respecto a la preferencia por ciertos dígitos					
n/a...	0.56	2.63	2.79	0.343	0.794	0.902	1.790	1.938
n/a...	0.45	2.38	1.98	2.968	1.847	1.387	2.205	0.527
n/a...	0.25	6.26	5.54	1.361	1.467 ^B	1.251	1.742	1.215

Al igualando el nivel justamente sobre los 40, correspondiendo a una θ_0 de 41, es decir, el mismo resultado obtenido mediante la aplicación del Caso Especial I.

El único caso en que se encontró un ajuste consistentemente bueno fue el de Siam, 1936. La técnica del Caso Especial I para estos datos mostró un valor de 1.28 para $A_3 B^2$ correspondiendo al nivel 22 y a $\theta_0 = 31$ con $r = 0.01$. La aplicación del Caso Especial II demostró: $m/n = 1.004$, justamente menos que el mínimo posible de 1.101 entre los niveles 35 y 40; $m/o = 1.091$, correspondiente al nivel 31, $\theta_0 = 36$ y $m/o = 1.095$, correspondiendo al nivel 26, $\theta_0 = 38$ y $r = 0.015$. Las tablas de vida nacionales muestran $\theta_0 = 37$ y las tasas de natalidad y mortalidad publicadas en la década de 1930 fueron de alrededor de 30 y 20 por mil respectivamente.

Resultará evidente que estos dos casos especiales de ningún modo agotan las posibles aplicaciones de la técnica general, que son innumerables. Sin embargo, se requiere algo más que experimentaciones a la ventura en una combinación de edades tras otra para superar lo que se cree son las dos dificultades que han alterado los resultados, vale decir, las grandes variaciones en los patrones de mortalidad y los datos erróneos. Sin duda que esas variaciones con respecto al modelo de población estable fueran la causa de algunos fracasos, particularmente por un elemento de migración, pero resulta dudoso que estos casos puedan tener algún tipo de solución. Pero es posible que las otras dificultades pudieran ser minimizadas mediante la selección de grupos de edades apropiados. En este sentido cabría mencionar que las pruebas realizadas dieron la impresión de que el criterio "m/o" proporcionaba una respuesta razonable más a menudo que cualquier otro.

Nota: La traducción del artículo, "Una nota sobre la estimación de la mortalidad y otras características de la población, dadas las defunciones por edad", de H.H. Carrier, ha sido realizada por el Curso Avanzado de 1971.

Referencias

- 1/ Por ejemplo, suponiendo que r es pequeña y d_x es una función lineal de x , tratando x como una variable continua o bien que discreta, y expandiendo todas las funciones involucradas en potencias de r , se demuestra fácilmente que si $(1-r)^{a,b}$ es estimado como el valor ya sea en la edad media, o en la edad media de los fallecidos en el intervalo (a,b) , en cualquiera de los dos casos el error introducido es proporcional al cuadrado de la longitud de (a,b) .
- 2/ The Registrar General's Decennial Supplement, 1921, parte III, págs. 27-34.
- 3/ Obsérvese que en el intervalo 60-64 por ejemplo, $(1-r)^{\frac{a+b-1}{2}}$ es $(1-r)^{62}$ y no $(1-r)^{62\frac{1}{2}}$.
- 4/ Esta debe obtenerse para una tabla de vida que se considere similar a la involucrada. Para las Tablas de Vida Inglesas 3 (en que se muestran l_x y q_x), $(c + \frac{1}{2})$ se obtiene de $\left[\frac{1}{a}(a+q_a) - \frac{1}{b}(b+q_b) \right] / (l_a - l_b)$. Para las tablas modelo de vida de Naciones Unidas, véase págs. 72-81 en la publicación citada, la fórmula (a) es la más apropiada. Donde $\frac{1}{5}x$ ha sido calculada como $5(l_x + l_{x+5})/2$; esta fórmula dará $(x + \frac{1}{2})$ como la edad media de los fallecidos para un grupo quinquenal. Obsérvese que ambas fórmulas dan $c + \frac{1}{2}$ no c .
- 5/ La edad media de los fallecidos en los grupos de edades pequeños o moderadamente grandes no resulta muy crítica para los niveles de mortalidad. En Inglaterra y Gales desde 1850 y 1920, en tanto que la esperanza de vida al nacer para los hombres se elevó de 39,91 a 55,52, la edad media de los fallecidos bajó de 2,31 a 2,22 para las edades de 1-4, y se elevó de 76,30 a 80,22 para los 45 y más, y de 83,64 a 88,72 para los 65 y más.
- 6/ De estas tres asociaciones, solamente dos son independientes.
- 7/ No cae dentro del ámbito de esta nota la discusión, el debate de los métodos precisos para analizar estas situaciones. Brevemente, la primera puede por lo general resolverse con suficiente precisión suponiendo una función cuadrática de frecuencia para las defunciones por edad, determinando los coeficientes mediante la integración de la función a lo largo de tres intervalos apropiados para los cuales se conocen las defunciones. Intercalados estos coeficientes se hace la integración a lo largo del intervalo requerido y desconocido. Este enfoque es naturalmente conocido. La segunda situación -dada con un alza y baja claramente alternados- se considera en una nota

se está preparando el autor. En el Estudio de Naciones Unidas de Methods for Age-Adjustment Projections, págs. 12
y 13, se recomienda una simple fórmula de ajuste la cual, estoy de acuerdo, tiene mucho de recomendable.
Para datos con un alza y baja alternas y marcadas, sin embargo, puede preferirse otra técnica. Esto involucra
la agregación de los datos originales, datos en grupos quinquenales, en grupos decenales que van desde las edades
que terminan en el dígito cinco hasta aquellas que terminan en cuatro, v.g., 45-54, 55-64, etc. Estos pueden
luego separarse en grupos quinquenales usando la fórmula que se da en el Estudio de Naciones Unidas ya mencionado,
sección 101.

Nota: La traducción del artículo de W.H. Carrier "Una nota sobre la estimación de la mortalidad y otras
características de la población, dadas las defunciones por edad" ha sido realizada por el Curso
Avanzado 1971. ~~Adicionalmente, la traducción de la última parte, que corresponde al Apéndice, también
fue realizada por el Curso Avanzado 1971.~~

27 de septiembre de 1971