Construcción de tablas de mortalidad

Por **D. Miguel Saura del Campo,**
Ingeniero Industrial. — Estadístico Facultativo del Instituto Nacional de Estadística. —De la Dirección General de Aduanas.—De la «Powers-Samas Accounting Machines».

I.—Probabilidades de muerte.

Sean mis primeras palabras de agradecimiento. Ocupar un puesto, aunque sea el último, en las tareas anuales del Instituto de Actuarios, es honor y distinción, que yo, desde luego, no merezco. Oír las frases de elogio que de mí se han hecho, rebasa, con mucho, lo que de mi torpe expresión y de mis modestos conocimientos puede esperarse. Sea, pues, mi gratitud a todos.

Antes de entrar en el tema, trataré de disculpar el atrevimiento que he tenido opinando sobre una técnica que, como la de mortalidad, tan célebres actuarios, estadísticos y demógrafos formaron.

Por pequeñas dudas en los métodos censales de ajuste, comenzó esta investigación. De las dudas pasé a lo que yo creo errores, para llegar más tarde a soluciones propias. Acogido cariñosamente mi primer trabajo, por el Ilmo. Sr. Director general de Estadística se nombró una ponencia (*) del Cuerpo Facultativo para su conocimiento. Anima-

---

(*) La ponencia, presidida por el Ilmo. Sr. Director general de Estadística, se constituyó por:

Excmo. Sr. D. Antonio de Miguel Martín.
D. José Irizar Egui.
D. Adolfo Melon Ruiz de Gordejuela.
D. Javier Ruiz Almansa.
D. José Sáenz García.
D. José Ros Jimeno.

(Sigue la nota en la página siguiente.)

do tanto por mis colaboradores, como por Jefes y compañeros, me vi obligado a seguir su estudio.

Los métodos de ajuste me llevaron a las perecuaciones, siguiendo, más tarde, por los gráficos a las tasas de muerte, funciones de población y mortalidad; para finalizar en las funciones de tasas y función de supervivencia. De esta forma, y sin yo pretenderlo, fuí penetrando poco a poco en la técnica de la mortalidad, y de tal forma que, cuando quise recapitular, encontré que poco o nada había quedado sin remover, de los procedimientos de formación para tablas nacionales de mortalidad.

Por último, debo declarar que mis debilidades de padre bondadoso no me llevaron hasta el punto de pasarme inadvertido que en una investigación como esta, que tanto tiene de original, se habrán deslizado errores que mi limitado bagaje de conocimientos no permitió ver, pero que ustedes, el más alto tribunal que puede entender sobre tablas de mortalidad, juzgará.

En el bosquejo de esta conferencia incluí un estudio previo de los trabajos que acerca de la mortalidad han sido hechos, y también incluía las orientaciones y puntos de vista distintos, desde los que se ha enfocado el asunto; pero llevado a cuartillas el guión, desistí de este propósito para no agotar la paciencia de la Asamblea. Así, pues, siéndoles perfectamente conocido cuanto yo pudiera decir a este respecto, comienzo el estudio de la mortalidad por su representación gráfica.

REPRESENTACIONES GRÁFICAS.—*Esquema de Lexis.*—Ante la complejidad de los elementos que intervienen en el estudio de la mortalidad se hace precisa una representación gráfica que aclare la marcha del fenómeno, siendo el llamado *esquema de Lexis* el que resuelve el problema.

Lexis fundamenta su construcción gráfica en las llamadas *líneas de vida,* que no son más que las huellas dejadas en un sistema cartesiano de la continuidad de vida de un conjunto de habitantes. Lo mismo que el barógrafo o el termógrafo marcan la curva de las presiones o de las temperaturas sobre un papel registrador, así el *esquema de Lexis* marca

---

D. Andrés Cerdán Moreno.
D. José María Huarte Baztán.
Doña Victoria Beatriz Baylos Corroza.
D. Glicerio Bermúdez Romero.
D. Juan Cruz San Román Ramírez.
D. Manuel García Alvarez.

sobre el plano la duración de todas las vidas que forman un conjunto de habitantes.

Dos ejes cartesianos son divididos en segmentos iguales entre sí, e iguales en los dos ejes, llamando al eje de abcisas *línea de tiempos* y al de ordenadas *línea de edades*. La *línea de tiempos* no es más que la representación rectilínea del tiempo, contado a partir de un momento que se toma como origen de coordenadas, llamado *origen de tiempo*.

En la *línea de tiempos* está el comienzo de todas las *líneas de vida*, que empiezan por nacimiento, desarrollándose éstas según ordenadas

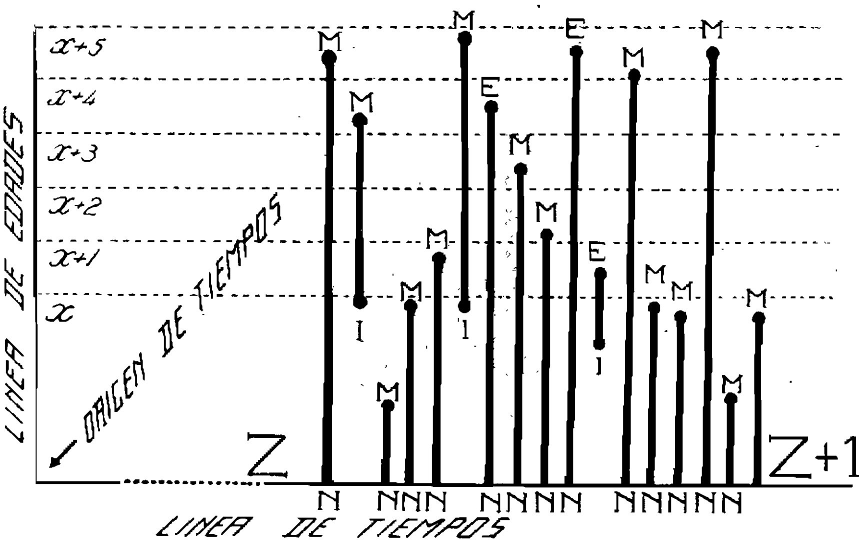


FIG. 1

que parten de aquélla y precisamente en los puntos cuya fecha corresponde a la del nacimiento. Así un individuo cuyo origen de registro o nacimiento se encuentre en el punto N (fig. 1), irá describiendo la recta N M en el transcurso del tiempo.

La recta N M corta sucesivamente a las paralelas por $x, x+1, x+2$..., etcétera, a la *línea de tiempos* a medida que las edades del individuo considerado pasen por esos valores. Esta recta-registro concluye por dos causas, o bien porque el individuo emigre del conjunto en observación, línea N E, o porque muera, línea N M. Las *líneas de vida* pueden también comenzar fuera de la *línea de tiempos*; caso que ocurre cuando los individuos que ingresan en el conjunto observado lo hacen en fecha posterior a la de su nacimiento, tales son las líneas de vida I M e I E.

Hasta este momento sólo hemos considerado una sola *línea de vida*; pero el *esquema de Lexis* registra todas las que forman un conjunto en observación. En la figura se ha representado el caso para la generación comprendida en los tiempos $(z) \rightarrow (z + 1)$.

Es condición constructiva del *esquema de Lexis* que, tiempos iguales contados sobre cualquiera de los ejes, o una parte sobre la *línea de*

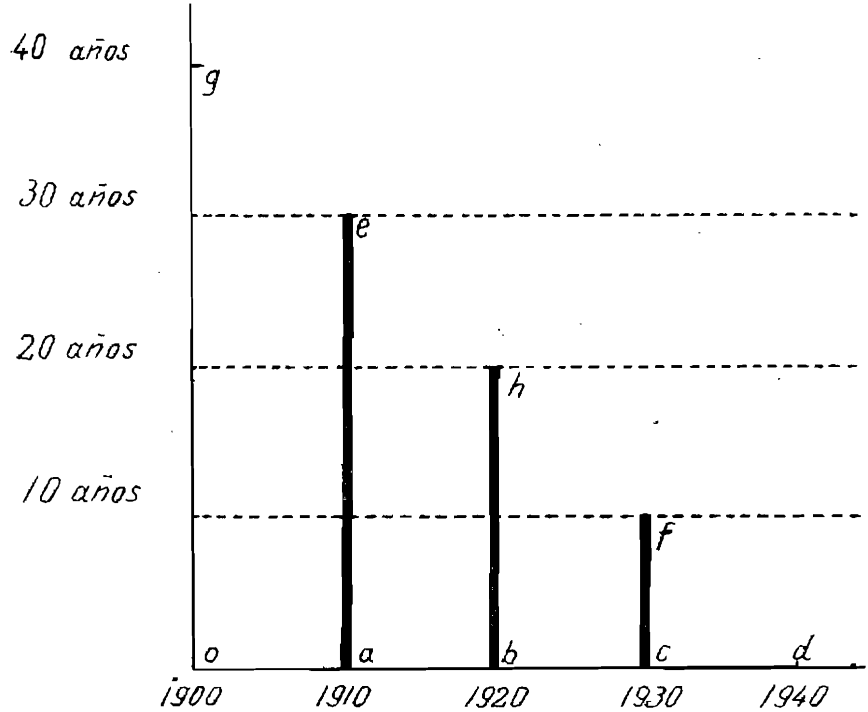


FIG. 2

tiempos y otra sobre la *línea de vida*, se expresen por segmentos también iguales. En consecuencia, la *línea de vida*, que comienza por nacimiento en el principio del 1910 (fig. 2), tendrá treinta años de edad en el principio del año 1940 y se efectuará que:

$$oa + ae = og = od = oa + ad.$$

Lo mismo ocurrirá para cualquier otra *línea de vida*, *cf* en que:

$$oc + cf = og = od = oc + cd,$$

igualdad de la que se deduce que:

$$cf = cd,$$

es decir, que cuando el tiempo llega al comienzo del año 1940, el individuo que nació en el primer instante del año 1930 tiene una *línea de vida* cf igual al intervalo cd.

Limitando el caso a la observación de un año, es decir, al conjunto de nacidos (despreciando la migración para mayor sencillez) durante el

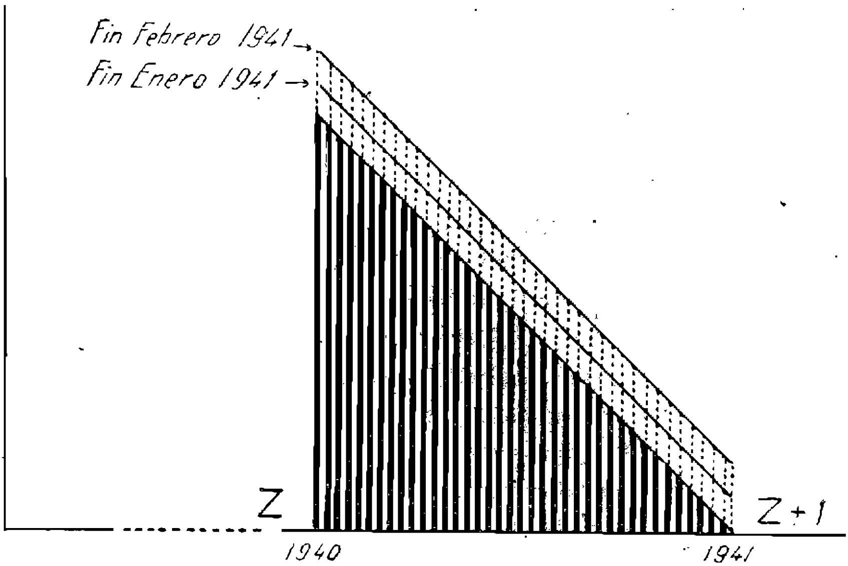


FIG. 3

transcurso del año 1940, veremos (fig. 3) que las *líneas de vida* de todos los nacidos en ese año tienen sus extremos en una recta que forma un ángulo de 45° con la *línea de tiempos* en 1.° de enero de 1941.

A partir de primero del año 1941 continuamos la observación. Las *líneas de vida* del conjunto observado aumentan segmentos iguales en períodos de tiempo iguales, es decir, que los extremos de las *líneas de vida* se desplazan, según rectas que forman ángulos de 45°, con la *línea de tiempos*, a medida que transcurre la observación. Estas líneas, en las que en todo momento reposan los extremos de las *líneas de vida*, son las llamadas *isócronas*.

Con los elementos anteriores se expresa los valores que intervienen en la mortalidad, y se hace asimilando las *líneas de vida* a líneas de fuerza y conjunto de habitantes a conjunto de *líneas de vida,* siendo dadas las distintas modalidades que puede tomar aquél por las *líneas de vida* cortadas por las rectas ya descritas.

Existen en un momento dado tantos habitantes de una generación cuantas sean las *líneas de vida* cortadas por la *isócrona* correspondiente. De igual forma existen o existieron tantos habitantes, de edad x, procedentes de una generación cuantas fueran las *líneas de vida* cortadas por la recta de edad x.

Los muertos son determinados por diferencia entre los conjuntos de *líneas de vida* cortadas por las rectas que limitan las condiciones o fechas en que ocurrieron las defunciones.

Teniendo presente el posterior desarrollo de este tema, se limitará el *esquema de Lexis* al campo de dos generaciones anuales y a cuatro edades (fig. 4).

Debajo de cada línea del esquema está colocado el valor que representan las *líneas de vida* cortadas por ella, siendo el valor de las defunciones el correspondiente al triángulo en que están colocados.

Empleamos la notación E_x^z para los habitantes que cumplen x años durante el año z; P_x^z para los habitantes que en primero del año z tienen x años; $\alpha\, D_x^z$ para los que mueren con x años, durante el año z, procedentes de los habitantes que en principio del mismo año tienen $x - 1$ años, y $\delta\, D_x^z$ para los fallecimientos de edad x ocurridos en el año z, procedentes de los habitantes que en primero del mismo año tienen x años.

Por no alejarme del tema principal a tratar, omitiré las variaciones introducidas por otros estadísticos en el *esquema de Lexis.*

Como hemos visto, se fundamenta el *esquema de Lexis* en las *líneas de vida,* expresando los distintos valores que intervienen en la mortalidad por conjuntos de *líneas de vida* cortadas. Acostumbrado, en el campo de la demografía, a la representación por líneas, áreas o volúmenes de los conjuntos estadísticos, intenté llevar este esquema a una representación gráfica del tipo clásico, pero el primer inconveniente que se encuentra (si de un gráfico de área se trata) es el de que las superficies limitadas por el eje de abcisas, las ordenadas y las *isócronas* correspondientes son menores a medida que el grupo de población que se consideraba es de menor edad. Hecho que no permite establecer una escala con la realidad, por ser contrario a ella. Este inconveniente pudiera ser salvado mediante un cambio de ejes, introduciendo ciertos convenios en el gráfico. Pero a lo que no encontré solución fué a que las líneas

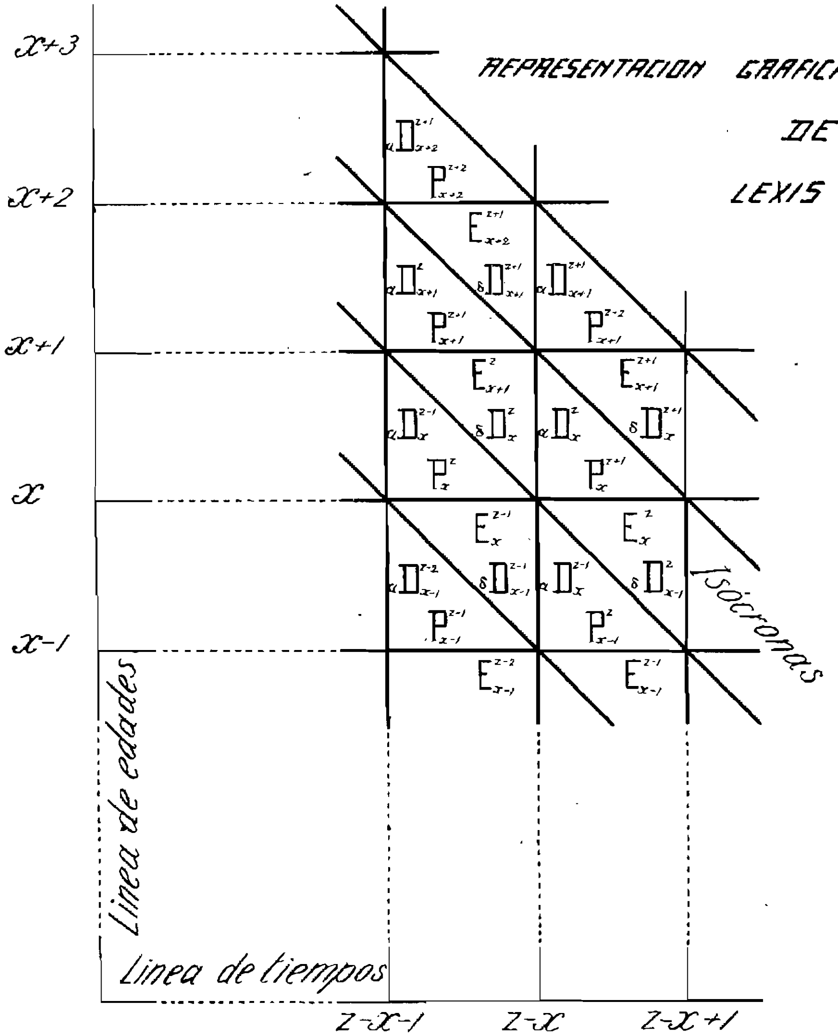


FIG. 4

isócronas (tengan o no la inclinación de 45°) sean rectas; pues admitir esto en un gráfico de superficies sería tanto como decir que la población se distribuye por edades según los términos de una progresión aritmética, supuesto que no creo acepte ningún demógrafo.

Por todo lo anteriormente expuesto, me he permitido sustituir el *esquema de Lexis* por el gráfico que seguidamente presento, en que, además de venir expresado por áreas, los elementos precisos en los cálculos de unas *Tablas de Mortalidad,* permite ver fácilmente el desarrollo y obtención de las *funciones de población, mortalidad, tasas de muerte,* etc.; *funciones* que como novedad tendré el honor de presentarles en la segunda conferencia.

Representación gráfica del conferenciante.—Finalizadas las labores de tabulación, de una inscripción censal o de un conjunto de pólizas de vida, podremos formar un gráfico (fig. 5) en que, dividido el eje de abcisas, que representa edades, en intervalos iguales, se ha construído, sobre ellos, rectángulos tales que, el área de cada uno, a la escala adoptada, represente el número de habitantes o de asegurados de cada edad. Expresando el intervalo $(x) \to (x + 1)$ edades comprendidas entre x y $x + 1$ años.

Tomemos ahora (fig. 6) uno de estos rectángulos, el P_x por ejemplo, y las papeletas o cédulas que les correspondan y clasifiquémoslas en dos grupos, compuesto el primero por edades de x a x más seis meses y el segundo, por edades de x más seis meses a x más un año. Elevemos sobre cada medio intervalo (ya que se comprenden seis meses) un rectángulo, cuyas áreas correspondan a estas subdivisiones, quedando el rectángulo primitivo, P_x, subdividido en dos y cuya área será, evidentemente, igual a la suma de los dos nuevos.

Si continuamos haciendo subdivisiones, por trimestres, meses, días, etcétera, llegaremos a un gráfico (fig. 7) en que el área viene limitada, no por una recta sino por una quebrada confundible, en el límite, con una curva.

En esta función en escalera, no es posible eliminar los escalones, por tratarse de la unidad demográfica, habitante, muerto, etc., pero cuya continuidad se admite en Estadística, y más en este caso, en que, pudiendo crecer la variable independiente, x, por incrementos infinitamente pequeños por ser tiempo, el momento estático de estas áreas es función continua, por serlo los tiempos vividos por los componentes del conjunto de habitantes, que aquél expresa.

Lo mismo que hemos hecho con una edad, haremos con todas y entonces, la población, ya no vendrá representada por una serie de rectángulos sino por una curva (fig. 8) cuya expresión matemática,

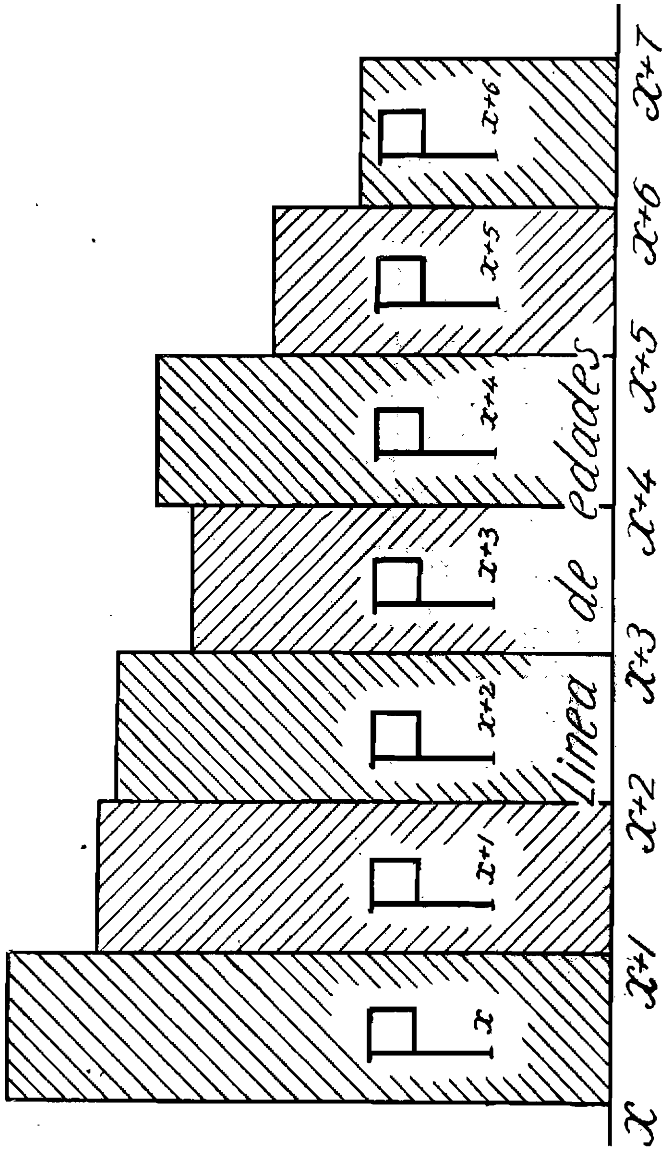


FIG. 5

P(x), nos dará, por su $\int_{x}^{x+1} P(x)\,dx$, los habitantes de edad x, lo mismo que antes nos lo daba el área de cada rectángulo.

La integral de P (x), no nos da más que la instantánea del conjunto, en un momento dado, fecha del cierre o inscripción censal, pero la

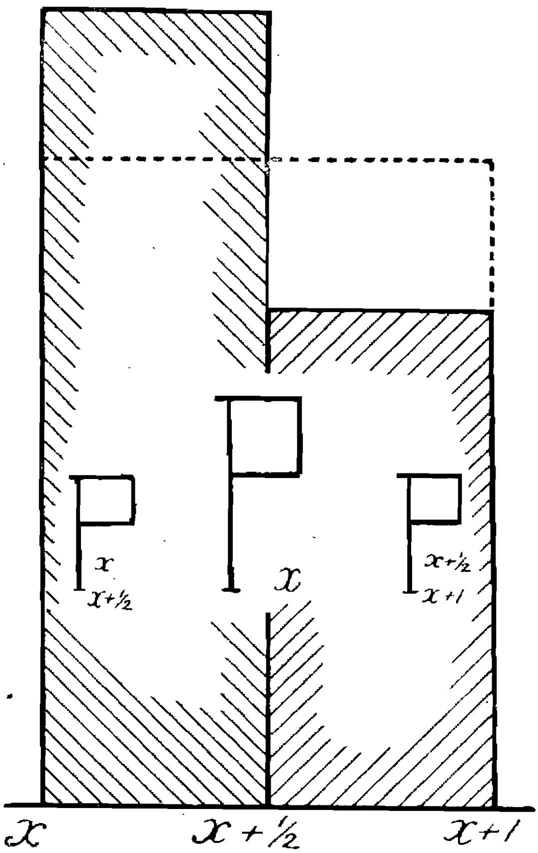


FIG. 6

población, durante la observación a que se la somete, sufre variaciones que son motivadas por fallecer habitantes, por cumplir años, por emigrar del conjunto observado y por inmigrar en el mismo, aunque prescindiremos de la consideración de estas dos últimas, dada su escasa importancia en los cálculos de unas Tablas nacionales de Mortalidad.

Las defunciones que se producen en el conjunto (si éste es suficientemente grande) durante un mes, afectarán a todas las edades, produciéndose un descenso en la curva que limita las áreas, P (x) (fig. 9) obteniéndose la P_1 (x), que tendrá la forma que sea y de la cual no

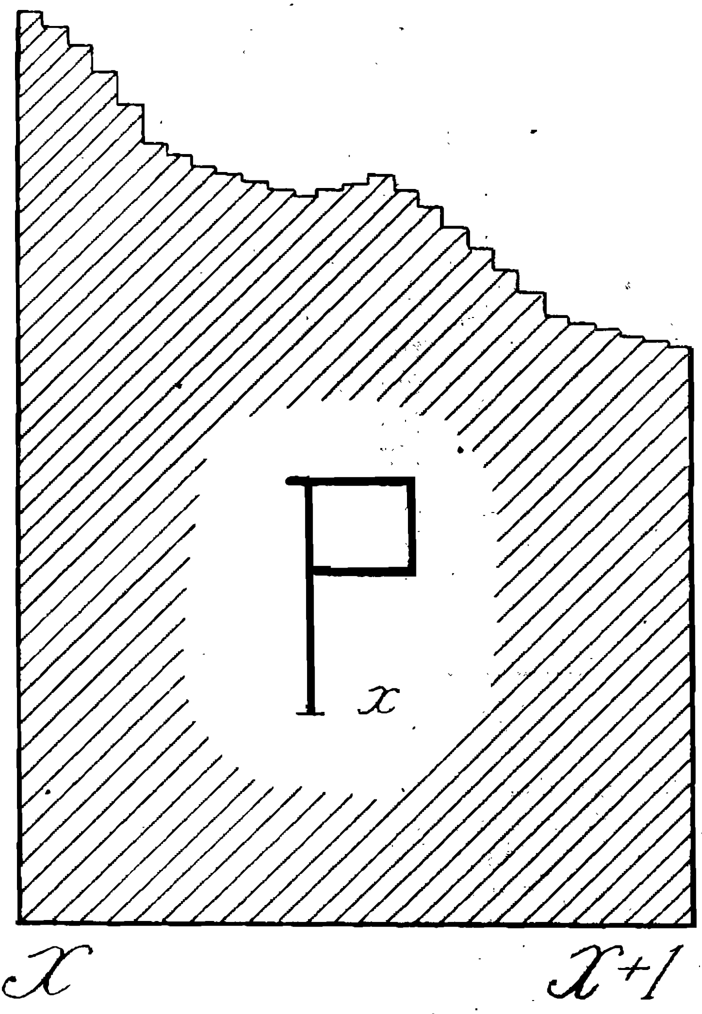


FIG. 7

vamos a ocuparnos por ahora. Lo mismo que ocurre al final del primer mes, ocurrirá en los siguientes, llegando, por último, al final del año a la P_{12}(x). Dándonos el área comprendida entre las ordenadas x y $x + 1$ y las dos curvas P (x) y P_{12} (x) las defunciones ocurridas durante un año, en el conjunto de habitantes, que en principio de observación tienen edades comprendidas entre x y $x + 1$ años.

Símultánea con esta variación, es el cumplimiento de años de los individuos. Una cuarta parte del intervalo $(x) \rightarrow (x + 1)$ pasa (fig. 10) a la edad $x + 1$, durante el primer trimestre, otra cuarta parte en el segundo, hasta que todo el conjunto pasa a la edad $x + 1$ al final del año de observación.

Las dos variaciones, descritas, son simultáneas y actúan de modo

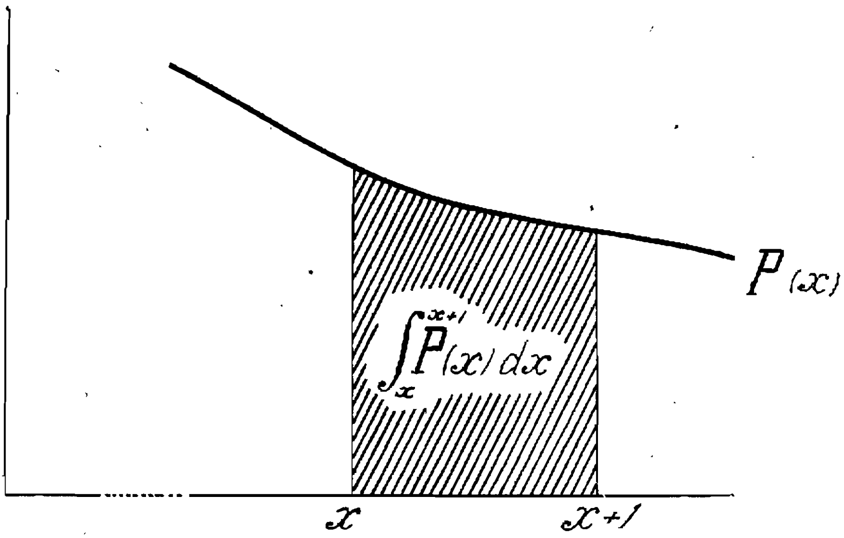


FIG. 8

igual sobre cada punto R de la curva P (x), a como lo harían los dos vectores T_1, T_2, obligando el T_1 a descender la curva (deformación aparte) a causa de la mortalidad y el vector T_2 a desplazarse horizontalmente, por cumplir edades superiores a la inicial. En la misma figura puede verse que, del área limitada por las ordenadas que pasan por $x + {}^3/_4$ y $x + 1$, la curva P (x) y el eje de abcisas, una parte pasa viva a la edad $x + 1$, la $\int E(x)\,d\,x$ y otra parte muere antes de pasar a esta edad, la $\int D(x)\,d\,x$, debido esto a los dos movimientos de la curva, reflejo del fenómeno real.

A la figura 11 han sido llevadas todas las variaciones que experimenta un conjunto de habitantes de varias edades en dos años y su construcción es la siguiente: En principio del año z y en vista de los resultados de un censo, calculamos la curva $P^z(x)$ (ya veremos cómo en la segunda conferencia), que nos da, por las superficies que ella limita, en unión de las ordenadas y de la línea edades-tiempo, los habitantes que de cada edad existen en la nación. Al cabo de un año esta curva habrá descendido y ocupará una posición $P^{z+1}(x)$, descenso que es motivado por la disminución de superficies, debida a las muertes ocu-

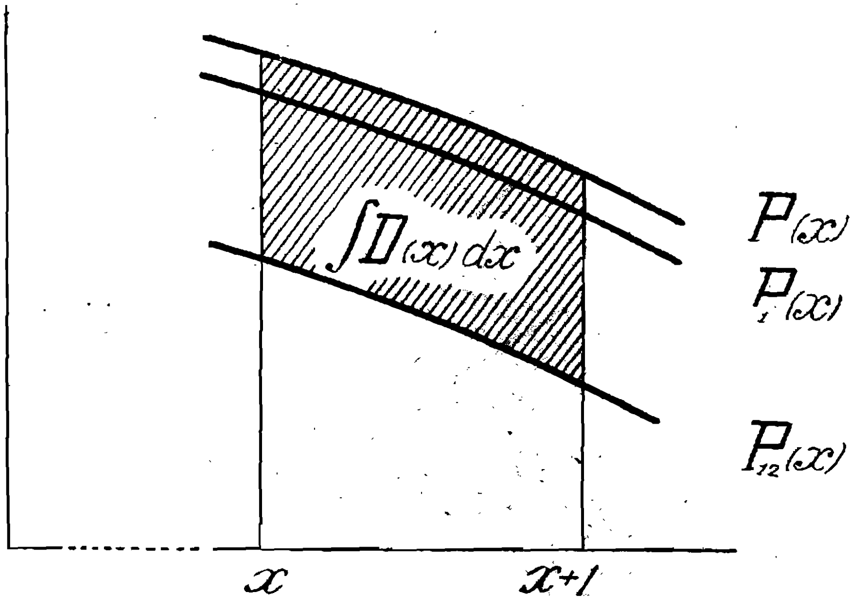


FIG. 9

rridas durante ese año y que a todas las edades afecta. Del mismo modo habrán sufrido un desplazamiento de una unidad todas las ordenadas, ya que las superficies o bien los habitantes que representan habrán aumentado en un año su edad.

El área P^z_x de la superficie limitada por la curva $P^z(x)$, las ordenadas que pasan por x y $x+1$ en principio del año z y el eje de edades-tiempo, viene a quedar reducida, a primeros del año $z+1$, al área P^{z+1}_{x+1}, limitada ahora por la curva $P^{z+1}(x)$ y las mismas ordenadas que han tomado los valores $x+1$ y $x+2$. La diferencia entre las áreas P^z_x y P^{z+1}_{x+1} es a causa de las defunciones habidas durante el año z en el conjunto de habitantes P^z_x.

A partir del tiempo inicial z, observemos el fenómeno en un tiempo Δz. La curva $P^z(x)$ sufrirá un descenso elemental que la transforma en la $\overset{z+\Delta z}{P}(x)$ y las ordenadas se desplazan Δx. Durante el tiempo elemental Δz mueren los habitantes expresados por el área de S N O Q, que se descompone en las dadas por el área de S N R Q que mueren con $x - 2$ años y la dada por el área de N O R que mueren con $x - 1$ años, ya

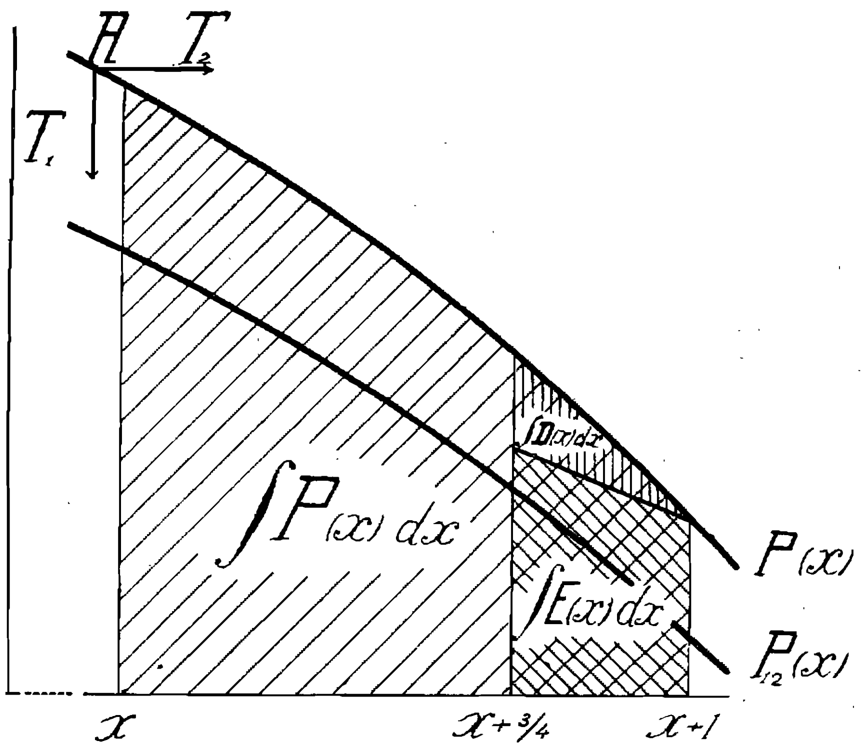


FIG. 10

que cumplieron dicha edad cuando se produjo la muerte. En los sucesivos tiempos elementales, las fajas elementales S N R Q disminuyen en valor, al paso que aumentan las correspondientes a N O R. Así el fenómeno varía desde morir todos con $x - 2$ años en el primer instante del año z, hasta morir todos con $x - 1$ al final de dicho año, siendo la línea N M la que marca la divisoria entre los muertos con esas edades, procedentes todos del conjunto de habitantes P^z_{x-2}.

Volviendo al conjunto P^z_x, veremos que durante el año z mueren

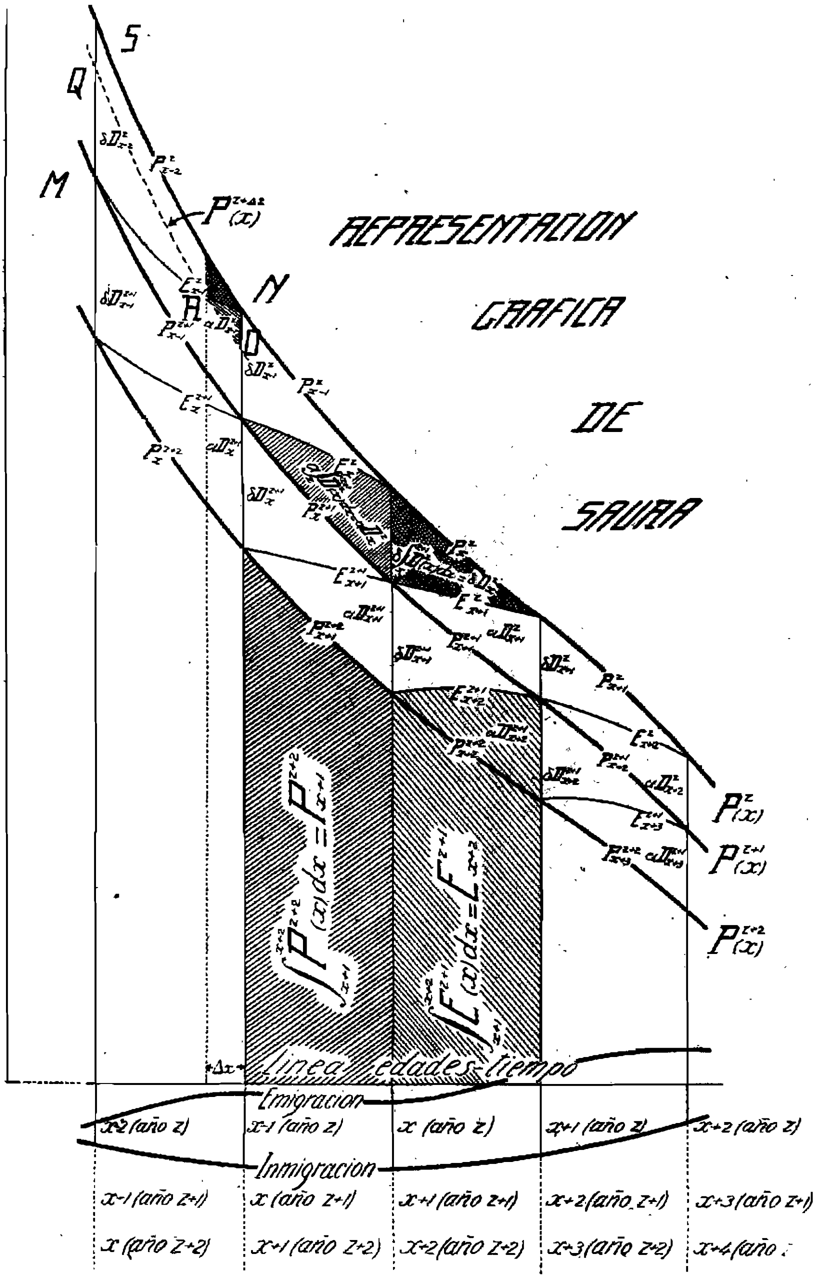
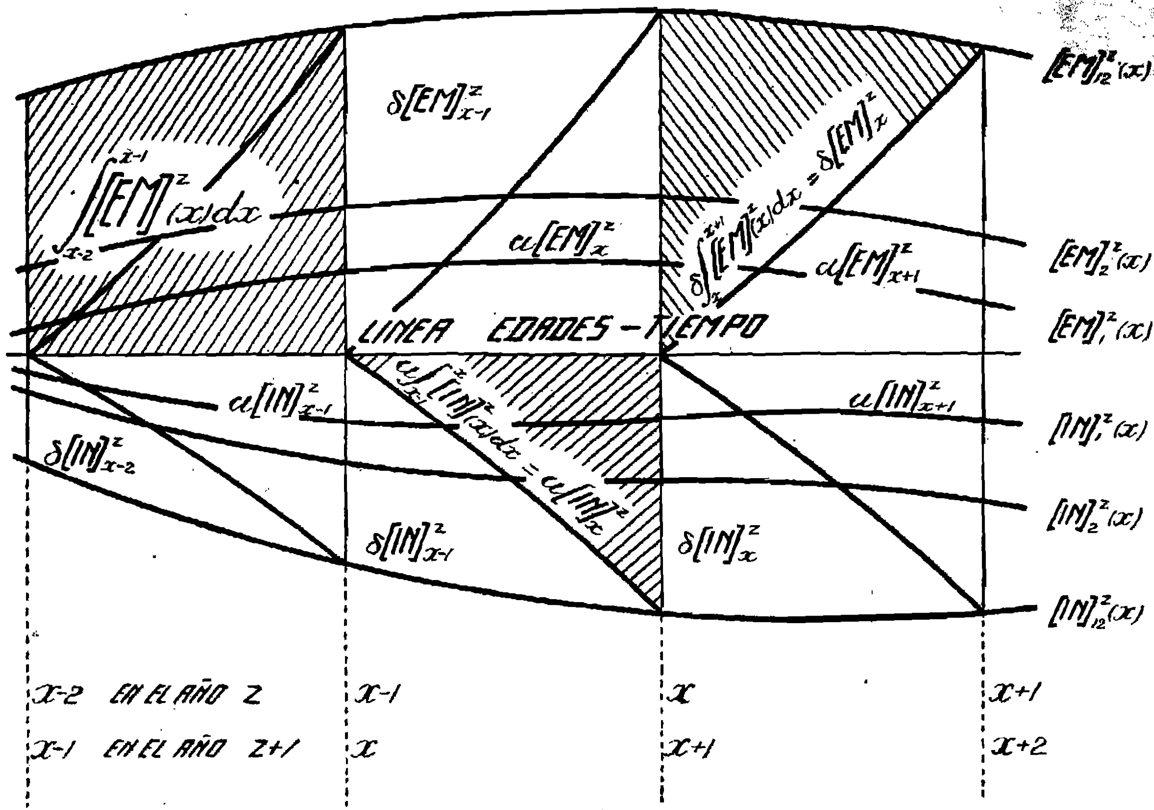

REPRESENTACION
GRAFICA
DE
SAURA
Linea edades-tiempo
Emigracion
Inmigracion
x-2 (año z)
x-1 (año z)
x (año z)
x+1 (año z)
x+2 (año z)
x-1 (año z+1)
x (año z+1)
x+1 (año z+1)
x+2 (año z+1)
x+3 (año z+1)
x (año z+2)
x+1 (año z+2)
x+2 (año z+2)
x+3 (año z+2)

FIG. 11



$\delta[EM]^z_{x-1}$
$\int_{x-2}^{x-1}[EM]^z(x)dx$
$\alpha[EM]^z_x$
$\delta\int_x^{x+1}[EM]^z(x)dx = \delta[EM]^z_x$
$\alpha[EM]^z_{x+1}$
$[EM]^z_{12}(x)$
$[EM]^z_2(x)$
$[EM]^z_1(x)$
LINEA EDADES - TIEMPO
$\alpha[IN]^z_{x-1}$
$\alpha\int_{x-1}^{x}[IN]^z(x)dx = \alpha[IN]^z_x$
$\alpha[IN]^z_{x+1}$
$\delta[IN]^z_{x-2}$
$\delta[IN]^z_{x-1}$
$\delta[IN]^z_x$
$[IN]^z_1(x)$
$[IN]^z_2(x)$
$[IN]^z_{12}(x)$
$x-2$ EN EL AÑO Z
$x-1$ EN EL AÑO Z+1
$x-1$
x
x
$x+1$
$x+1$
$x+2$

FIG. 12

δD_x^z con la edad x y αD_{x+1}^z con la edad $x+1$. En consecuencia, los que cumplen la edad $x+1$ durante el año z, se mueran o no después de cumplirlos, son los E_{x+1}^z.

Para mayor claridad han sido rayados algunos valores, viniendo designados éstos por la notación ya usada en el *esquema de Lexis.*

Por la poca influencia que en los cálculos tiene el movimiento migratorio, no lo hemos considerado anteriormente, pero el gráfico permite su representación. Sobre la línea edades-tiempo (fig. 12) se llevan las curvas (E M) de emigración. Al final del primer mes los que han emigrado de la nación son clasificados por edad y construída la curva $[E\,M]_1^z\,(x)$, que nos da por su $\int (E\,M)_1^z\,(x)\,d\,x$, los que de cada edad han emigrado. Al final del segundo mes esta curva se habrá desplazado a $[E\,M]_2^z\,(x)$ y al final del año a la $[E\,M]_{12}^z\,(x)$.

Lo mismo que en el caso general, las ordenadas sufren un desplazamiento de una unidad al cabo del año z.

Los habitantes que habrá que restar a los P_{x+1}^{z+1} serán los dados por la suma:

$$\alpha\,[E\,M]_x^z + \delta\,[E\,M]_x^z$$

y los que se restaran de E_{x+1}^z serán los $\alpha\,[E\,M]_x^z$.

Lo mismo, salvo ser incrementos, podemos decir de los valores de las funciones $[I\,N]_x^z$ de inmigración.

En el gráfico general (fig. 11) se han llevado las dos curvas, de emigración e inmigración, a la parte inferior de la línea edades-tiempo, restándose las dos áreas para así obtener en una sola superficie los valores totales de P_x^z y E_x^z.

Colectivos de habitantes y colectivos de muertos.—Es indudable que la mortalidad puede ser estudiada por el número absoluto de los sucesos ocurridos de esta clase. Así podremos ver qué cifras máximas alcanza en el transcurso de los años, o bien si se trata de un período prefijado de tiempo, cual es la edad en la que más defunciones se producen o en la que menos, y aun se puede llegar a la ley que rige la mortalidad por tiempos o por edades. Pero lo que no podemos hacer, sin recurrir a otros hechos, es comparar los valores que la mortalidad toma en varias localidades, en varias épocas, en varias edades, etc., ya que para ello será preciso trabajar con valores que relacionen las defunciones ocurridas con las poblaciones en que se produjeron. Estas relaciones han sido fijadas siguiendo criterios distintos, por lo que es preciso considerar conjuntos o *colectivos* de habitantes y de muertos diferentes.

Los *colectivos* de población más importantes son:

$$P_x^z \quad , \quad P_x^{z+1} \quad , \quad E_x^z$$

y entre los de muertos lo son los:

$$\alpha D_x^z + \delta D_x^{z+1} \quad , \quad \delta D_x^z + \alpha D_{x+1}^z \quad , \quad \alpha D_x^z + \delta D_x^z .$$

Recordemos (fig. 11) que P_x^z es la población que tiene x años en principio del año z. P_x^{z+1} es la población que tiene x años en principio del año $z + 1$. E_x^z son los habitantes que teniendo $x - 1$ años de edad en principio del año z, cumplen x años durante el mismo.

En las defunciones αD_x^z son los que mueren con x años durante el año z, de los que en principio de dicho año tienen $x - 1$ años. δD_x^{z+1} los que fallecen con x años durante el año $z + 1$, de los que en principio del año $z + 1$ tienen x años, siendo la suma $\alpha D_x^z + \delta D_x^{z+1}$ los fallecimientos ocurridos durante los años z y $z + 1$ con la edad x procedentes de los E_x^z. Los δD_x^z son los fallecidos durante el año z con la edad x, y αD_{x+1}^z son los fallecidos durante el año z con la edad $x + 1$, procedentes ambos de los P_x^z. Por último, los αD_x^z , δD_x^z, ya definidos cuya suma no procede de ningún *colectivo* real.

TASAS DE MORTALIDAD.—La mayor o menor «repetición» del suceso, muerte, en cada colectivo de habitantes, caracteriza la mortalidad del mismo. Estudiados los habitantes por colectivos de edad, su mortalidad vendrá definida, no sólo por la importancia de la «repetición» del suceso, sino también por la importancia que tenga el colectivo de habitantes que se estudia.

Las debidas relaciones entre los colectivos de muertos y los colectivos de habitantes reciben el nombre de *tasas de muerte*.

Tanto en libros nacionales como en libros extranjeros, he visto usado, indistintamente, los términos *tasa de mortalidad* y *probabilidad de muerte*, a los que habremos de añadir el estadístico de *frecuencia relativa*. Unido esto a que figuras de tan relevante prestigio como FRECHET, DE MISES, CANTELLI, DE FINETTI, TORNIER, entre otros, encuentran posibilidades de discusión sobre el concepto de probabilidad, me plantearon un problema, que yo, modesto operador del azar, no me creo apto para resolver. Sin embargo, al tratar de formar unas *Tablas nacionales de Mortalidad*, tenía necesariamente que decidirme por una *tasa* que me sirviera de base en el cálculo de las mismas. Siendo la distinción de los tres términos, *tasa de mortalidad*, *probabilidad de muerte* y *frecuencia relativa*, lo que me permitió la elección.

Como norma de selección consideré el hecho de que no siempre la *frecuencia relativa* es o tiende a ser una *probabilidad*. Aceptando como *tasas de mortalidad* aquellas relaciones que, a más de ser *frecuencias relativas*, fueran o tendieran a ser *probabilidades de muerte*.

La razón por la que adopte este criterio es la de que, debiendo tomar en el campo actuarial la *tasa de mortalidad* un carácter de *frecuencia relativa*, por otro lado, por depender de un suceso, muerte, de azar, debía tender a una *probabilidad*.

Es indudable que para aceptar como *probabilidad* una expresión, se debe fijar antes las condiciones del suceso y en vista de ellas definir la *probabilidad* del mismo. Bien conocidas son las paradojas que el *Cálculo de Probabilidades* presenta, como la de BERTRAND, en las probabilidades geométricas, consecuencia de una incompleta definición del problema. Algo de esto, creo, ha debido pasar en la aceptación, como *probabilidades de muerte*, del gran número de fórmulas que se han utilizado o se han propuesto para el cálculo de *tasas de mortalidad*.

Todo ello ha contribuído a que el estudiante, en sus primeras lecturas sobre *probabilidades de muerte*, y ante el gran número de fórmulas que pretenden serlo, se desoriente y dude de la utilidad y de la ciencia de un *Cálculo de Probabilidades*, que permite tan variadas soluciones para un problema que él cree único.

En vez de establecer una relación y definirla como *probabilidad de muerte*, he analizado el problema no sólo en su marcha real, sino he descendido en el cálculo de la probabilidad a las *probabilidades unitarias*, como último elemento separable del conjunto probabilidad. Ello me ha permitido ver que el *Cálculo de Probabilidades* no es de una utilidad muy discutible en este problema de azar, sino que hasta la fecha se habían considerado como *probabilidades de muerte*, relaciones que, ni aun como *frecuencias relativas*, podían ser aceptadas:

Los tres tipos fundamentales de *tasas de mortalidad*, son:

$$\text{Tipo} \quad \mathrm{I} = \frac{\alpha D_x^z + \delta D_x^{z+1}}{E_x^z}$$

$$\text{Tipo II} = \frac{\delta D^z_x + \alpha D^z_{x+1}}{P^z_x}$$

$$\text{Tipo III} = \frac{\delta_{,} D^z_x + \alpha D^z_x}{P^z_x}$$

siendo ya dada la significación de los elementos que en ella intervienen.

TRES COLECTIVOS DE POBLACIÓN Y TRES COLECTIVOS DE MUERTOS.—Voy a separarme del camino seguido hasta ahora en la exposición para emprender otro que nos llevará fácilmente a la selección de tasas de muerte.

Comenzaré por presentar tres colectivos teóricos de población con sus correspondientes de muertos, producidos éstos en los primeros durante un período de observación.

Los colectivos son:

Colectivo 1.º—Lo constituye una población de P habitantes (fig. 13) que durante la observación sólo sufre las bajas producidas por muerte. Pasa, por tanto, el colectivo por los estados P, P — 1, P — 2, ... P — D. Es decir, comenzando por P habitantes, termina por P — D, si D es el número de muertes registradas en su seno durante el período de observación.

Notaremos que aquí los valores del colectivo población no vienen expresados en el gráfico por áreas, sino por ordenadas, indicándonos los intervalos en la abcisa, que durante el tiempo correspondiente a ellos el colectivo de población se ha mantenido constantemente igual a P, P — 1, P — 2, ... P — D.

Colectivo 2.º—Está formado por una población de P habitantes, que durante el período de observación sufre bajas por defunción, pero cubiertas inmediatamente que se producen por otros habitantes que entran a formar parte del colectivo. El colectivo, por tanto, comienza por P habitantes y termina también con P, siendo D el número de fallecidos durante el tiempo de observación.

Por medio de horizontales se ha establecido la separación entre los habitantes que han entrado en el colectivo para suplir las bajas producidas en el mismo y los que quedan de los que formaron el colectivo inicial.

Colectivo 3.º—Este comienza con un valor cero y luego ingresan individuos en él; a medida que transcurre el período de observación y cuando llega a su mitad en que toma el valor P, empiezan a salir habi-

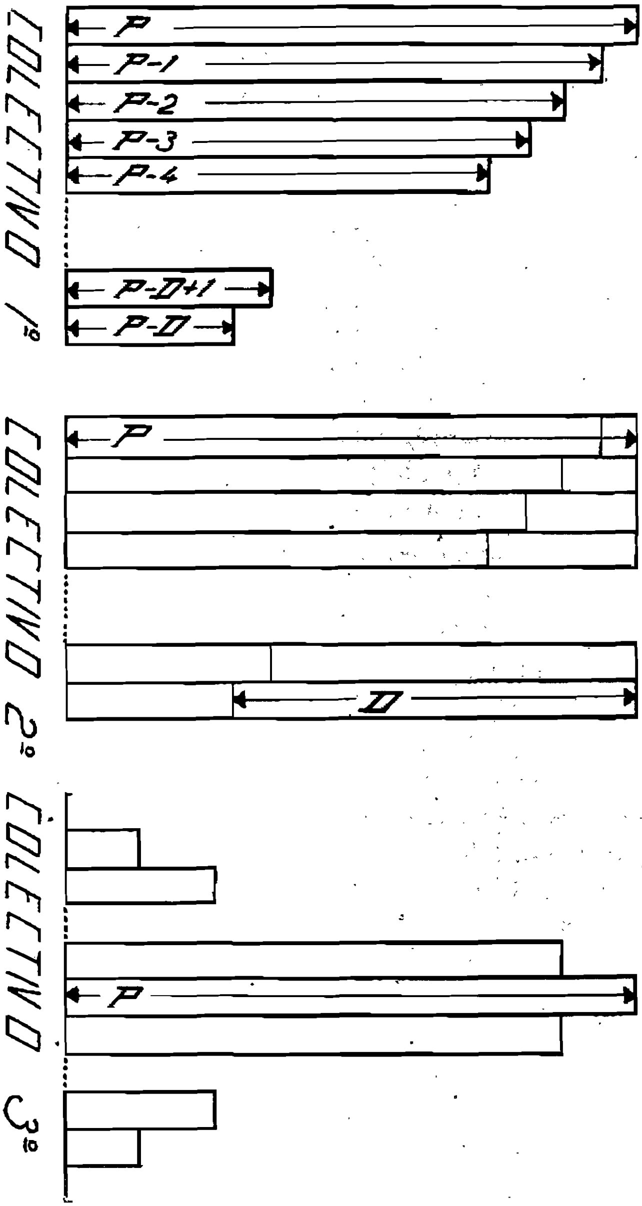

COLECTIVO 1º
P
P-1
P-2
P-3
P-4
P-D+1
P-D
COLECTIVO 2º
P
D
COLECTIVO 3º
P

FIG. 13

tantes del colectivo hasta que, al final, habiendo salido todos, toma el valor cero. A más de estas variaciones se producen D defunciones durante la observación.

Los valores de P, así como los de D que figuran en estos colectivos teóricos, se suponen ser iguales en los tres y el período de observación el mismo. En estas condiciones nos hacemos la pregunta: ¿La relación:

$$\frac{D}{P}$$

expresa la probabilidad de muerte en los tres colectivos, es su frecuencia relativa o no expresa nada?

La forma de variar los tres colectivos teóricos de población es la misma, en líneas generales, que la de los conjuntos de expuestos al riesgo de morir que sirven de base para establecer las tasas-tipo de mortalidad.

En la tasa de muerte segunda varía el conjunto de habitantes expuestos al riesgo de morir, lo mismo que lo hace el primer colectivo teórico de población, salvo que P y D toman los valores P_x^z y $\delta D_x^z + \alpha D_{x+1}^z$, admitiéndose como probabilidad de muerte la:

$$\frac{D}{P}$$

o sea en este caso

$$\frac{\delta D_x^z + \alpha D_{x+1}^z}{P_x^z}$$

La tercera tasa de muerte varía su conjunto de habitantes, lo mismo que lo hace el segundo colectivo teórico de población, tomando los P y D los valores P_x^z y $\delta D_x^z + \alpha D_x^z$, admitiéndose como probabilidad de muerte la misma

$$\frac{D}{P}$$

que aquí toma el valor

$$\frac{\delta D_x^z + \alpha D_x^z}{P_x^z}$$

Por último, en la tasa primera varía su colectivo de habitantes, lo mismo que el tercero de los teóricos de población, siendo E_x^z y $\alpha D_x^z +$

δD_x^{z+1} los valores que toma P y D en este caso. Se admite, como en las anteriores, que su probabilidad de muerte es la misma

$$\frac{D}{P}$$

correspondiente a

$$\frac{\alpha D_x^z + \delta D_x^{z+1}}{E_x^z}$$

para esta tasa.

Colectivos reales de población.—Los colectivos teóricos sólo permiten un primer examen comparativo de las tres tasas-tipo de mortalidad, pero no un análisis del fenómeno por diferir en pequeños detalles de la realidad del mismo. Profundizando en la investigación, vamos a estudiar las variaciones reales que experimentan los colectivos de población, expuestos al riesgo de morir, en cada tasa-tipo de mortalidad.

Comenzaré por la tasa segunda. En mi gráfico, y mejor aún en la figura 14, se aprecia la variación del colectivo P_x^z de población, que esta tasa considera como expuesto al riesgo de morir. Sometidos a la observación, los habitantes P_x^z, que tienen edades comprendidas entre x y $x + 1$ años, en principio del año z, disminuyen en número durante él por las muertes que en ellos se producen. El área de ABFH $= P_x^z$ va decreciendo por fajas elementales hasta llegar, al final del año, al área de CGFH $= P_{x+1}^{z+1}$. El área de ABGC nos da la diferencia $P_x^z - P_{x+1}^{z+1}$, es decir, el número de muertes ocurridas en el transcurso del año z en el colectivo inicial P_x^z. Este área se descompone en las dos ACB $=$ δD_x^z y BGC $= \alpha D_{x+1}^z$, representación, la primera, de los muertos con edad x y de los muertos con edad $x + 1$, la segunda. Proceden las dos del colectivo P_x^z.

Esto mismo puede verse (fig. 15) en el fragmento del *esquema de Lexis*. Las líneas de vida cortan simultáneamente a la línea P_x^z y cuyo número nos da el de habitantes en principio del año z. Continuando su movimiento ascendente, algunas terminan dentro del triángulo δD_x^z, las que nos dan los fallecidos con la edad x. De una manera sucesiva cortan las líneas de vida a la recta E_{x+1}^z, empezando por la izquierda, pasando los habitantes que representan a la edad $x + 1$. Dentro del triángulo αD_{x+1}^z terminan algunas líneas de vida, cuyo número es el de fallecidos a la edad $x + 1$. Por último, y simultáneamente llegan a la [illegible] dando el conjunto de líneas de vida que la cortan el número de habitantes que en principio del año $z + 1$ quedan de los que iniciaron el año.

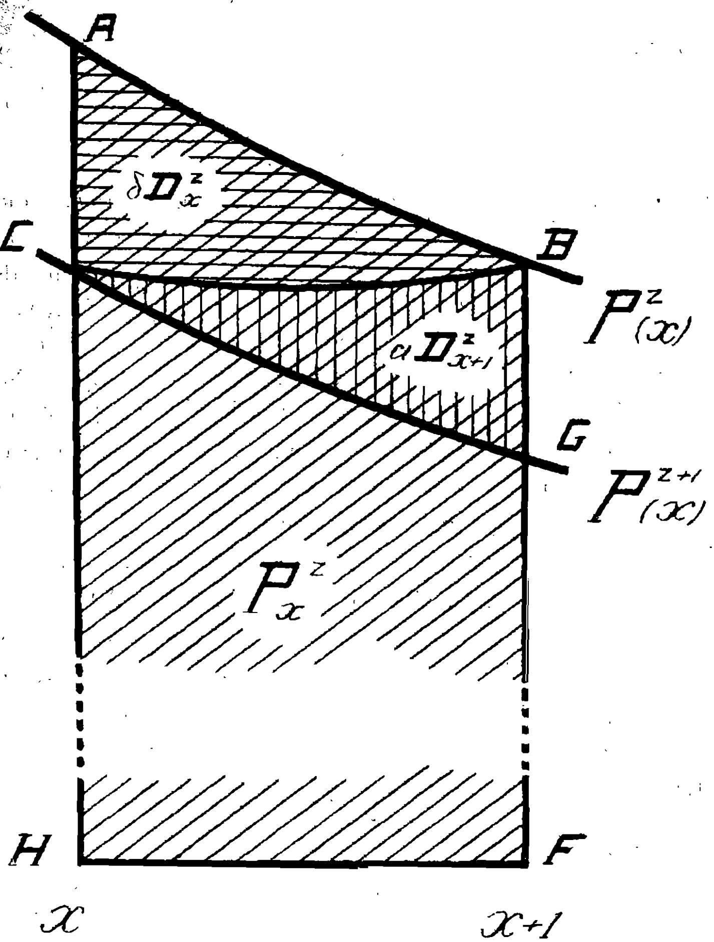
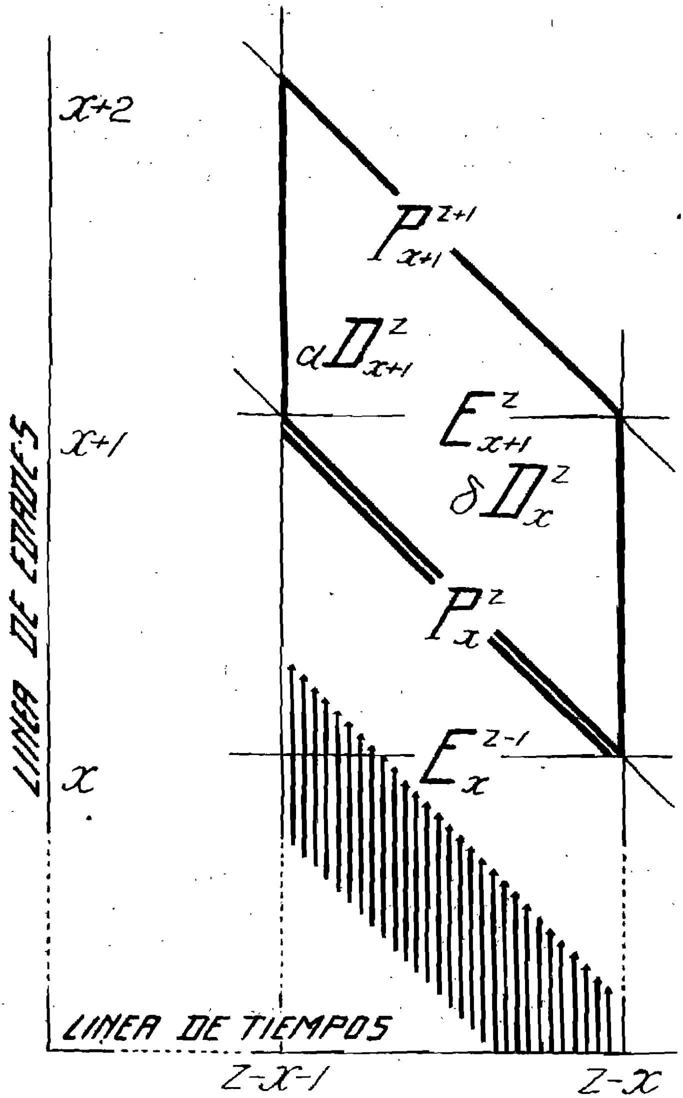


FIG. 14

Tanto en un gráfico como en otro las únicas variaciones en cantidad que se aprecian son las de muerte y éstas en forma sucesiva.

El colectivo teórico primero de población tiene una pequeña diferencia con el real (fig. 16), y es la de haber supuesto en aquél que los muertos se producían a intervalos constantes de tiempo, no siendo así en la realidad.



LINEA DE EDADES
x+2
x+1
x
LINEA DE TIEMPOS
Z-X-1
Z-X
P^{z+1}_{x+1}
aD^{z}_{x+1}
E^{z}_{x+1}
δD^{z}_{x}
P^{z}_{x}
E^{z-1}_{x}

Fig. 15

A la tasa de mortalidad tercera habremos de oponerla que tanto en

$$\frac{\alpha\, D^z_x + \delta\, D^z_x}{P^z_x}$$

mi gráfico (fig. 17) como en el *esquema de Lexis* (fig. 18) puede verse, que si bien los $\delta\, D^z_x$ muertes proceden de los P^z_x habitantes, no ocurre lo mismo con las $\alpha\, D^z_x$, que proceden de los P^z_{x-1}. Como quiera que P^z_x

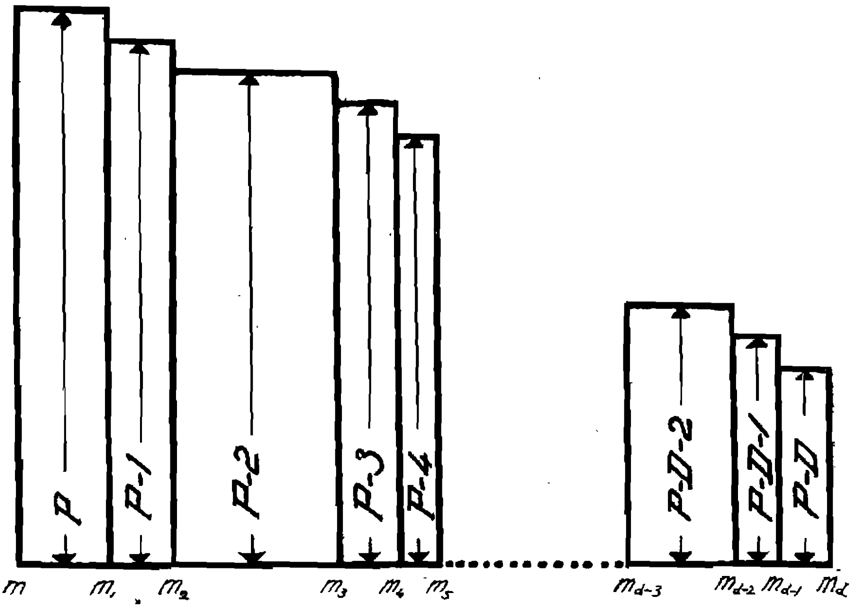


FIG. 16

y P^z_{x-1} son dos colectivos de población, de edades consecutivas no ligadas por ninguna relación a los muertos de uno de ellos, no nos es lícito (sin otras razones) aceptar como tasa la llamada tasa tercera de mortalidad.

Se puede deducir la tasa tercera después de establecer la hipótesis

$$P^z_x = P^z_{x-1}$$

y las dos frecuencias relativas

$$\frac{\alpha\, D^z_x}{P^z_{x-1} = P^z_x} \qquad \text{y} \qquad \frac{\delta\, D^z_x}{P^z_x}$$

darán, por su suma, la tasa tercera de mortalidad. Aunque lo lícito de esta transformación es muy discutible, la consideraremos como válida.

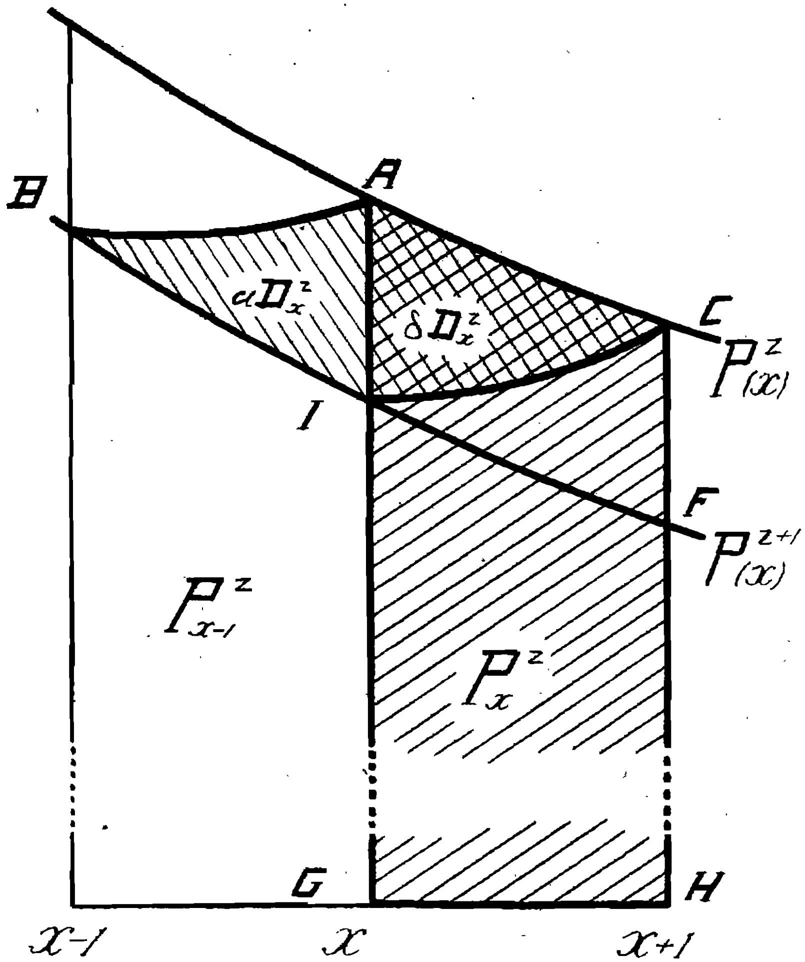


FIG. 17

Los P_{x-1}^z habitantes vendrán disminuídos, a principio del año $z+1$ en los muertos producidos en ellos, por lo cual

$$P_x^{z+1} < P_{x-1}^z$$

y como consecuencia de la hipótesis admitida

$$P_x^{z+1} < P_x^z$$

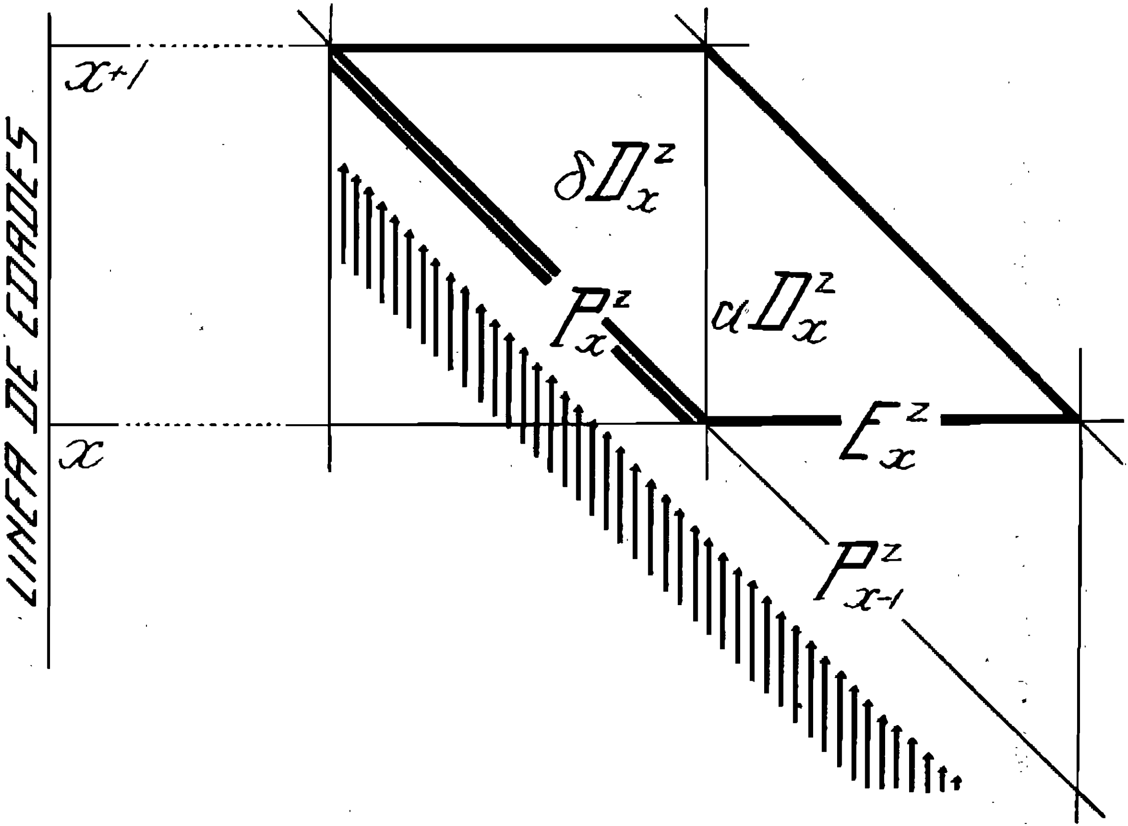


FIG. 18

Los dos miembros de esta desigualdad expresan los habitantes, que de la misma edad x, existen en la Nación, en dos años consecutivos z y $z + 1$ y según ella son en mayor número los de un año que los del año siguiente. Hecho contrario a la realidad demográfica, que establece el incremento anual de población. En consecuencia la hipótesis de igualdad de población no es posible admitirla.

Aun es posible otra deducción de la fórmula, fundada en la manera de variar su colectivo básico de población. Este colectivo viene dado por el conjunto de habitantes que tienen x años y sujetas sus variaciones a los necesarios cambios, para que en todo el año z no figuren, en el colectivo, más que habitantes de x años. Comienza el colectivo por los P^z_x. Salen después habitantes, tanto por fallecer como por cumplir $x + 1$ años. Los fallecidos serán, siempre, con la edad x por no componerse de otros el colectivo. A medida que salen habitantes entran otros que procedentes de P^z_{x-1} cumplieron los x años. Ahora hacemos la hipótesis, de que los que entran son en igual número a los que salen, por tanto, la población del colectivo básico permanece constantemente igual, en número, a P^z_x de donde procederán los muertos $\alpha\, D^z_x$ así como los $\delta\, D^z_x$ y la relación entre muertos y población dará la tasa tercera. Pero la hipótesis de que la población de x años, sujeta al riesgo de morir, permanece constante, es falsa. Veamos (fig. 19) las variaciones reales del colectivo básico. Comienza éste por $P = P^z_x$. Al final del tiempo $m\, m_1$ se produce la primera muerte y la segunda, después del intervalo $m_1\, m_2$ y así sucesivamente hasta la última. Por cumplir años, salen los marcados en el dibujo por las letras C. Los que entran, por cumplir los x años, son en tal número, que no sólo cubren las bajas producidas por C y D sino que originan un aumento continuado, del colectivo básico de población. Aumento que viene designado en el dibujo por las A. De esta forma, el colectivo que empezó con P^z_x habitantes no se mantiene constante, en número, durante el año z. Esta manera de variar el colectivo es la real que presenta el fenómeno demográfico del incremento anual de la población.

Al ser falsos los supuestos admitidos para explicar la tasa tercera de mortalidad, pensé que ella carece de tal carácter y que como consecuencia de lo expuesto ni aun como frecuencia relativa puede aceptarse.

Voy, por último, a estudiar, en la tasa primera de mortalidad, las variaciones reales de su colectivo básico.

En mi gráfico (fig. 20) tenemos representado el colectivo E^z_x por el

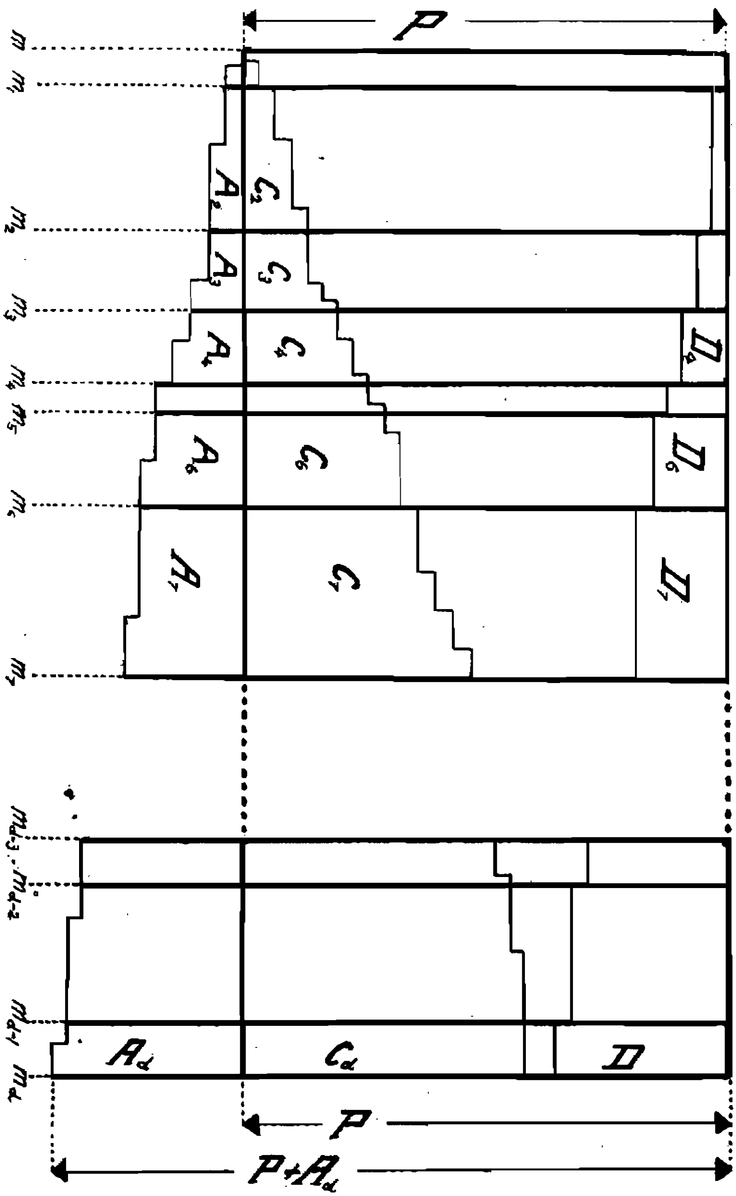

P
m
m1
m2
m3
m4
m5
m6
m7
A2
C2
A3
C3
A4
C4
D4
A6
C6
D6
A7
C7
D7
md-3
md-2
md-1
md
Ad
Cd
D
P
P+Ad

FIG. 19

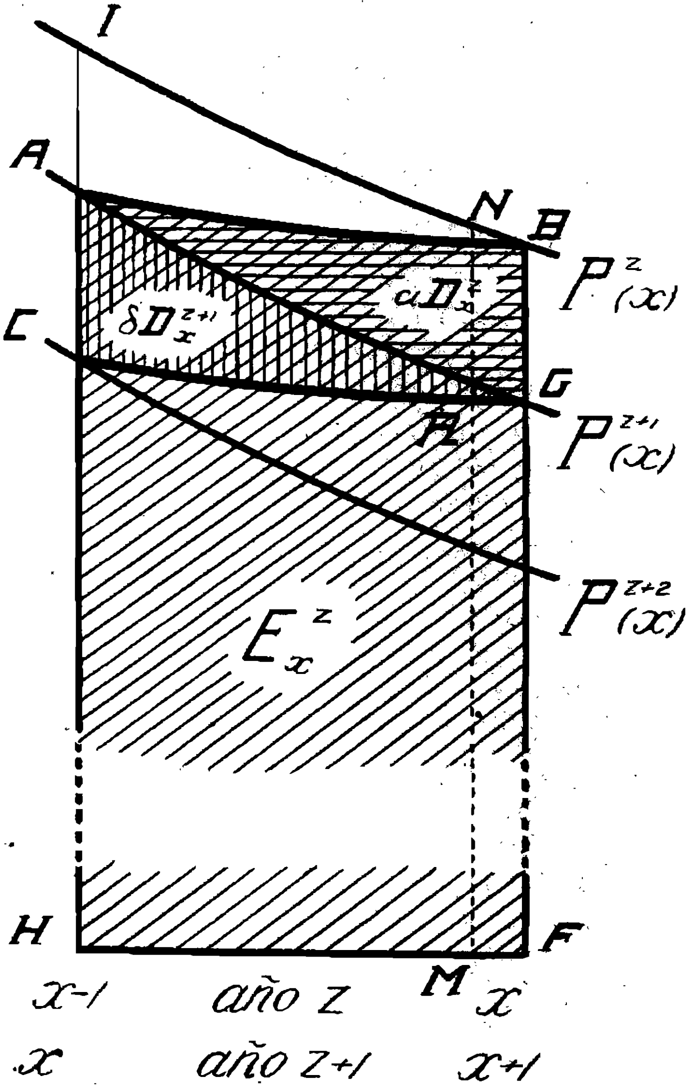

I
A
N
B
C
G
R
H
F
M
$P_{(x)}^{z}$
$P_{(x)}^{z+1}$
$P_{(x)}^{z+2}$
δD_x^{z+1}
$a D_x^z$
E_x^z
$x-1$
año z
x
x
año z+1
$x+1$

FIG. 20

área de A B F H y engendrado por la variación que experimenta el P_{x-1}^{z} durante el año z, a medida que fajas elementales como la N B F M pasan a la edad x. Empieza, por tanto, el colectivo por cero y pasan E_x^z habitantes a la edad x durante el año z. Mueren en el mismo año αD_x^z y quedan al final del año z

$$E_x^z - \alpha D_x^z = P_x^{z+1}$$

Durante el segundo año de observación $z+1$ no pueden pasar más habitantes al colectivo E_x^z por haberlo efectuado todos los P_{x-1}^{z} durante el primer año z. Comienzan a salir habitantes del colectivo básico por cumplir los $x+1$ años, haciéndolo por fajas elementales como la R G F M, que se restan del área de A G F H o valor de P_x^{z+1}. Mueren durante este segundo año, con la edad x, los δD_x^{z+1}, dados por el área de A G C. El colectivo varía, por tanto, desde cero al valor máximo P_x^{z+1}, que se da a finales del año z, decreciendo hasta llegar otra vez a cero al final del año $z+1$.

En el *esquema de Lexis* (fig. 21) las líneas de vida cortan, simultáneamente, a la recta P_{x-1}^{z} y sucesivamente a la E_x^z, empezando a cumplir años, los habitantes que representan, a partir de la izquierda hasta que al final del año z todas las líneas de vida han cortado a la E_x^z y se encuentran, sus extremos, en la recta P_x^{z+1}. Hasta esta recta sólo llegan una parte de las líneas de vida que cortaron a E_x^z. Siendo las que faltan, αD_x^z, muertos con x años durante el año z. A partir de la posición P_x^{z+1}, de los extremos de las líneas de vida comienzan éstas a cortar sucesivamente a la línea E_{x+1}^{z+1}, cumpliendo $x+1$ años los habitantes que representan. Como quiera que no todas las líneas de vida que pasaron por P_x^{z+1} llegan a cortar a E_{x+1}^{z+1}, las que faltan, δD_x^{z+1}, nos dan los fallecidos, en el segundo año de observación, con la edad x.

El colectivo real de población (fig. 22) comienza por cero, punto m, aumenta por ingresar habitantes hasta el punto m_1 en que se produce la primera muerte. Del mismo modo se registran los sucesivos cambios, hasta que al llegar al final del año z el colectivo toma el valor:

$$E_x^z - \alpha D_x^z = P_x^{z+1}$$

disminuyendo, a partir de aquí, hasta llegar de nuevo al valor cero.

Debo hacer notar que debía figurar en el dibujo el caso de descender, en una unidad, las ordenadas siguientes a la que registra una muerte

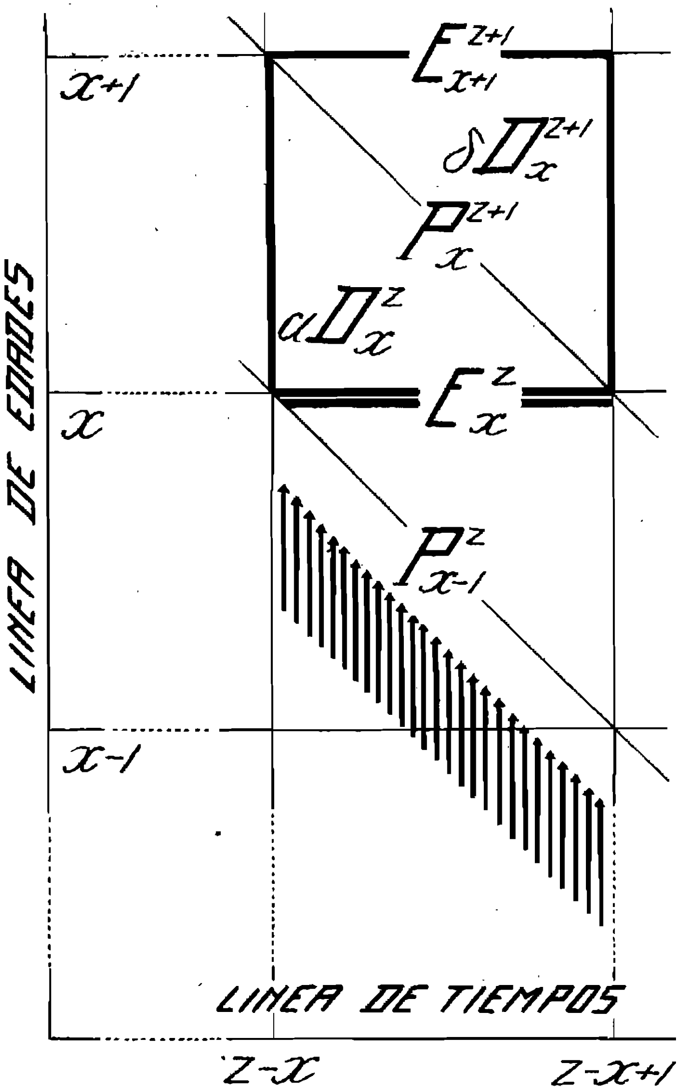


FIG. 21

(puntos *m*) en la parte ascendente del colectivo, caso que se presenta, cuando no ingresa un habitante inmediatamente que se produce una muerte. En figuras posteriores se incluirá este caso.

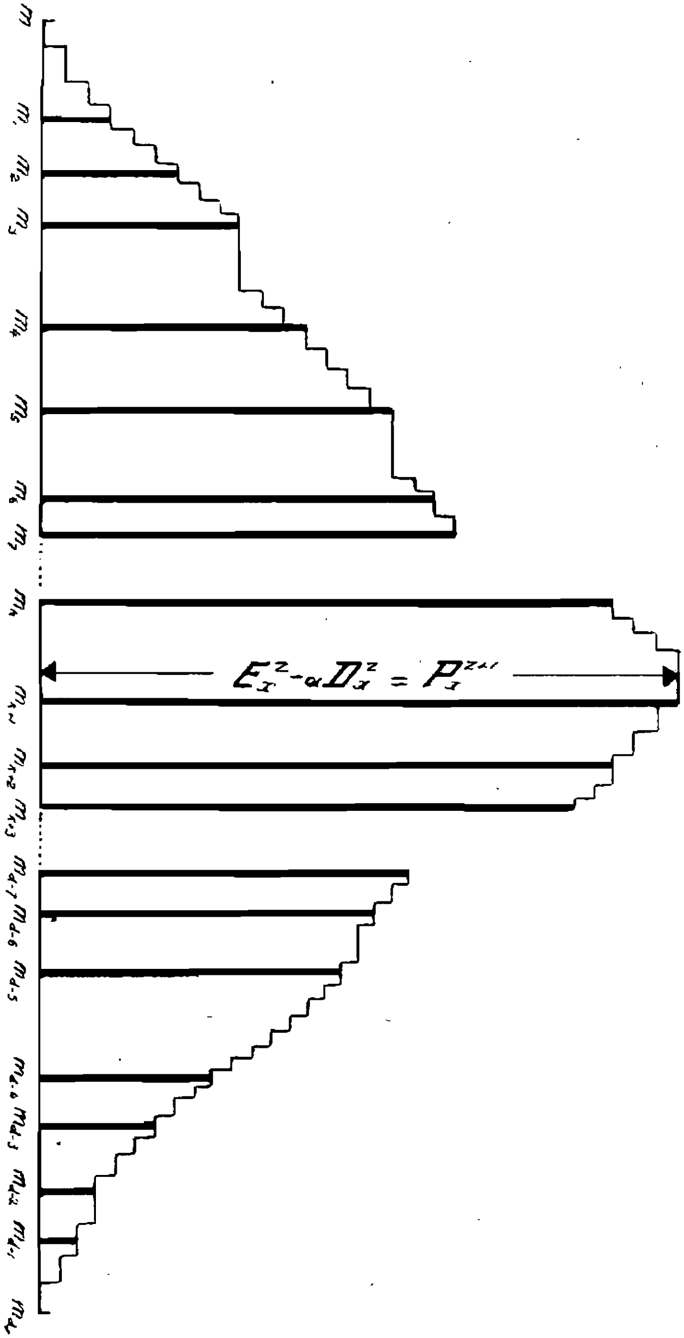


FIG. 22

De este previo análisis deducimos que la tasa segunda de mortalidad relaciona un colectivo de habitantes con todos los muertos producidos en su seno durante un período de observación, por lo que la acepto como frecuencia relativa, a resultas de poderla considerar también como probabilidad de muerte.

La tasa de mortalidad tercera, ni por la falta de conexión entre los colectivos de población y de muertos que relaciona, ni por la falsedad de las hipótesis sustentadas para explicarla, podemos incluirla entre las frecuencias relativas.

Por último, la tasa de mortalidad primera relaciona los habitantes que cumplen los x años con los muertos que de ellos se producen, a esa edad, en un período de tiempo. Tiene, por tanto, carácter de frecuencia relativa.

Probabilidades de muerte.—Vamos a pasar al cálculo de la probabilidad que de morir tienen los habitantes en los colectivos reales de población.

Calcularemos, en primer lugar, la probabilidad de muerte que corresponde al colectivo real de población de la tasa segunda de mortalidad. Las variaciones de su colectivo (fig. 16) ya han sido descritas.

Cada uno de los habitantes que forman el colectivo de población puede morir o sobrevivir al término del año de observación. Si sólo estudiamos la probabilidad de morir, descartemos la de ser superviviente al final del año, que podrá ser complementaria de la de morir (existen casos en la mortalidad que así no ocurre). Un habitante del colectivo es considerado al azar y calculada la probabilidad que tiene de morir durante un año, sabiendo que se producen D muertes. Esta probabilidad será la misma que tienen todos los habitantes en principio del año de observación.

El habitante considerado puede ser el fallecido cuando ocurre la primera muerte, la segunda... o la última. Estará, por tanto, expuesto al riesgo de morir D veces.

La probabilidad que de morir tiene cuando ocurre la primera muerte, es:

$$\frac{1}{P}$$

La de morir en segundo lugar lleva implícita la de no haber muerto en primero

$$\frac{P-1}{P}$$

afectada de la probabilidad libre de morir en segundo lugar

$$\frac{1}{P - 1}$$

[después de la primera muerte el colectivo sólo se compone de P — 1 habitantes].

La probabilidad de morir en segundo lugar será

$$\frac{P-1}{P} \cdot \frac{1}{P-1} = \frac{1}{P}$$

De igual forma la probabilidad de morir en tercer lugar se formará por las probabilidades libres de sobrevivir en la primera y segunda muerte y la de morir en la tercera, o sea

$$\frac{P-1}{P} \cdot \frac{P-2}{P-1} \cdot \frac{1}{P-2} = \frac{1}{P}$$

Las probabilidades de morir en cuarto, quinto..., último lugar serán todas, como las tres primeras, iguales a

$$\frac{1}{P}$$

Son sucesos incompatibles el fallecer el habitante en cada una de las D muertes registradas. Todos los habitantes están desde el principio de la observación expuestos al riesgo de morir en todas las muertes ocurridas.

Como el suceso, muerte del habitante, puede producirse por D causas diferentes e incompatibles, su probabilidad será la suma de las probabilidades que resultan al actuar cada causa:

$$\frac{1}{P} + \frac{1}{P} + \frac{1}{P} + \dots\dots = \frac{D}{P}$$

en la que poniendo los valores que P y D toman en este caso, tendremos como probabilidad de muerte en el colectivo básico de la tasa segunda

$$\frac{1}{P_x^z} + \frac{1}{P_x^z} + \frac{1}{P_x^z} + \dots\dots = \frac{\delta\, D_x^z + \alpha\, D_{x+1}^z}{P_x^z}$$

Ya vimos que esta fórmula era frecuencia relativa y ahora la volvemos a encontrar como probabilidad de muerte, por lo que la admitimos como tasa de mortalidad y nos da la *probabilidad que de morir en un año tienen los habitantes* P_x^z.

Aceptada esta tasa en el campo actuarial y en el demográfico como probabilidad de muerte, sólo he pretendido al calcularla reafirmar tal carácter poniendo de manifiesto la propiedad aditiva que tiene.

Confirman también el carácter de probabilidad algunos resultados a que conducen determinados valores del colectivo de muertos. Cuando D toma el valor cero, ninguna muerte se produce en el colectivo, luego el suceso morir el habitante entra en lo «imposible», hecho que pone de manifiesto la tasa al dar a D el valor cero. También cuando D se hace igual a P mueren todos los habitantes del colectivo, luego tenemos la «certeza» de producirse el suceso, es decir, morir el habitante, sea cual fuere el considerado. A la misma conclusión llegamos al hacer D = P en la tasa.

Pudiera pensarse que los sumandos que nos dan la probabilidad del suceso no debieran ser probabilidades ligadas, sino probabilidades libres; entonces la probabilidad vendría dada por

$$\frac{1}{P} + \frac{1}{P-1} + \frac{1}{P-2} + \ldots\ldots + \frac{1}{P-D+1}$$

en la que comparando cada uno de sus sumandos con los

$$\frac{1}{P} + \frac{1}{P} + \frac{1}{P} + \ldots\ldots = \frac{D}{P}$$

vemos son mayores que éstos y en igual número; por tanto, la probabilidad será mayor que

$$\frac{D}{P}$$

dándose entonces el caso de que, al establecer el supuesto de morir todos sus habitantes, la probabilidad sería mayor que la unidad, es decir, superior a la «certeza», consecuencia absurda debida a lo erróneo de admitir probabilidades libres en vez de ligadas.

Por último, si hubiera sido calculada la probabilidad del suceso por una probabilidad simple, ésta nos daría la misma que la ya obtenida. La analogía de resultados también expresa aquí analogía en los fenómenos. Hecha abstracción de los intervalos de tiempo, no consignando de él, más que la condición de producirse en un año los muertos registrados, lo mismo da estudiar las muertes una a una que en bloque. Esto no es más que el clásico ejemplo de las bolas en las pruebas repetidas, salvo que aquí no se repone bola después de sacada.

El estudio de la probabilidad de muerte reviste más importancia que en la anterior tasa, en la tasa primera de mortalidad, ya que de ella se deriva la fórmula

$$\frac{D}{P - \frac{D}{2}}$$

que ha servido de base para el cálculo de la mayoría de las Tablas nacionales de Mortalidad.

Es puesto de manifiesto en los libros que de la tasa primera tratan, el inconveniente que presenta su cálculo al exigir la observación de dos años de muertos. A la mayor dificultad de cálculo se une el que trabajando sobre dos años de observación, se aplica después a períodos de un año. Este inconveniente puede, en parte, ser salvado, suponiendo que las defunciones que se registran a la edad x, en un año, son la mitad de las producidas en los dos años en el colectivo P_{x-1}^{z} con las edades $x - 1$, x y $x + 1$ años.

No voy a calcular la probabilidad de muerte en el colectivo básico de habitantes, sólo le compararé con otro cuya probabilidadde muerte nos es ya conocida. Con lo que creo bastará para ver no puede admitirse como probabilidad de muerte, la tasa primera de mortalidad.

El colectivo buscado ha de ser tal, que variando como lo hace el colectivo básico de la tasa segunda, su probabilidad de muerte sea dada por la misma fórmula que expresa la tasa primera.

El colectivo representado por líneas punteadas de la figura 23, resuelve el problema. Por líneas gruesas viene dado el colectivo básico de la tasa primera, en el que he dibujado los casos de descender una unidad en la parte ascendente, caso que ya indiqué podía ocurrir.

Comienza el colectivo de comparación por E_{x}^{z} habitantes. Decrece este número, unidad a unidad, por las muertes que se supone ocurren en los mismos intervalos de tiempo en que así sucede en el colectivo básico de la tasa primera. Llega al final del año z habiéndose producido ${}_{\alpha}D_{x}^{z}$ muertes (las mismas que en el colectivo básico y en los mismos tiempos) quedando en aquél.

$$E_{x}^{z} - \alpha\, D_{x}^{z} = P_{x}^{z+1}$$

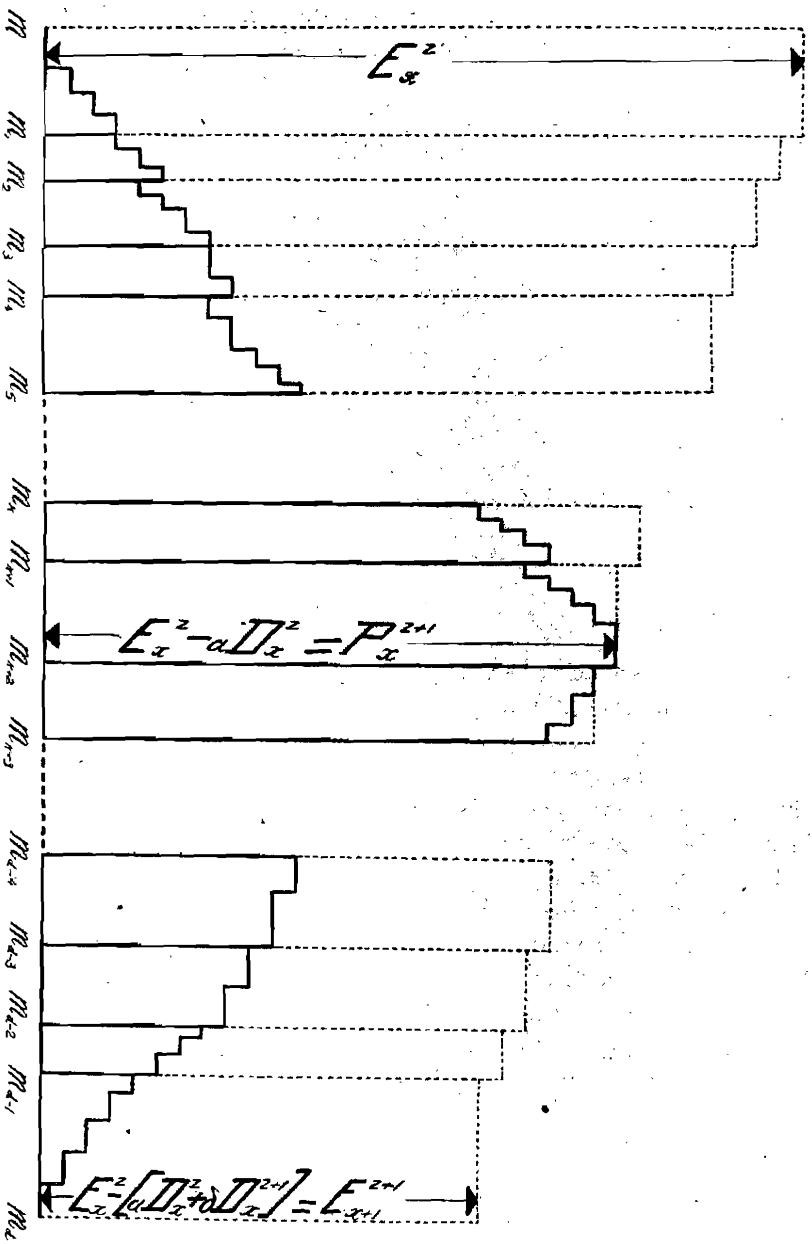

M
M_1
M_2
M_3
M_4
M_5
M_x
M_{x+1}
M_{x+2}
M_{x+3}
M_{d-4}
M_{d-3}
M_{d-2}
M_{d-1}
M_d
$E_x^{z'}$
$E_x^z - aD_x^z = P_x^{z+1}$
$E_x^z - [aD_x^z + bD_x^{z+1}] = E_{x+1}^{z+1}$

FIG. 23

habitantes. Continúa decreciendo hasta llegar al final del año $z + 1$ que, como han fallecido δD_x^{z+1} habitantes más (los mismos y en los mismos tiempos que en el colectivo básico), quedarán:

$$E_x^z - \alpha D_x^z - \delta D_x^{z+1} = E_{x+1}^{z+1}$$

Los dos colectivos, el básico y el de comparación, sólo tienen un punto de contacto que se presenta a final del año z y con un valor P_x^{z+1}, siendo en cualquier otro momento mayor el número de habitantes que registra el colectivo de comparación que el básico.

En la obtención de la probabilidad de muerte de la tasa segunda hemos visto que siempre que un colectivo varía, como el básico de ella, su probabilidad de muerte viene dada por la relación entre los habitantes iniciales y los muertos ocurridos en ellos durante el período de observación. Así la probabilidad de muerte del colectivo de comparación es:

$$\frac{\alpha D_x^z + \delta D_x^{z+1}}{E_x^z}$$

es decir, la misma que se supone ser probabilidad de muerte en el colectivo de líneas continuas de la figura 23, o sea la tasa primera de mortalidad.

Ante la igualdad del número de muertos ocurridos en los dos colectivos en los mismos períodos de tiempo y variando de manera tan diferente ambos, que siempre el de comparación tiene más habitantes que el básico, ¿se puede admitir que la misma fórmula exprese la probabilidad de muerte en los dos?

La simple inspección de la figura 23 bastaría para ver que no es así. No obstante, iniciaré el cálculo de las probabilidades elementales del colectivo básico de la tasa primera.

Al llegar el colectivo básico al punto m_1 (fig. 24) sucede la primera muerte y se encuentran en el C_1 habitantes. La probabilidad de muerte de los habitantes en este perío de tiempo, $m\,m_1$, es:

$$P_{c_1}^1 = \frac{1}{C_1} > \frac{1}{E_x^z} = P_E^1$$

En el punto m_2 ocurre la segunda muerte, y aquí hemos de considerar dos grupos de habitantes: unos los C_1, que estaban cuando ocu-

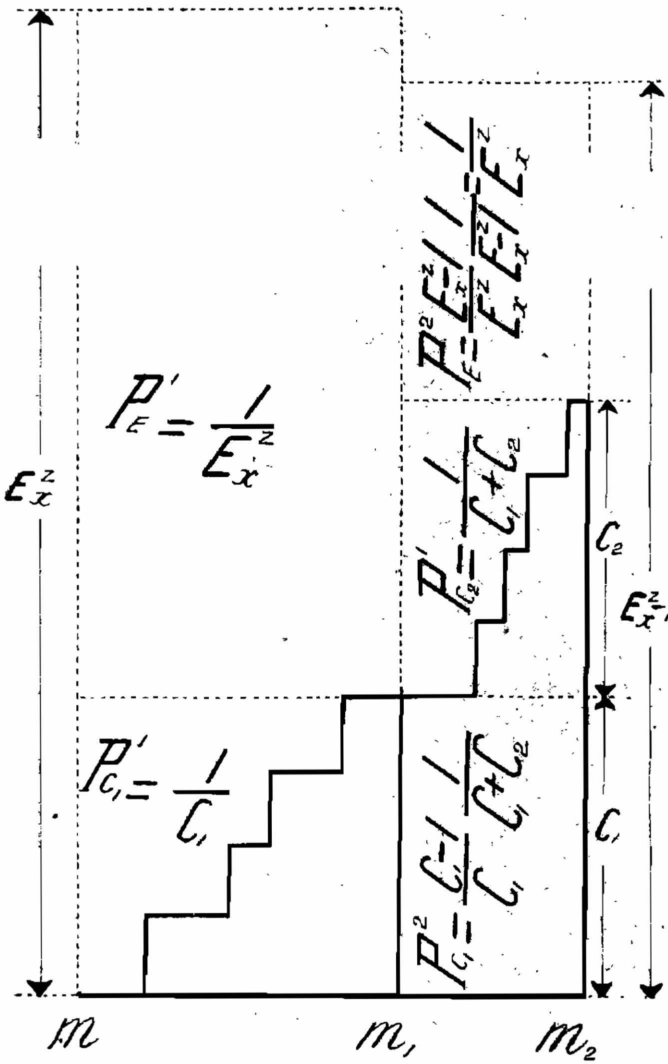

P'E = 1/Ex^z
Ex^z
P'C1 = 1/C1
P2C1 = (C1-1)/C1 · 1/(C1+C2)
P'C2 = 1/(C1+C2)
P2E = (Ex^z-1)/Ex^z · 1/(Ex^z-1) = 1/Ex^z
C2
C1
Ex^z-1
m
m1
m2

FIG. 24

rrió la primera muerte, y otro los C_2, que han entrado en el período de tiempo transcurrido entre la primera y segunda muerte $m_1 m_2$. Las dos probabilidades que de morir tienen los habitantes de los grupos C_1 y C_2 en la segunda muerte son:

$$P^2_{c_1} = \frac{C_1 - 1}{C_1} \cdot \frac{1}{C_1 + C_2} \;\;,\;\; P^1_{c_2} = \frac{1}{C_1 + C_2}$$

Las probabilidades de muerte en cualquier suceso del colectivo de comparación es

$$\frac{1}{E^z_x} = \frac{E^z_x - 1}{E^z_x} \cdot \frac{1}{E^z_x - 1} = \frac{E^z_x - 1}{E^z_x} \cdot \frac{E^z_x - 2}{E^z_x - 1} \cdot \frac{1}{E^z_x - 2} = \cdots\cdots$$

y como

$$C_1 < C_1 + C_2 < \cdots\cdots < E^z_x$$

tendremos

$$P^1_{c_1} = \frac{1}{C_1} > \frac{1}{E^z_x} = P^1_E$$

$$P^2_{c_1} = \frac{C_1 - 1}{C_1} \cdot \frac{1}{C_1 + C_2} > \frac{E^z_x - 1}{E^z_x} \cdot \frac{1}{E^z_x - 1} = \frac{1}{E^z_x} = P^2_E$$

$$P^1_{c_2} = \frac{1}{C_1 + C_2} > \frac{1}{E^z_x} = P^2_E$$

Que nos dice que todas las probabilidades que de morir tienen los habitantes de los distintos grupos del colectivo básico de la tasa primera son, en todo momento, mayores que las que tienen los habitantes del colectivo de comparación, y como la probabilidad de muerte de éste es:

$$\frac{\alpha\, D^z_x + \delta\, D^{z+1}_x}{E^z_x}$$

la que se trata de encontrar será mayor que ella, es decir, no podrá ser la misma.

Haremos notar que en ningún momento del año de observación el colectivo básico de la tasa primera tiene E_x^z habitantes.

Si lo que se pretendía con la tasa primera era calcular la *probabilidad que de morir, a los x años de edad, durante un año, tienen los habitantes de esa edad,* más fácilmente la obtendríamos por el cálculo de la probabilidad de muerte del colectivo real de habitantes (fig. 19) de la tasa tercera de mortalidad, en cuyo cálculo no entro por varias razones: una, por ser esta probabilidad de más utilidad demográfica que actuarial; otra, la misma por la que he evitado el cálculo de la probabilidad de muerte en el colectivo básico de la tasa primera, y es por la que tanto para un caso como para otro preciso presentar las *probabilidades unitarias* y sus *elementales,* que me obligarían a una ampliación en estas conferencias, removiendo conceptos y fórmulas tan fundamentales en el *Cálculo de Probabilidades* como la de las pruebas repetidas en el esquema de Bernoulli.

Como resultado del análisis hecho he deducido que *la única tasa de mortalidad que satisface las dos condiciones de ser frecuencia relativa y probabilidad de muerte, es la tasa de mortalidad tipo segundo* y, por tanto la aceptada para el cálculo de las *Tablas nacionales de Mortalidad.*

Tasa de mortalidad empleada.—Al tratar de mi representación gráfica, indiqué que se expondría un método en la segunda conferencia que permitiría obtener la función de población y mortalidad y, por consiguiente, la integral de la misma entre intervalos daría el número de habitantes y de muertos de que consta cada colectivo de edad.

En esas condiciones la tasa de mortalidad segunda

$$\frac{\delta D_x^z + \alpha D_{x+1}^z}{P_x^z}$$

tomará un carácter funcional.

Recordemos la hipótesis admitida en la teoría de la mortalidad y comprobada por la experiencia de que:

$$\delta D_x^z = \alpha D_x^z \qquad \text{y} \qquad \alpha D_{x+1}^z = \delta D_{x+1}^z$$

y que el número total de defunciones registradas en el movimiento demográfico durante el año z, con x años, es:

$$D_x^z = \delta D_x^z + \alpha D_x^z = 2 \, . \, \delta D_x^z$$

y con $x + 1$ años es

$$D^z_{x+1} = {}_1D^z_{x+1} + {}_\delta D^z_{x+1} = 2 \cdot {}_1 D^z_{x+1}$$

Calculadas las funciones de población y de mortalidad, los términos que entran en la tasa segunda toman los valores

$$\delta D^z_x = {}^1/_2 D^z_x = {}^1/_2 \int_x^{x+1} D(x)\, dx$$

$$\alpha \dot{D}^z_{x+1} = {}^1/_2 D^z_{x+1} = {}^1/_2 \int_{x+1}^{x+2} D(x)\, dx$$

$$P^z_x = \int_x^{x+1} P(x)\, dx$$

y la tasa segunda de mortalidad toma la forma

$$\frac{{}^1/_2 \int_x^{x+1} D(x)\, dx + {}^1/_2 \int_{x+1}^{x+2} D(x)\, dx}{\int_x^{x+1} P(x)\, dx} = \frac{\int_x^{x+2} D(x)\, dx}{2 \int_x^{x+1} P(x)\, dx} = T_m\ (x)$$

siendo T_m (x) la función de tasas de mortalidad y cuyo carácter funcional puede ponerse de manifiesto, dando un incremento a la edad

$$\frac{\int_{(x+\Delta x)}^{(x+\Delta x)+2} D(x)\, dx}{2 \int_{(x+\Delta x)}^{(x+\Delta x)+1} P(x)\, dx} = T_m (x + \Delta)$$

De la extraordinaria simplicidad de cálculo a que el uso de esta tasa conduce en la construcción de *Tablas nacionales de Mortalidad,* me ocuparé mañana, si ustedes tienen la paciencia de soportarme.

II.—Tablas de mortalidad.

En la conferencia de hoy vamos a pasar a un campo más operativo: al de la técnica de formación de tablas de mortalidad.

Aunque esta investigación es orientada fundamentalmente a la construcción de tablas nacionales, todo cuanto en ella se trata es también común a las tablas de experiencia restringida o de Compañías de Seguros, salvo, claro está, la parte de errores de inscripción de que éstas carecen.

Dos colectivos distintos pueden ser tomados como base para la formación de unas tablas: Unò el nacional, dado por un censo, en donde se incluyen todas las clases sociales, y otro, el de Compañías de Seguros, de límites más restringidos. Como es de suponer, los resultados difieren.

Sin embargo, las tablas nacionales de mortalidad tienen un interés en el campo actuarial. Los Seguros de Vida sin previo reconocimiento médico y el Seguro total sólo en ellas deben basarse, y, por otro lado, los colectivos de edades bajas son pequeños en las experiencias actuariales, siendo, por tanto, preciso conocer cómo se comportan en las tablas nacionales para servir de guía, al menos en esta parte, de la tabla de las Compañías de Seguros.

Errores censales de inscripcion.—Uno de los puntos de estudio más interesante en las tablas nacionales de mortalidad, es el relativo a errores censales de inscripción del dato edad.

Variadas son las causas a que éstos obedecen: unas veces por ignorancia del inscripto, por femenina coquetería y vanidad de anciano otras, y, las más, por ser facilitados o declarados los datos por personas ajenas al habitante. Como consecuencia de esto, los estados de clasificación por edades que salen de tabulación no pueden aceptarse como útiles para trabajos de laboratorio.

Es un hecho comprobado que cuando se desconoce la edad de una persona y sólo se tiene una idea aproximada de ella, se declaran siempre edades con determinada terminación, huyendo de otras terminadas en distinta cifra.

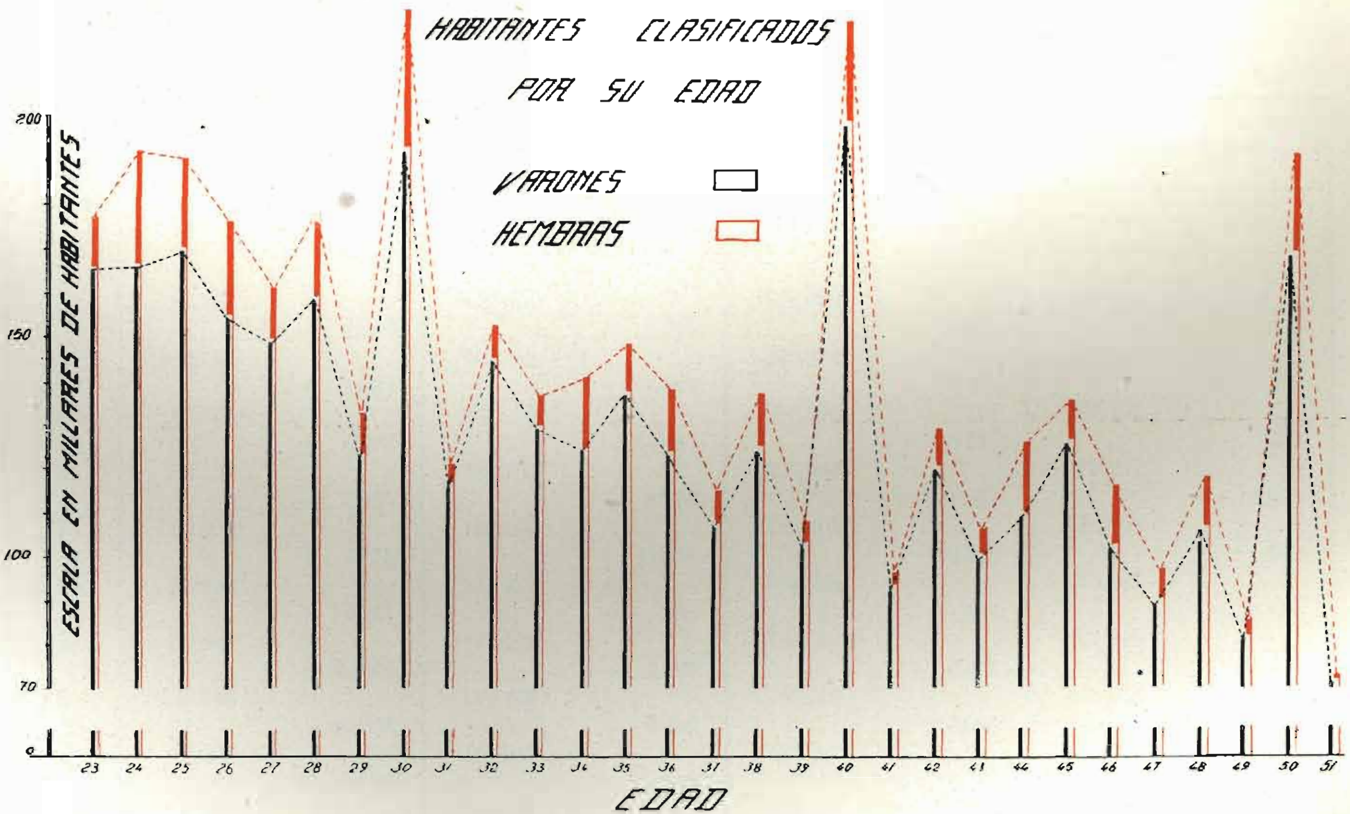
Lo mismo ocurre cuando voluntariamente se aumenta (aunque parezca raro se da el caso aún dentro del elemento femenino) o disminuye la edad.

En estadística demográfica, la aglomeración de habitantes en las edades terminadas en cero o cinco, es un hecho con el que siempre se ha de contar. Pero si esta influencia—error de edad—se ha de estudiar en su totalidad, no menos se ha de estimar la disminución en el número de habitantes que las edades terminadas en siete y tres presentan. Esto no es más que una tendencia de simpatía por parte del habitante a las terminaciones cero o cinco, a la vez que una tendencia contraria a las terminaciones tres y siete. No sólo son afectadas de tendencia las terminaciones cero, cinco, tres y siete de la edad, sino que todas lo son en mayor o menor cuantía.

Llevados a un gráfico (fig. 25) los valores, tal y como salen de tabulación, se aprecia una serie de máximos o «picos» en las edades de terminación cero o cinco, al mismo tiempo que unos mínimos, o depresiones, en las terminaciones impares, destacándose las terminaciones cero y cinco en los máximos y el siete en los mínimos. He llevado al gráfico sólo las edades 23 a 51 años, pero lo mismo ocurre, aunque más acentuado, en las edades sucesivas, y, en menor cuantía, en las edades de menos de 20 años.

El gráfico permite ver la mayor o menor simpatía que siente el habitante por esta o aquella terminación de edad, pero no sirve para apreciar este fenómeno en su totalidad. Para conseguir esto se han sumado los habitantes de todas las edades de la misma terminación registradas en el censo y obtenido su promedio. Al mismo tiempo se han calculado las primeras diferencias finitas (diferencia entre promedios de dos terminaciones consecutivas), las cuales nos dan la mayor o menor oscilación de una a otra terminación.

Estas cifras pueden verse en el cuadro I.



HABITANTES CLASIFICADOS
POR SU EDAD
VARONES
HEMBRAS
ESCALA EN MILLARES DE HABITANTES
200
150
100
70
0
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
EDAD

Fig. 25

Número promedio de habitantes varones registrados en los Censos de varias Naciones, clasificados por la cifra en que termina la edad declarada

(CUADRO I)

Terminación	ESPAÑA		FRANCIA		ITALIA		INGLATERRA	
	Promedio	Δ	Promedio	Δ	Promedio	Δ	Promedio	Δ
	99.152		210.654		217.494		176.774	
3		+ 2.975		— 5.396		— 7.821		+ 400
	102.127		205.258		209.673		177.174	
4		+ 4 980		+ 1.952		— 2.083		+ 318
	107.107		207.210		207.590		177.492	
5		— 8.979		— 18.450		— 3.731		— 280
	98.128		188.760		203.959		177.212	
6		— 7.205		— 6.170		— 21.551		— 4.958
	90.923		182.590		182.408		172.254	
7		+ 7.775		— 5.333		— 2.325		+ 1.368
	98.698		177.257		180.083		173.622	
8		— 18.051		— 865		— 361		— 5.177
	80.647		176.392		179.722		167.445	
9		+ 42.948		— 2.969		+ 6.496		— 461
	123.595		173.423		186.218		166.984	
0		— 49.307		+ 6.492		— 5.921		— 12 946
	74.288		179.915		180.297		154.038	
1		+ 13.590		— 1.354		— 4.561		+ 157
	87.878		178.561		175.736		154.195	
2								

También han sido llevados al mismo cuadro los resultados censales de otras naciones, y de la misma manera calculados promedios y diferencias finitas. Para la rápida percepción se ha construído (fig. 26) un gráfico en el que constan estos resultados. Se aprecia en él, que si bien España tiene una población censal mitad que las otras naciones, su terminación cero presenta una primera diferencia finita de extraordinaria importancia frente a la de las demás. He representado también el fenómeno en otros países para poner de manifiesto que, en mayor o menor cuantía, en todos ellos se produce.

Trabajando con cifras absolutas no se pueden percibir cuantitativamente las diferencias apuntadas, por lo que han sido registradas en el cuadro II las primeras diferencias finitas, máximas y relativas en tanto por mil.

Diferencias máximas, de habitantes varones, entre edades de terminación consecutiva. (Promedios.)

(CUADRO II)

NACION	Δ	
	máxima	en ‰
España	+ 42.948 — 49.307	44,619 51,226
Francia	+ 6.492 — 18.450	3,453 9,814
Italia	+ 6.496 — 21.551	3,378 11,206
Inglaterra	+ 1.368 — 12.946	0,806 7,628

La extraordinaria importancia de las cifras relativas de España frente a las de otras naciones, ya sugiere, sin apelar a otras razones, la necesidad de un estudio más minucioso que el efectuado en otros países para la realización de nuestros trabajos de laboratorio.

Habitantes Varones Clasificados Por La Cifra De Terminacion De La Edad Declarada
[PROMEDIOS]
ESPAÑA
FRANCIA
INGLATERRA
ITALIA
ESCALA EN MILLARES DE HABITANTES
210
200
190
180
170
160
120
110
100
90
80
75
0
3
4
5
6
7
8
9
0
1
2
TERMINACION

FIG. 26

Ajuste de cifras censales.—Tal como se presentan los resultados de un censo de población, después de la tabulación, no permiten su trabajo en operaciones de laboratorio.

Es necesario, por tanto, someterlos a un tratamiento matemático que, fundado en determinados supuestos, haga que estos resultados, a todas luces erróneos, se transformen en otros con más carácter de veracidad.

También se persigue, simultánea o sucesivamente, transformar la línea quebrada que une los extremos de las ordenadas representativas del número de habitantes de cada edad en otra que sea curva y presente la mayor «suavidad» posible.

Aún queda otra finalidad, también conseguida en esta investigación, y esta es la de encontrar la ley o función por la que se rige la distribución de habitantes por su edad.

Los procedimientos que el estadístico dispone para resolver estas cuestiones son los de interpolación, perecuación y gradación o ajuste mecánico. Estudiados conjuntamente en algunos tratados de Estadística, me voy a permitir una separación de los mismos, aclarando el objeto por ellos perseguido, al mismo tiempo que se distinguen los casos o fenómenos en que pueden ser utilizados.

Tratado este tema en mi libro, en preparación, «Métodos de Análisis Estadístico», traslado algunos párrafos que referentes a esta cuestión en él expongo y que de una manera directa han de intervenir en el posterior desarrollo de esta conferencia.

«En todo trabajo estadístico pueden distinguirse cuatro fases:

1.º Planteo, 2.º Observación, 3.º Tabulación y 4.º Análisis.»

. . . .

. . . .

«No obstante la importancia aislada de cada una de las cuatro fases en que descomponemos todo trabajo estadístico, la última de ellas, o análisis, es la de mayor valor, ya que, en definitiva, es la que, a modo de compendio o resumen de todas las anteriores, nos ofrece las leyes que rigen el hecho considerado.»

Valiéndonos de un ejemplo, diremos que la fase de análisis es a la Estadística lo que el levantamiento de un plano es a la labor del topógrafo. Efectuado el trabajo de campo, en el que, por medio de los correspondientes aparatos topográficos, se ha «observado» una parcela determinada de terreno y calculada la situación de los diversos puntos—que en ningún caso podrán ser todos los de la parcela—, se sitúan éstos sobre el papel. A primera vista nada puede ser deducido por un observador de esta serie de puntos trazados, al parecer, de una manera arbitraria, pero si por medio de líneas unimos los de igual cota, los que forman parte de un curso de agua, de una carretera, etc., aquellos primitivos puntos sin ilación habrán cobrado para nosotros un claro significado, que nos permitirá darnos exacta cuenta del relieve y características de la parcela estudiada.»

«Análogamente, en una estadística ya tabulada, el investigador se enfrenta con una serie de valores sin conexión alguna. Mediante la aplicación de los procedimientos que la Matemática nos facilita, podemos obtener las leyes a las que el hecho observado obedece.»

«Del estudio del cuadro o cuadros numéricos que resumen las observaciones de un hecho, el investigador debe deducir, como previa labor estadística, de qué modo el suceso se produce según cambia la cuantía de la variable considerada.»

«La simple inspección de los resultados poco ha de decirle; sólo podrá observar los valores mayores y los menores y si éstos siguen una serie creciente, decreciente u oscilante; pero lo que él busca en realidad, es decir, cómo varía el fenómeno o, mejor, qué ley sigue el mismo, no puede serle sugerida por la simple inspección de los estados resumen.»

«Unas veces con fines divulgadores y otras como previa especulación, se llevan las magnitudes del cuadro numérico estudiado a un gráfico en el que, tomando por abcisas los valores de la variable independiente, se sitúan en las ordenadas los correspondientes de la supuesta función. Este gráfico nos facilita solamente una serie de puntos desperdigados en el papel milimetrado, y no permite siquiera adquirir el concepto más elemental de la investigación: el de cómo en el fenómeno observado se pasa de un valor al siguiente. Es costumbre muy extendida unir estos puntos por rectas y suponer que el suceso varía, de un valor al consecutivo, en la forma reflejada por la recta que une dos puntos próximos. Aunque este procedimiento persigue una finalidad más vulgarizadora que científica, constituye en realidad el primer paso en la *interpolación*.»

«Conociendo sólo cuantías aisladas de un hecho, para poder conocer la ley que rige al mismo, *precisamos determinar todos los valores existentes entre dos consecutivos de la investigación;* y esto sólo es posible conseguirlo por medio de una línea recta o curva que pasando por los valores conocidos procedentes de la observación, nos dé los intermedios. Claro está que por estos dos puntos pasan infinitas curvas, y, por tanto, obtendremos infinitas soluciones del problema. Este no es otra cosa que el de la *interpolación*.»

«Cuando se precisa conocer un valor intermedio en una serie de puntos, se hace pasar por los adyacentes, o bien por todos, una curva, y de su intersección con la ordenada que pasa por el valor correspondiente de la variable se obtiene el punto deseado. Ahora bien; buscar una curva que pase por los valores observados, equivale (exceptuando claro está el método gráfico) a encontrar la expresión matemática de la que cumple estas condiciones, y ésta, que compendia todos los valores que el fenómeno puede tomar entre los límites registrados, es la primera ley que el investigador estadístico ha de obtener en el estudio de un fenómeno.»

«Si tenemos n puntos de observación y los unimos por medio de rectas (fig. 27) habremos de servirnos para el estudio total del fenómeno de $n - 1$ líneas, es decir, tendremos $n - 1$ leyes distintas del mismo, de la forma

$$y = a_0 + a_1 x$$

«Supongamos ahora que en lugar de usar rectas empleáramos parábolas de segundo grado, que tendríamos que hacer pasar, como es lógico, por tres puntos consecutivos; en este caso habría $\frac{n-1}{2}$ leyes diferentes de la forma

$$y = a_0 + a_1 x + a_2 x^2$$

«Del mismo modo, si fueran parábolas de tercer grado las adoptadas, se reduciría el número de leyes a $\frac{n-1}{3}$ de la forma

$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3$$

«Por último, si trabajáramos con parábolas de grado $n - 1$, el número total de parábolas usadas será $\frac{n-1}{n-1}$, es decir, una de la forma

$$y = a_0 + a_1 x + a_2 x^2 + \ldots\ldots + a_{n-1} x^{n-1}$$

«Esta es la que resuelve, en principio, el problema de manera más satisfactoria, pues a la cualidad de ser una sola ley, une la de venir influída no sólo por los dos valores adyacentes, sino por todos, hecho digno de tenerse en cuenta.»

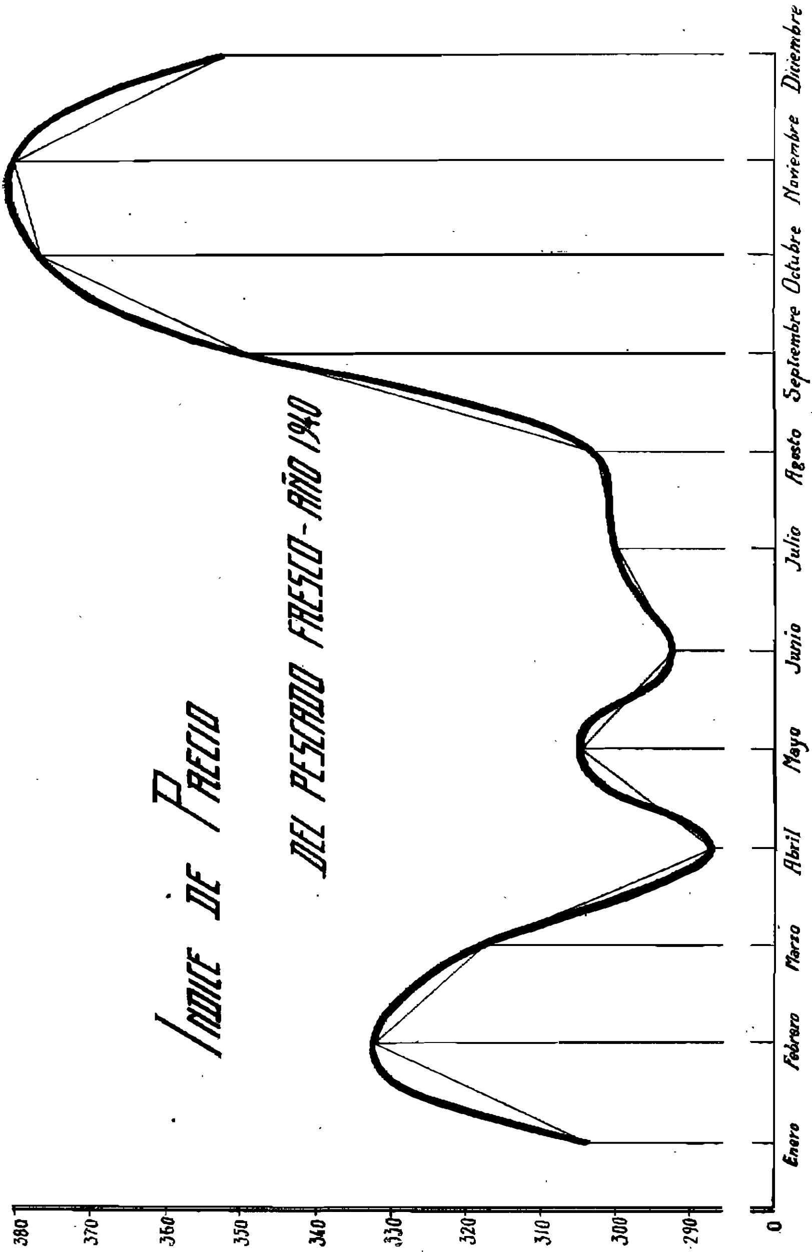

INDICE DE PRECIO
DEL PESCADO FRESCO - AÑO 1940
380
370
360
350
340
330
320
310
300
290
0
Enero
Febrero
Marzo
Abril
Mayo
Junio
Julio
Agosto
Septiembre
Octubre
Noviembre
Diciembre

FIG. 27

«Con el fin de dejar bien sentados los conceptos, añadiremos que a la operación de fijar esta curva se la distingue con el nombre de *interpolación*, y la función así obtenida se denomina *función interpolatriz*, pudiendo ser ésta diferente de la parabólica.»

«Haremos notar que al determinar la curva pasando por las cuantías de observación, implícitamente aceptamos que se pueden calcular y *tener significado todos los valores intermedios* y, además, que utilizando una función continua para obtener la ley del fenómeno, éste *debe ser también función continua de la variable que del mismo se considera*. Circunstancias necesarias que ha de cumplir el hecho en observación para que la *interpolación* le pueda ser aplicada.»

«Hemos visto que la curva producida por la supuesta ley del fenómeno ha de pasar por los valores observados; luego será igualmente condición necesaria que estos valores sean puntos de la misma, es decir, que *se trabaje con valores exactos*, abandonando los métodos de *interpolación* cuando los del hecho observado sean el resultado de medidas u observaciones imprecisas o erróneas, tanto por los medios en que se observa como por los aparatos usados en la observación, ya que, en caso contrario, el resultado obtenido como consecuencia de nuestros trabajos sería erróneo, por serlo los datos básicos utilizados para nuestra labor.»

«El estadístico, ya en posesión de la curva o su expresión matemática, puede determinar los valores máximos o mínimos del hecho, los crecimientos, la velocidad del incremento, las variaciones de esta velocidad, etc., y, en general, cuantos valores el matemático obtiene del estudio analítico de una curva.»

«Si conocida la ley de un suceso en un período de observación tratamos de conocer la importancia que el mismo toma, antes o después del intervalo estudiado, habremos de dar valores inferiores o superiores, en la variable independiente, a los límites del hecho que primeramente se estudió y calcular los que tomaría la función con arreglo a la ley descubierta; operación ésta conocida con el nombre de *extrapolación*. Es preciso advertir la poca consistencia que ofrecen los resultados obtenidos por la *extrapolación*, debiendo ser utilizados éstos solamente a título indiciario.»

«Como consecuencia de todo lo anterior podemos, pues, decir que, en Estadística, se conoce con el nombre de *interpolación* una serie de métodos que, determinando todos los valores intermedios de un fenómeno de ley continua en un período de observaciones exactas, nos proporciona la ley a que obedece el mismo y permite su estudio matemático.»

....
....

«No siempre es posible operar en Estadística con valores exactos; la mayor parte de las veces los resultados obtenidos vienen influídos por errores de observación o por causas accidentales cuyas influencias han de eliminarse en la investigación.»

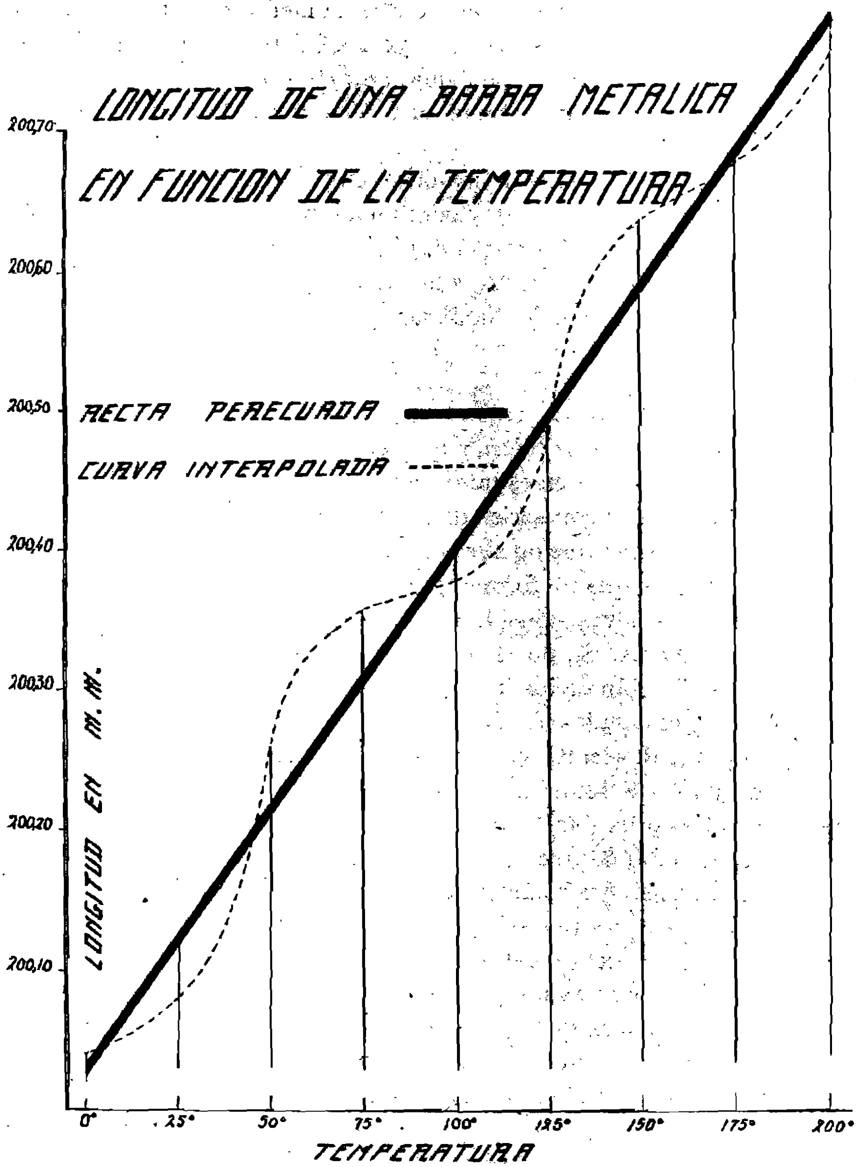
«Bien conocidos son los errores de observación producidos por los aparatos de medida que, en más o menos, afectan a las mismas, extendiendo, por consiguiente, su campo a cuantas observaciones se practican por medio de aparatos. No menos errores engendran las imperfecciones físicas de los sentidos, cuya influencia se hace sentir en los resultados. Ambas causas dan origen a que una serie de observaciones en que intervengan aparatos o apreciaciones sensitivas del operador no pueda considerarse como una serie exacta.»

«Otras series exactas han de considerarse como erróneas a los efectos de obtener su ley. Tal es el caso del crecimiento de la población cuando los años de observación vengan influídos por una guerra o una epidemia; los índices de precios, cuando una causa accidental pudiera afectar a su cuantía; la producción de una fábrica, cuando una avería origine un paro momentáneo, etc.»

«Otra serie de observaciones, que también hemos de considerar como errónea, es la obtenida en un medio inadecuado, tal es la observación de la mortalidad en un pequeño pueblo, en el cual, por la poca importancia numérica de los casos registrados, el fallecimiento de un individuo, de una edad o de otra, puede influir considerablemente en la ley que se obtenga.»

«En estos casos, la ley que mejor expresa el fenómeno, no es la originada por la curva que pasa por los n puntos observados, como ocurría en la interpolación, sino otra distinta que, situada entre los mismos, satisface a ciertas condiciones. Esta curva, que está comprendida entre los valores erróneos (fig. 28), «centrada» entre los mismos, tiende a corregir las medidas u observaciones realizadas con exceso, compensándolas con las verificadas en defecto. Al conjunto de métodos empleados para obtener la ley de un fenómeno con datos erróneos pero de ley continua, se llama *perecuación,* y la curva que la satisface se llama *perecuatriz.*»

«A veces, en los trabajos estadísticos, no es preciso conocer la ley que sigue un fenómeno, sino si éste tiene una «tendencia» a crecer o a



LONGITUD DE UNA BARRA METALICA
EN FUNCION DE LA TEMPERATURA
200,70
200,60
200,50
200,40
200,30
200,20
200,10
RECTA PERECUADA
CURVA INTERPOLADA
LONGITUD EN M. M.
0°
25°
50°
75°
100°
125°
150°
175°
200°
TEMPERATURA

FIG 28

decrecer y determinar la velocidad de este aumento o disminución; en este caso basta *perecuar* una recta, cuyo coeficiente angular mide esta velocidad. A esta *perecuatriz* se llama *línea de tendencia* y puede aplicarse a fenómenos de ley continua con observaciones tanto erróneas como exactas.»

«*Perecuación* es la serie de métodos empleados para determinar la ley que rige un fenómeno de ley continua, cuyos datos de observación se suponen erróneos o inexactos o, también, para obtener la *tendencia* que presenta el mismo, cuando sus datos sean o no exactos; estableciendo ciertas condiciones o hipótesis a que se han de someter los mismos.»

....

....

«Lo mismo en la *interpolación* que en la *perecuación* era condición precisa que el fenómeno en estudio fuera de ley continua, pero, a veces, es necesario trabajar con series discretas, obligando esta circunstancia al investigador a tener que operar con hechos de los que sólo se pueden determinar valores que no forman parte de ninguna ley.»

«Bastan unas leves consideraciones para darse perfecta cuenta de esta clase de valores. Si, por ejemplo, en un censo de la población se estudia la distribución de los habitantes por su edad, expresada en años cumplidos, los resultados forman una serie discreta, ya que los intervalos, comprendidos entre cada dos edades consecutivas, carecen de significado y aun de sentido. Hemos de observar además, en este caso, que las cuantías primitivas sobre las que vamos a operar son erróneas, debido a que, como consecuencia de un fenómeno psicológico, el declarante muestra, en gran número de casos, una marcada simpatía sobre las denominadas cifras redondas; circunstancia que origina el acrecentamiento de éstas a expensas de las restantes. Ahora bien, si las edades consecutivas son agrupadas juiciosamente y se someten a un reparto lógico (fig. 29), los resultados obtenidos—que no pueden asegurarse sean los verdaderos—presentan un carácter de mayor realidad que los primitivos y son, por ello, de mayor utilidad para ser empleados en posteriores trabajos de investigación, ya que el error psicológico, que en un principio fué cometido por el declarante, ha sido subsanado, en su mayor parte, por el investigador.»

«El problema de cómo sería posible operar o discurrir sobre las cuantías primitivas de una serie discreta, afectadas de un error inicial, podemos resolverle mediante el uso de la *gradación*, que no es más que

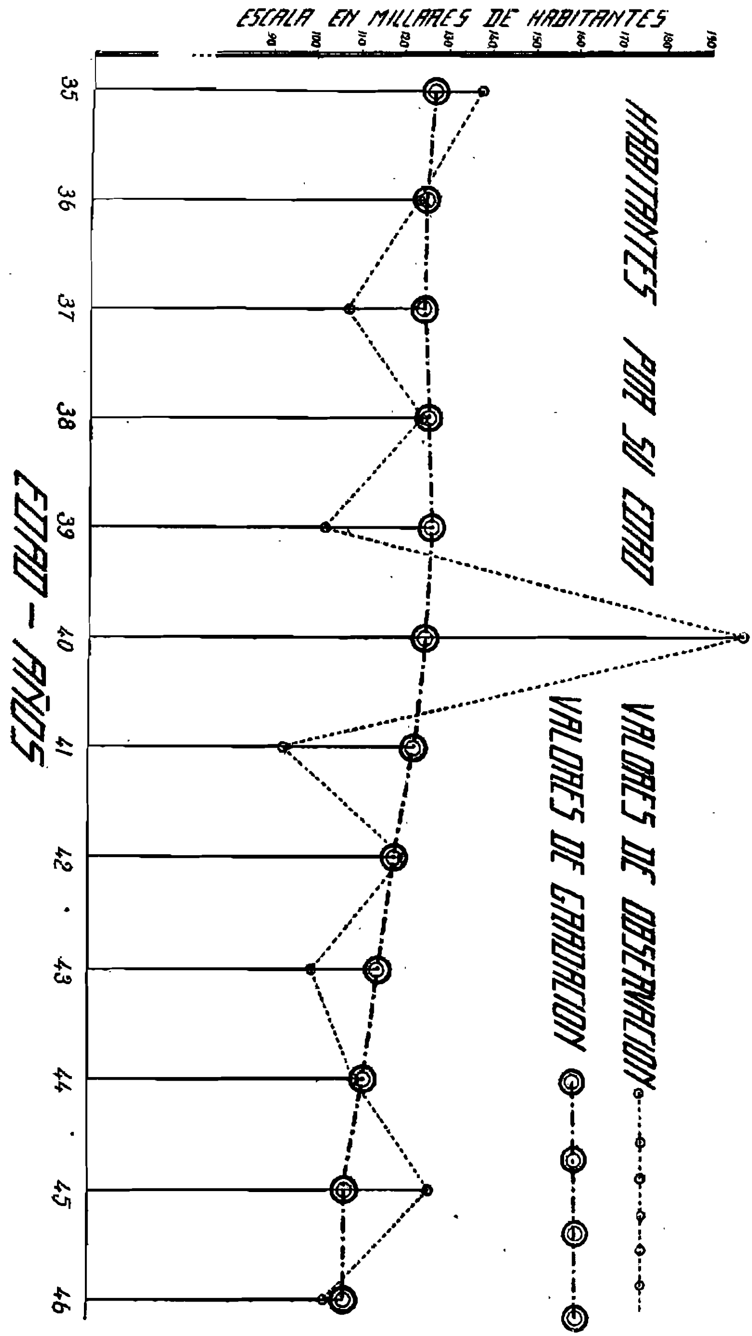

ESCALA EN MILLARES DE HABITANTES
90
100
110
120
130
140
150
160
170
180
190
HABITANTES POR SU EDAD
VALORES DE OBSERVACION
VALORES DE GRADUACION
35
36
37
38
39
40
41
42
43
44
45
46
EDAD - AÑOS

FIG. 29

el conjunto de métodos empleados para sustituir unos valores de observación, erróneos y discontinuos, por otros, también discontinuos, pero que se aceptan como más plausibles.»

....

....

«Existe una clase de series estadísticas, las acumulativas, a las que no es posible aplicar directamente los métodos de interpolación, aunque sus datos sean exactos y expresados por número suficiente para poderlos considerar como formando parte de una ley. En estas series, las cifras de observación vienen dadas por la suma de los hechos o valores que toma el fenómeno, a través de un período de tiempo.»

«De esta clase son las series de nacimientos, defunciones, etc., cuando vienen expresadas por el total de hechos ocurridos en períodos de años, meses, etc. También lo son las de producción, consumo, comercio, etc., cuando, como las anteriores, vienen dadas por años, trimestres o meses.»

«Es costumbre muy extendida considerar estos valores como formando parte de una curva y operar con ellos como tales; obteniendo una ley, una función interpolada o una línea de tendencia, completamente falsa.»

«Aclararemos esto con la exposición de un caso: Sea la serie de las defunciones ocurridas en varios años, clasificadas por el año en que ocurrieron; supongamos que los valores acumulados, por año, han sido llevados a un gráfico por medio de ordenadas distanciadas a intervalos iguales y, por último, interpolada la curva (fig. 30)

$$y = F(x)$$

que, por tratarse de una serie de valores estimados como exactos, pasara por los extremos de aquellas ordenadas.»

«Siendo la función interpolada continua entre los límites de observación, cualquier valor que se dé a la variable independiente—aquí, tiempo—producirá en la función el correspondiente al que tome el fenómeno—número de muertos—. Así, si $F(x_0)$ son las defunciones ocurridas durante el año 1940, $x_0 \rightarrow x_1$ el intervalo adoptado para representar un año—1941—, y $F(x_1)$ las defunciones registradas desde primeros a fines del año 1941, un valor intermedio tal como *m n*, situado en el punto medio del intervalo $x_0 \rightarrow x_1$ representará las defunciones ocurridas durante los seis primeros meses del año 1941. En casos como el de la figura, en que por ser

$$F(x_0) > F(x_1)$$

la función es decreciente, resultará que

$$m\,n > F(x_1)$$

es decir, que las defunciones ocurridas en los seis primeros meses del

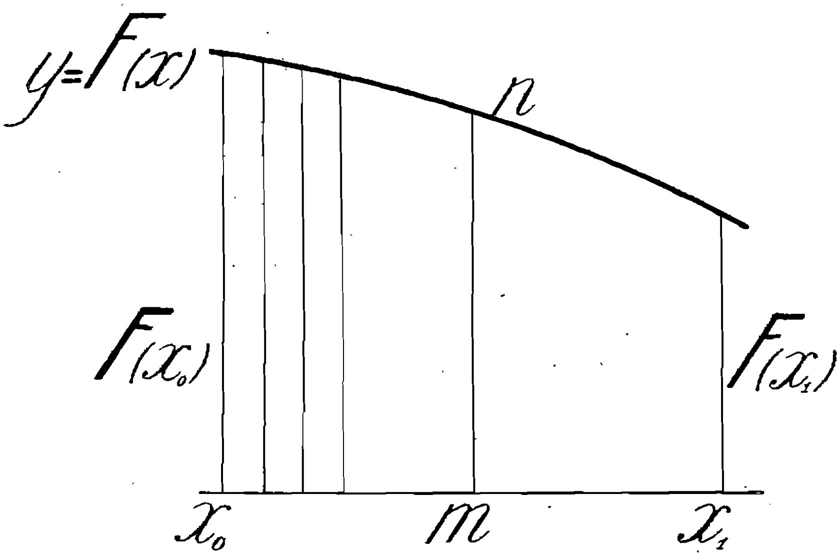


FIG. 30

año 1941 son en mayor número que las ocurridas en todo el mismo año. Resaltará más el anterior absurdo si se determinan valores de la función por meses, cuya suma será, en mucho, superior a $F(x_1)$.»

«A todo esto, inmediatamente se argumentará que no es lícito pensar que los valores intermedios de la función, entre x_0 y x_1, puedan expresar estados intermedios del fenómeno, y esto es lo que únicamente se quería poner de relieve, pues si la función carece de significado en la mayor parte de sus puntos, no existe tal función, quedando reducida la pretendida interpolación a una mera gradación en la que, por no tener significado más que en las ordenadas que pasan por x_0, x_1, se ha obtenido, como final de los cálculos, los mismos valores—datos de observación—que ya se conocían.»

«Se acostumbra también a tratar estas series situando el valor de observación durante un período $x_0 \rightarrow x_1$, en el punto medio n de dicho intervalo, e interpolar después. Este proceder adolece del mismo inconveniente que el anterior, pues siguen careciendo de significado los valores intermedios de la función. Cabe, en esta forma de presentar el fenómeno, considerar que estas ordenadas centrales $m\,n$ son las alturas (fig. 31) de los rectángulos que, construídos sobre los intervalos, repre-

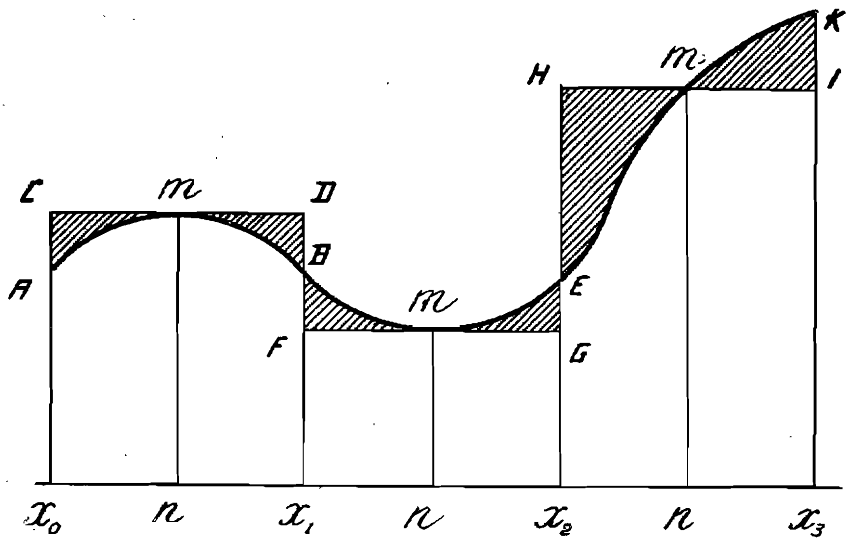


FIG. 31

sentan los valores del fenómeno cuando el intervalo es la unidad, y considerar la curva interpolada como límite de áreas equivalentes a las de observación. En casos como el presentado en el intervalo $x_2 \rightarrow x_3$, el área E H m que se resta del área de observación para obtener la interpolada, se acepta como compensando a la que se suma: m K I; pero esto, aunque muy discutible en este caso, es completamente inadmisible en los otros intervalos en que las áreas A C m + m D B y la B F m + m G E no tienen compensación dentro de cada intervalo, por lo que no es posible aceptar esta última interpretación en la representación de las series acumulativas.»

«Los mismos razonamientos pueden ser empleados en el caso de la perecuación, de uso más extendido, especialmente, en líneas de tendencia.»

«Si bien estas series, ya ajustadas, pudieran ser objeto de correcciones posteriores, como éstas no formarían ley, sino sólo serían incrementos calculados en función de valores próximos, entraríase de lleno en la gradación.»

«Sin embargo, los hechos demográficos no carecen de ley en función del tiempo; lo que ocurre es que anteriormente no fueron correctamente expresados.»

«Dos representaciones pueden ser utilizadas: Una, la acumulada (figura 32), en que las sucesivas ordenadas son suma de los hechos ocurridos desde un origen de observación hasta la fecha indicada por la abcisa. Aquí cualquier valor intermedio tiene un claro y distinto significado. Otra forma es por áreas. Se determina la curva $y = F(x)$ de tal manera (fig. 33), que la integral, con límites dados por los extremos del intervalo, equivalga al valor de observación entre los mismos límites. La integral de límites intermedios entre los de observación, da también valores intermedios del fenómeno, así como usando límites fuera de lo observado dará valores de extrapolación.»

«Tanto una como otra representación permiten lo mismo la interpolación que la perecuación.»

Los métodos usados en el ajuste de la población han sido casi exclusivamente de gradación, variando, en cuanto a dificultad de cálculo, desde el de Finlaison, que sólo precisa de cinco edades para el ajuste de una, hasta el de King, que utiliza para el mismo objeto cuarenta y tres. En consecuencia, no se puede ajustar antes de los dos años por el primero ni antes de los veintidós por el segundo.

Pueden también usarse los métodos de perecuación, pero todos ellos, salvo el de Cantelli, no producen más que valores discretos, viniendo a dar los mismos resultados que la gradación, con la única diferencia de que la linea que une los extremos de las ordenadas representativas de los habitantes es continua, aunque carece de significado entre ordenadas.

El método de Cantelli, con carácter funcional, presenta grandes dificultades de cálculo, pues de emplearse clases quinquenales es preciso resolver un sistema de 20 ecuaciones con parámetros de 40 o más cifras.

Unicamente Gini logra una perecuatriz viable valiéndose de trazos de parábola que comprenden cinco edades, precisando, por tanto, de 17 funciones distintas para el ajuste de edades comprendidas entre 10 y 95 años.

Errores que los métodos clásicos de ajuste introducen en su apli-

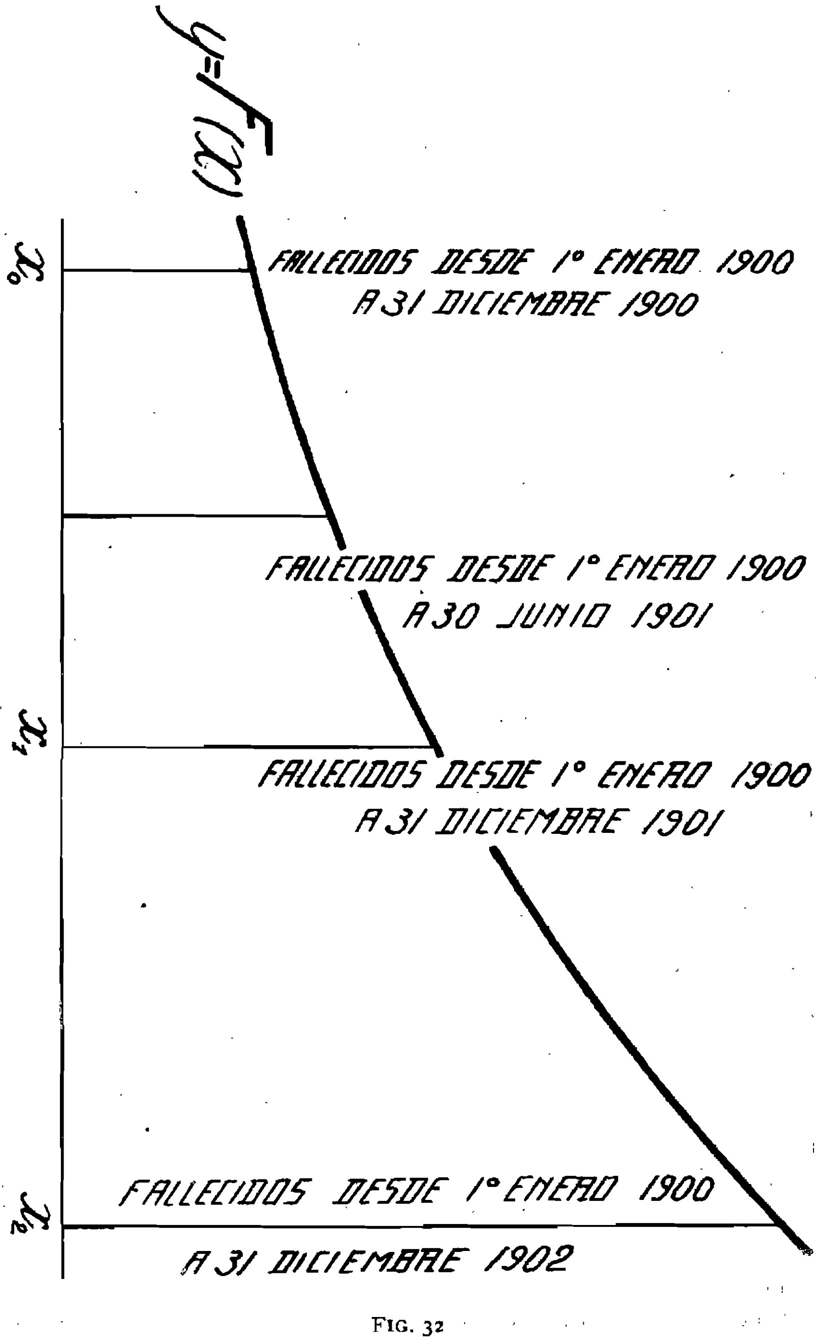

y=f(x)
x0
x1
x2
FALLECIDOS DESDE 1° ENERO 1900
A 31 DICIEMBRE 1900
FALLECIDOS DESDE 1° ENERO 1900
A 30 JUNIO 1901
FALLECIDOS DESDE 1° ENERO 1900
A 31 DICIEMBRE 1901
FALLECIDOS DESDE 1° ENERO 1900
A 31 DICIEMBRE 1902

FIG. 32

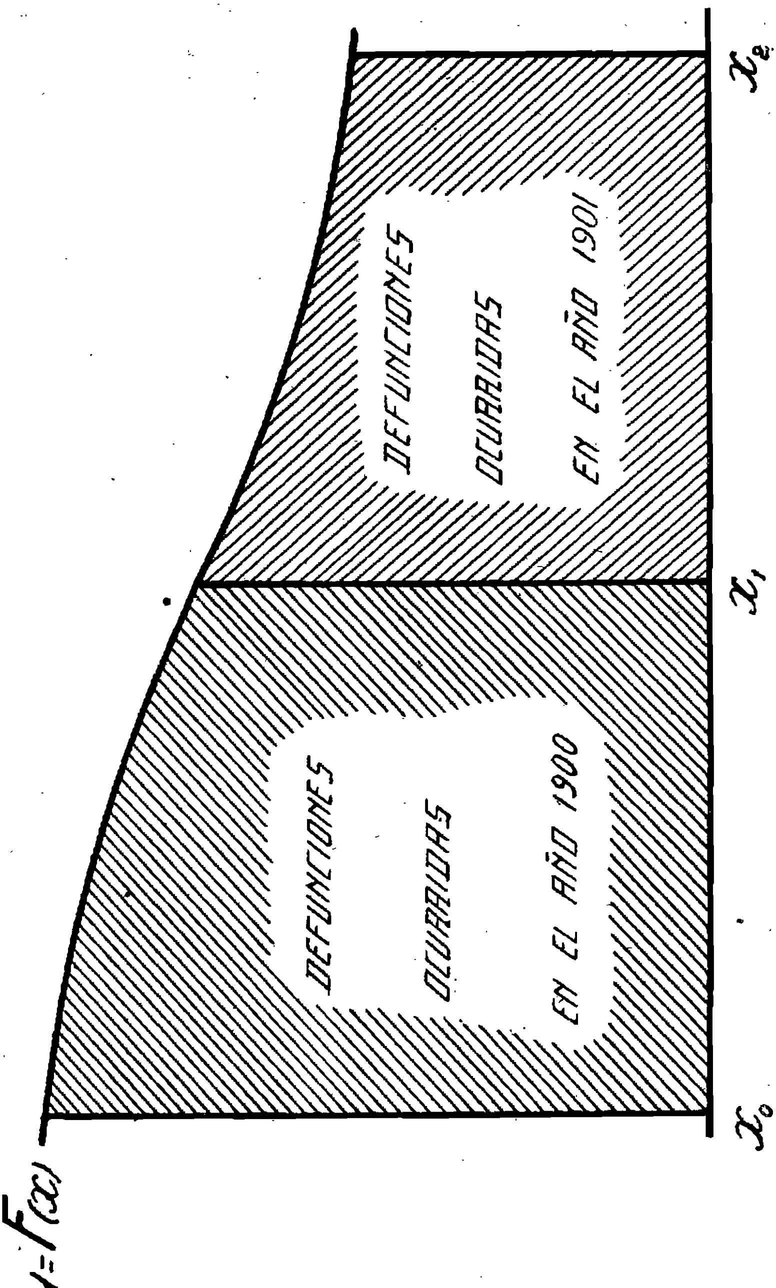

y=F(x)
DEFUNCIONES
OCURRIDAS
EN EL AÑO 1900
DEFUNCIONES
OCURRIDAS
EN EL AÑO 1901
x_0
x_1
x_2

FIG. 33

CACIÓN A LA DEMOGRAFÍA ESPAÑOLA.—En la publicación que con el título *La Demografía Española* hizo el Laboratorio de la Dirección General de Estadística, puede verse, en la página 54, un estado de los varones ajustados por varios métodos, estado que reproducimos en el cuadro III.

Las cifras consignadas en este cuadro presentan anomalías consistentes en que algunas edades, después de ajustadas, tienen mayor número de habitantes que otras que les preceden. Aunque esto puede ser posible, no es muy lógico que, siendo la población de una nación decreciente desde cero años de edad al límite humano de vida, se presente así después de ajustada. Las cifras a que me he referido son las subrayadas en el cuadro III. La duda que se me presentó al estudiar esta serie, entre lo lícito o no de aceptar este ajuste, fué la causa de toda la investigación que ante ustedes expongo.

En la misma publicación, y en la página 57, puede verse un estadillo en el que se suman las desviaciones por grupos de edad, y que he reproducido en el cuadro IV. En él se aprecia que, a partir del grupo 28-37 años, las desviaciones totales son negativas, al paso que en los grupos anteriores, son positivas. Esto parece, como si los métodos de ajuste trasladaran una parte de los habitantes de más de 37 años a los de menos de esa edad.

Dándose los valores que creo anómalos en las proximidades de las edades terminadas en cero o cinco, llevé los valores ajustados a un gráfico (fig. 34) en el que se aprecia claramente una ondulación en la línea que pasa por los extremos de la población ajustada. Ondulación que presenta sus máximos cerca de los ceros y mínimos en las proximidades de los cincos, y que, aunque el gráfico sólo comprende las edades 33 a 52 años, lo mismo se produce en las sucesivas. A esta ondulación es debida la irregularidad de cifras que hicimos notar.

Recordando el cuadro II pensé que esta ondulación fuera debida a la gran oscilación que entre edades próximas al cero o cinco presenta el Censo español y, aún más, a la diferencia entre cifras correspondientes a ceros y cincos consecutivos que también presenta el mismo.

Antes de pasar adelante, voy a tratar del ajuste de una recta en los valores de observación; recta que, tanto por sí, como por los segmentos que ella determina, ha de gozar un papel importante en el sucesivo desarrollo de esta conferencia.

RECTA AJUSTADA.—Para el cálculo de la recta indicada supongamos (fig. 35) que el segmento $3 - 3_1$ representa el promedio de habitantes de

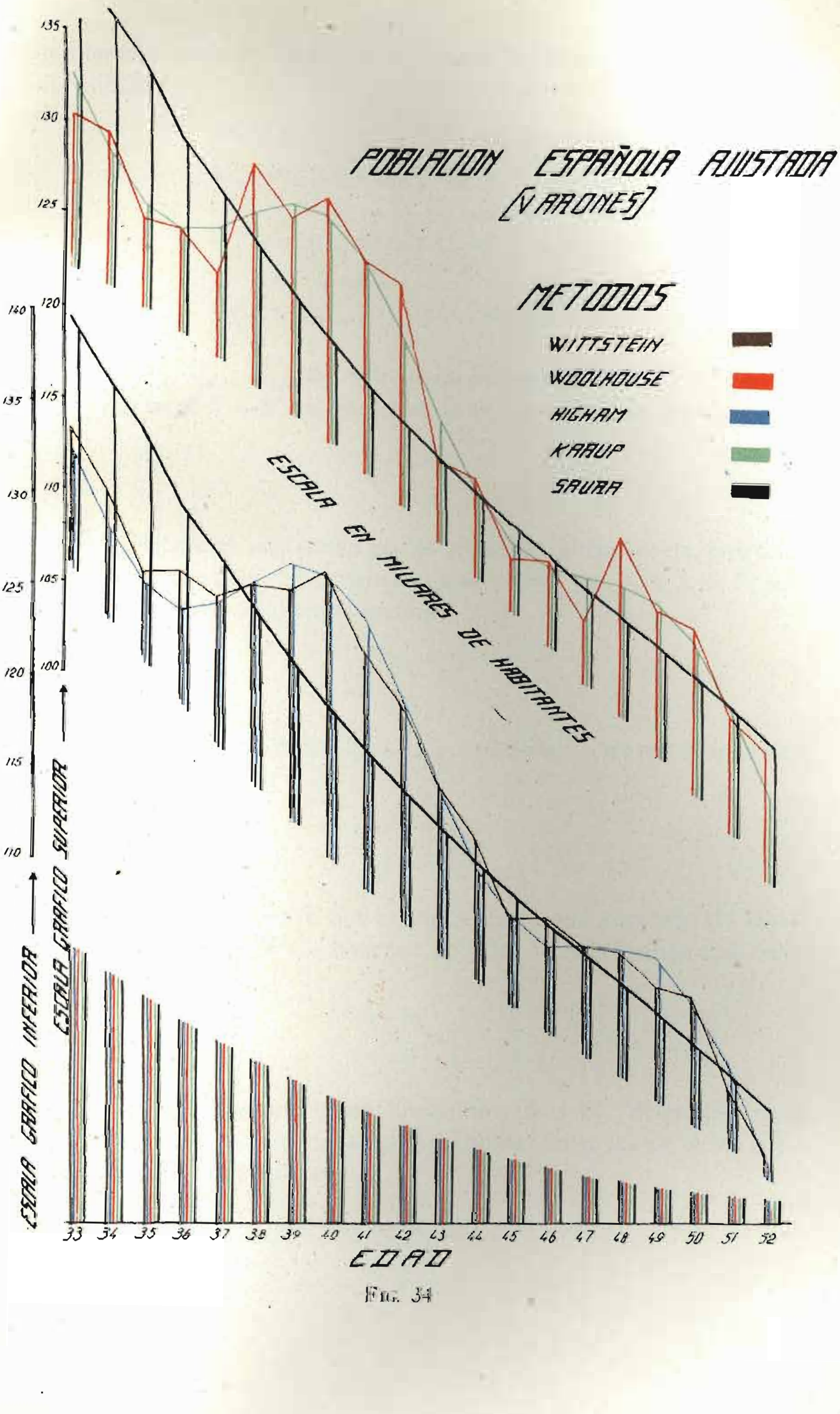

POBLACION ESPAÑOLA AJUSTADA
(VARONES)
METODOS
WITTSTEIN
WOOLHOUSE
HIGHAM
KARUP
SAURA
ESCALA EN MILLARES DE HABITANTES
ESCALA GRAFICO INFERIOR
ESCALA GRAFICO SUPERIOR
135
130
125
120
115
110
105
100
140
135
130
125
120
115
110
33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52
EDAD

FIG. 34

edad terminada en 3; desde 3 a 93 años y, lo mismo, los demás y cuya expresión es:

$$U_3^P = {}^1/_{10}\,[U_3 + U_{13} + U_{23} + \ldots\ldots + U_{93}]$$
$$U_4^P = {}^1/_{10}\,[U_4 + U_{14} + U_{24} + \ldots\ldots + U_{94}]$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$U_{13}^P = {}^1/_{10}\,[U_{13} + U_{23} + U_{33} + \ldots\ldots + U_{103}]$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$U_{22}^P = {}^1/_{10}\,[U_{22} + U_{32} + U_{42} + \ldots\ldots + U_{113}]$$

Siendo U_3, U_4 ... U_{118} los valores de observación.

El segmento 5 — 5′ es determinado por la condición de ser igual a

$$^1/_5 \sum_{x=3}^{x=7} U_x^P$$

Todas las rectas que pasan por el punto 5′, gozan de la propiedad de que limitan sobre las ordenadas que pasan por 3, 4, 5, 6, 7; segmentos cuya suma es siempre igual a

$$\sum_{x=3}^{x=7} U_x^P$$

El punto 10′ es fijado por análoga condición; la de ser el segmento 10—10′ igual a

$$^1/_5 \sum_{x=8}^{x=12} U_x^P$$

y, por tanto, los segmentos que cualquier recta que pase por 10′ determina en las ordenadas trazadas por 8, 9, 10, 11, 12 forman una suma siempre igual a

$$\sum_{x=8}^{x=12} U_x^P$$

La recta C D, pasando por los puntos 5′ y 10′, determina unos segmentos sobre las ordenadas comprendidas entre las elevadas por 3 y 12, cuya suma es siempre

$$\sum_{x=3}^{x=12} U_x^P$$

Resumen de los resultados obtenidos ajustando, por diversos métodos, la población por edades de 9 a 91 años dada por el Censo de 1920

Varones

(CUADRO III)

Edad	Valores observados	WITTSTEIN		WOOLHOUSE		HIGHAM		KARUP	
x	u_x	u'_x	$u'_x - u_x$	u'_x	$u'_x - u_x$	u'_x	$u'_x - u_x$	u'_x	$u'_x - u_x$
9	226 583	234 912	+ 8 329	236 144	+ 9 561	236 831	+ 10 248	236 623	+ 10 040
10	249 530	234 259	— 15 271	236 486	— 13 044	235 651	— 13 879	235 594	— 13 936
11	217 970	231 488	+ 13 518	232 698	+ 14 728	233 580	+ 15 610	233 493	+ 15 523
12	242 827	228 957	— 13 870	231 296	— 11 531	230 353	— 12 474	230 401	— 12 426
13	218 645	224 501	+ 5 856	225 693	+ 7 048	226 433	+ 7 788	226 380	+ 7 735
14	227 820	220 769	— 7 051	222 233	— 5 587	221 613	— 6 207	221 639	— 6 181
15	213 355	214 615	+ 1 260	215 407	+ 2 052	216 213	+ 2 858	216 024	+ 2 669
16	206 236	209 132	+ 2 896	211 003	+ 4 767	209 742	+ 3 506	209 621	+ 3 385
17	204 631	202 204	— 2 427	200 731	— 3 900	202 724	— 1 907	202 538	— 2 093
18	210 206	195 811	— 14 395	195 702	— 14 504	195 085	— 15 121	195 162	— 15 044
19	168 507	187 917	+ 19 410	187 238	+ 18 731	187 788	+ 19 281	187 884	+ 19 377
20	193 978	182 254	— 11 724	181 739	— 12 239	180 973	— 13 005	181 365	— 12 613
21	158 350	176 374	+ 18 024	174 295	+ 15 945	175 512	+ 17 162	175 716	+ 17 366
22	179 818	172 236	— 7 582	171 997	— 7 821	170 888	— 8 930	171 070	— 8 748
23	164 256	167 809	+ 3 553	166 388	+ 2 132	167 323	+ 3 067	167 218	+ 2 962
24	164 946	164 954	+ 8	164 566	— 380	163 882	— 1 064	163 841	— 1 105
25	169 196	160 506	— 8 690	160 063	— 9 133	160 666	— 8 530	160 472	— 8 724
26	154 012	157 939	+ 3 927	158 215	+ 4 203	157 214	+ 3 202	157 315	+ 3 303
27	147 769	154 006	+ 6 237	152 679	+ 4 910	154 223	+ 6 454	154 232	+ 6 463
28	157 611	151 211	— 6 400	152 677	— 4 934	151 088	— 6 523	151 191	— 6 420
29	121 549	147 536	+ 25 987	147 320	+ 25 771	148 260	+ 26 711	148 142	+ 26 593
30	191 452	145 689	— 45 763	145 548	— 45 904	145 061	— 46 391	144 996	— 46 456
31	115 527	140 787	+ 25 260	140 871	+ 25 344	141 617	+ 26 090	141 180	+ 25 653
32	144 175	137 731	— 6 444	139 033	— 5 142	137 134	— 7 041	136 948	— 7 227
33	128 955	133 341	+ 4 386	130 424	+ 1 469	132 631	+ 3 676	132 576	+ 3 621
34	123 843	129 909	+ 6 066	129 427	+ 5 584	128 207	+ 4 364	128 574	+ 4 731

36	123 167	125 638	+ 2 471	124 114	+ 947	123 514	+ 347	124 215	+ 1 048
37	107 239	124 239	+ 17 000	121 744	+ 14 505	123 940	+ 16 701	124 220	+ 16 981
38	124 256	124 809	+ 553	127 627	+ 3 371	125 030	+ 774	124 971	+ 715
39	101 514	124 548	+ 23 034	124 809	+ 23 295	126 082	+ 24 568	125 497	+ 23 983
40	197 324	125 467	— 71 857	125 849	— 71 475	125 373	— 71 951	124 914	— 72 410
41	92 190	121 135	+ 28 945	122 410	+ 30 220	123 175	+ 30 985	122 466	+ 30 276
42	119 525	118 454	— 1 071	121 246	+ 1 721	118 886	— 639	118 718	— 807
43	98 794	114 079	+ 15 285	111 592	+ 12 798	114 105	+ 15 311	114 304	+ 15 510
44	108 653	111 029	+ 2 376	110 707	+ 2 054	109 581	+ 928	110 217	+ 1 564
45	126 050	106 593	— 19 457	106 285	— 19 765	106 855	— 19 195	107 198	— 18 852
46	102 464	106 705	+ 4 241	106 153	+ 3 689	105 222	+ 2 758	105 780	+ 3 316
47	89 479	105 131	+ 15 652	102 931	+ 13 452	105 180	+ 15 701	105 228	+ 15 749
48	106 464	104 794	— 1 670	107 484	+ 1 020	105 117	— 1 347	104 875	— 1 589
49	81 109	102 907	+ 21 798	103 511	+ 22 402	104 516	+ 23 407	103 897	+ 22 788
50	168 497	102 295	— 66 202	102 564	— 65 933	102 097	— 66 400	101 779	— 66 718
51	70 600	97 211	+ 26 611	97 726	+ 27 126	98 592	+ 27 992	98 116	+ 27 516
52	90 347	93 899	+ 3 552	95 892	+ 5 545	93 718	+ 3 371	93 726	+ 3 379
53	75 372	89 657	+ 14 285	86 735	+ 11 363	89 067	+ 13 695	89 342	+ 13 970
54	85 457	86 882	+ 1 425	86 043	+ 586	85 124	— 333	85 742	+ 285
55	99 124	83 098	— 16 026	82 750	— 16 374	83 097	— 16 027	83 433	— 15 691
56	83 053	83 704	+ 651	82 974	— 79	82 285	— 768	82 801	— 252
57	67 969	82 734	+ 14 765	80 884	+ 12 915	82 944	+ 14 975	83 006	+ 15 037
58	80 940	82 599	+ 1 659	85 757	+ 4 817	83 497	+ 2 557	83 238	+ 2 298
59	64 596	81 205	+ 16 609	82 450	+ 17 854	83 252	+ 18 656	82 624	+ 18 028
60	139 478	80 395	59 083	81 068	— 58 410	80 908	— 58 570	80 512	— 58 966
61	56 122	74 952	+ 18 830	76 455	+ 20 333	76 967	+ 20 845	76 464	+ 20 342
62	66 481	70 527	+ 4 046	73 139	+ 6 658	71 279	+ 4 798	71 214	+ 4 733
63	55 019	65 172	+ 10 153	63 267	+ 8 248	65 216	+ 10 197	65 454	+ 10 435
64	58 329	60 437	+ 2 108	60 069	+ 1 740	59 432	+ 1 103	59 946	+ 1 617
65	64 000	54 890	— 9 110	54 980	— 9 020	55 046	— 8 954	55 290	— 8 710
66	54 647	52 661	— 1 986	51 902	— 2 745	51 506	— 3 141	51 823	— 2 824
67	41 751	49 350	+ 7 599	47 539	+ 5 788	48 958	+ 7 207	48 979	+ 7 228
68	44 608	46 533	+ 1 925	47 981	+ 3 373	46 542	+ 1 934	46 369	+ 1 761

Edad	Valores observados	WITTSTEIN		WOOLHOUSE		HIGHAM		KARUP	
x	u_x	u'_x	$u'_x - u_x$	u'_x	$u'_x - u_x$	u'_x	$u'_x - u_x$	u'_x	$u'_x - u_x$
69	31 739	43 253	+ 11 514	43 481	+ 11 742	43 908	+ 12 169	43 559	+ 11 820
70	71 489	40 703	− 30 786	40 378	− 31 111	40 348	− 31 141	40 220	− 31 269
71	26 115	36 006	+ 9 891	36 087	+ 9 972	36 409	+ 10 294	36 215	+ 10 100
72	29 361	32 175	+ 2 814	33 001	+ 3 640	32 018	+ 2 657	32 023	+ 2 662
73	22 623	28 285	+ 5 662	26 840	+ 4 217	27 800	+ 5 177	27 949	+ 5 326
74	23 416	25 001	+ 1 585	24 269	+ 853	24 038	+ 622	24 315	+ 899
75	25 290	21 506	− 3 784	21 235	− 4 055	21 237	− 4 053	21 350	− 3 940
76	19 246	19 733	+ 487	19 225	− 21	19 002	− 244	19 109	− 137
77	15 471	17 686	+ 2 215	16 652	+ 1 181	17 262	+ 1 791	17 219	+ 1 748
78	15 800	15 808	+ 8	16 144	+ 344	15 586	− 214	15 452	− 348
79	9 827	13 726	+ 3 899	13 693	+ 3 866	13 801	+ 3 974	13 612	+ 3 785
80	22 132	11 963	− 10 169	11 646	− 10 486	11 637	− 10 495	11 593	− 10 539
81	5 550	9 588	+ 4 038	9 313	+ 3 763	9 445	+ 3 895	9 413	+ 3 863
82	5 770	7 644	+ 1 874	7 625	+ 1 855	7 299	+ 1 529	7 343	+ 1 573
83	3 891	5 931	+ 2 040	5 141	+ 1 250	5 453	+ 1 562	5 537	+ 1 646
84	4 224	4 612	+ 388	4 029	− 195	3 988	− 236	4 097	− 127
85	3 681	3 387	− 294	3 041	− 640	3 026	− 655	3 066	− 615
86	2 598	2 832	+ 234	2 453	− 145	2 392	− 206	2 406	− 192
87	1 873	2 308	+ 435	1 824	− 49	1 979	+ 106	1 953	+ 80
88	1 808	1 883	+ 75	1 784	− 24	1 659	− 149	1 618	− 190
89	982	1 482	+ 500	1 366	+ 384	1 382	+ 400	1 335	+ 353
90	1 971	1 180	− 791	1 070	− 901	1 059	− 912	1 059	− 912
91	431	880	+ 449	770	+ 339	789	+ 358	787	+ 356

Población masculina de los grupos de edad que se mencionan, dada por la observación y por los métodos que se indican

(CUADRO IV)

GRUPOS DE EDAD	Población observada	WITTSTEIN		WOOLHOUSE		HIGHAM		KARUP	
		Población calculada	Diferencia entre la población calculada y observada	Población calculada	Diferencia entre la población calculada y observada	Población calculada	Diferencia entre la población calculada y observada	Población calculada	Diferencia entre la población calculada y observada
De 9 a 17 años...	2 007 597	2 000 837	— 6 760	2 011 691	+ 4 094	2 013 140	+ 5 543	2 012 313	+ 4 716
— 18 a 27 — ...	1 711 038	1 719 806	+ 8 768	1 712 882	+ 1 844	1 713 554	+ 2 516	1 714 275	+ 3 237
— 28 a 37 — ...	1 350 873	1 361 573	+ 10 700	1 355 906	+ 5 033	1 356 669	+ 5 796	1 357 567	+ 6 694
— 38 a 47 — ...	1 160 249	1 157 950	— 2 299	1 159 609	— 640	1 159 489	— 760	1 159 293	— 956
— 48 a 57 — ...	927 992	927 181	— 811	926 563	— 1 429	926 557	— 1 435	926 717	— 1 275
— 58 a 67 — ...	681 363	672 188	— 9 175	676 626	— 4 737	676 061	— 5 302	675 544	— 5 819
— 68 a 77 — ...	309 358	310 881	+ 1 523	309 149	— 209	308 564	— 794	308 328	— 1 030
— 78 a 87 — ...	75 346	77 799	+ 2 453	74 909	— 437	74 606	— 740	74 472	— 874
— 88 a 91 — ...	5 192	5 425	+ 233	4 990	— 202	4 889	— 303	4 799	— 393
De 9 a 91 años...	8 229 008	8 233 640	+ 4 632	8 232 325	+ 3 317	8 233 529	+ 4 521	8 233 308	+ 4 300

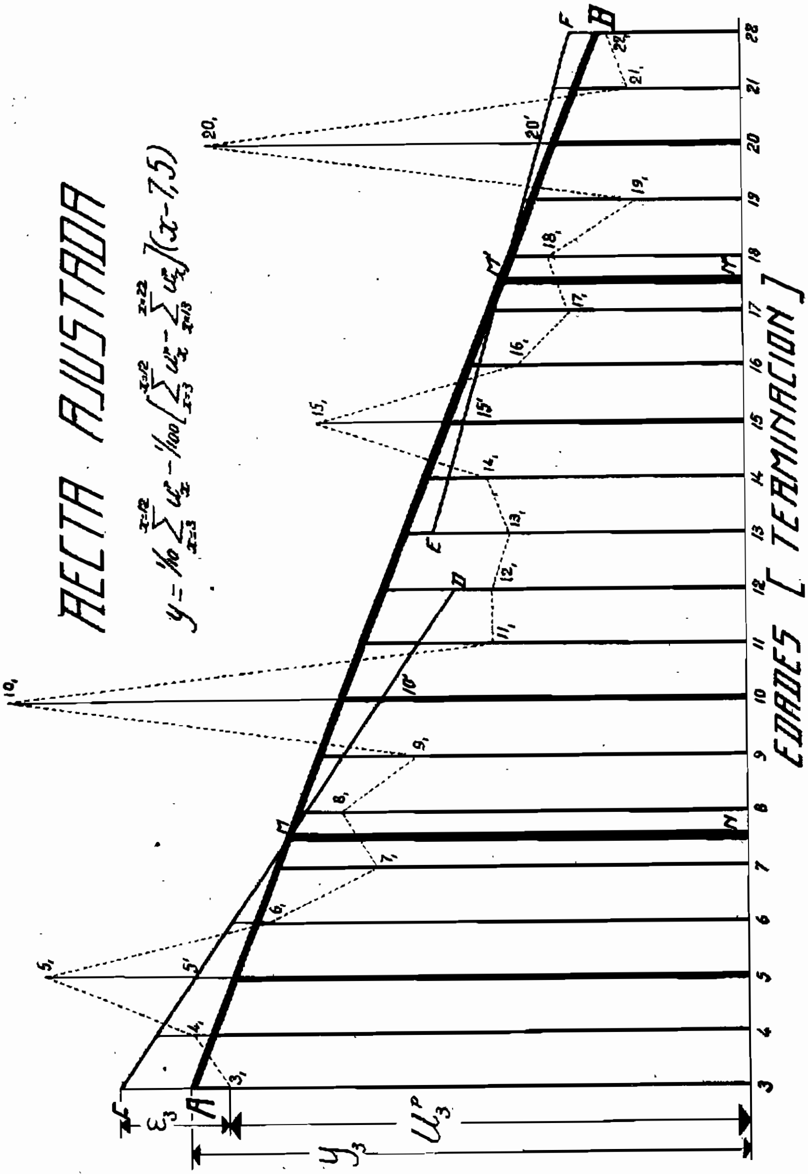

RECTA AJUSTADA
$y = \frac{1}{10}\sum_{x=3}^{x=12} u_x^p - \frac{1}{100}\left[\sum_{x=3}^{x=12} u_x^p - \sum_{x=13}^{x=22} u_x^p\right](x-7,5)$
EDADES [TERMINACION]

FIG. 35

Ocurre lo mismo con la recta E F, en la segunda parte de la figura.

La recta C D limita sobre la ordenada que pasa por el punto central 7,5 un valor igual a

$$\tfrac{1}{10} \sum_{x=3}^{x=12} U_x^P$$

siendo propiedad del punto M determinar cualquier recta que por él pase, segmentos en las ordenadas comprendidas entre 3 y 12, cuya suma es siempre

$$\sum_{x=3}^{x=12} U_x^P$$

Lo mismo ocurre con las rectas que pasan por el punto M′ y la suma invariable es aquí

$$\sum_{x=13}^{x=22} U_x^P$$

Si, por último, unimos M con M′ tendremos la recta A B que, manteniendo la constancia de suma en los dos grupos de diez ordenadas que comprende, lo verifica, por tanto, con todos los de observación.

Como M N vale

$$\tfrac{1}{10} \sum_{x=3}^{x=22} U_x^P$$

y el coeficiente angular de la recta M M′ es

$$\tfrac{1}{100} \left[\sum_{x=3}^{x=12} U_x^P - \sum_{x=13}^{x=22} U_x^P \right]$$

la ecuación de la misma, será

$$y = \tfrac{1}{10} \sum_{x=3}^{x=12} U_x^P - \tfrac{1}{100} \left[\sum_{x=3}^{x=12} U_x^P - \sum_{x=13}^{x=22} U_x^P \right] (x - 7{,}5)$$

Ondulación de ajuste.—Continuando la marcha seguida en la investigación, nos encontramos ante la creencia de que los métodos de gradación producen, en el censo español, una ondulación en la línea que une los extremos de las ordenadas representativas de las edades ajustadas. Para poner de manifiesto esta ondulación he trabajado, no sobre cifras aisladas, sino sobre promedios de todo el censo, agrupando previamente las edades por su terminación, y refiriendo los valores ajustados de estos promedios a la recta que anteriormente calculamos, y que por tener una inclinación promedia entre todas las que presentan cada dos grupos de diez edades consecutivas, descarta el enmascaramiento de la ondulación producida por el sucesivo descenso del valor en las ordenadas representativas de edades sucesivas.

En vez de sumar valores ya ajustados y obtener luego promedios, he preferido ajustar valores de observación, después de referirlos a la recta a que ya hicimos alusión.

En la figura 35 se ve que todo valor promedio de observación

$$U_3^P = y_3 + \varepsilon_3$$
$$U_4^P = y_4 + \varepsilon_4$$
$$\cdots\cdots\cdots\cdots$$
$$U_{22}^P = y_{22} + \varepsilon_{22}$$

Siendo $y_3, y_4 \ldots y_{22}$ las ordenadas de la recta ajustada y $\varepsilon_3, \varepsilon_4 \ldots \varepsilon_{22}$ las diferencias entre los dos valores: el ajustado de la recta y el promedio de observación.

Aplicando la fórmula de ajuste mecánico del método de Wittstein a estos valores, tendremos:

$$V_x^P = 0{,}2 \cdot U_x^P + 0{,}16\left(U_{x-1}^P + U_{x+1}^P\right) + 0{,}12\left(U_{x-2}^P + U_{x+2}^P\right) +$$
$$0{,}08\left(U_{x-3}^P + U_{x+3}^P\right) + 0{,}04\left(U_{x-4}^P + U_{x+4}^P\right)$$

o también:

$$V_x^P = 0{,}2\left[y_x + \varepsilon_x\right] + 0{,}16\left[(y_{x-1} + \varepsilon_{x-1}) + (y_{x+1} + \varepsilon_{x+1})\right] +$$
$$0{,}12\left[(y_{x-2} + \varepsilon_{x-2}) + (y_{x+2} + \varepsilon_{x+2})\right] + 0{,}08\left[(y_{x-3} + \varepsilon_{x-3}) + (y_{x+3} + \varepsilon_{x+3})\right] +$$
$$0{,}04\left[(y_{x-4} + \varepsilon_{x-4}) + (y_{x+4} + \varepsilon_{x+4})\right]$$

en que, agrupando las ordenadas de la recta ajustada

$$V_x^P = 0{,}2\, y_x + 0{,}16\,(y_{x-1} + y_{x+1}) + 0{,}12\,(y_{x-2} + y_{x+2}) +$$
$$0{,}08\,(y_{x-3} + y_{x+3}) + 0{,}04\,(y_{x-4} + y_{x+4}) + 0{,}2 \cdot \varepsilon_x +$$
$$0{,}16\,(\varepsilon_{x-1} + \varepsilon_{x+1}) + 0{,}12\,(\varepsilon_{x-2} + \varepsilon_{x+2}) + 0{,}08\,(\varepsilon_{x-3} + \varepsilon_{x+3}) +$$
$$0{,}04\,(\varepsilon_{x-4} + \varepsilon_{x+4})$$

Para toda ordenada y_x limitada por la recta AB, se verifica:

$$2\,y_x = y_{x-1} + y_{x+1} = y_{x-2} + y_{x+2} = y_{x-3} + y_{x+3} = y_{x-4} + y_{x+4}$$

con lo que, finalmente, se tendrá

$$V_x^P = y_x + 0{,}2\,\varepsilon_x + 0{,}16\,(\varepsilon_{x-1} + \varepsilon_{x+1}) +$$

$$0{,}14\,(\varepsilon_{x-2} + \varepsilon_{x+2}) + 0{,}08\,(\varepsilon_{x-3} + \varepsilon_{x+3}) + 0{,}04\,(\varepsilon_{x-4} + \varepsilon_{x+4})$$

fórmula que nos da los excesos $V_x^P - y_x$ sobre la recta ajustada AB de las cifras censales promedias ajustadas por el método de Wittstein.

De análoga forma se han obtenido los excesos sobre la recta AB de las cifras censales promedias, ajustadas por los métodos de Woolhouse y Karup. (Cuadro V.)

Llevados los valores así obtenidos a un gráfico (fig. 36), se aprecia claramente la ondulación; presentando sus máximos en los ceros y sus mínimos en los cincos y que, por haber trabajado con cifras promedias, a todo el censo afecta.

Para comprobar si se presenta ondulación al ajustar la población censal española por los métodos de King y Gini, he utilizado un método auxiliar que mantiene, aproximadamente, la constancia de suma por grupos quinquenales de edad que, tanto la primera fase del método de King como el método de Gini establecen. Este método, designado por el número 3 de gradación, consiste en hacer pasar parábolas de segundo grado por cada tres puntos consecutivos: 5′, 10′ y 15′ de la figura 35, y da como cifra ajustada los valores representados por los segmentos que estas parábolas determinan en las ordenadas de las tres clases quinquenales que comprenden. Ajustado así el censo, se ha llevado a un gráfico (fig. 40), en el que se ve la ondulación que, lo mismo que los anteriores métodos de gradación, presentan los basados en mantener la constancia de valor por clases quinquenales de edad.

Ensayos para ajustar la población y la mortalidad por nuevos métodos de gradación.—La ondulación en el ajuste se produce con un semiperíodo de cinco edades; luego si hacemos pasar parabolas por edades que se diferencien no en cinco años, sino en diez, habremos encerrado dentro del intervalo los nodos correspondientes a una onda completa y, por tanto, no podrá producirse ondulación en el ajuste ni, como consecuencia, unas edades ya ajustadas tendrán más habitantes que las que les preceden.

Varios fueron los métodos que empleé, y cuya creación fué hecha

Diferencias sobre la recta $y = {}^1/_{10} \sum_{i=3}^{i=12} u_i - {}^1/_{100} \left[\sum_{i=3}^{i=12} u_i - \sum_{i=13}^{i=22} u_i \right] (i - 7{,}5)$ **de valores promedios ajustados por los métodos de Wittstein, Woolhouse y Karup, clasificados por la terminación de la edad declarada**

(CUADRO V)

METODO	Terminación									
	2	3	4	5	6	7	8	9	0	1
Wittstein.	+ 468	— 505	— 705	— 1.746	— 583	— 185	+ 523	+ 944	+ 1.970	+ 1.117
Woolhouse. . . .	+ 1.621	— 1.786	— 938	— 1.899	— 842	— 1.521	+ 1.752	+ 976	+ 1.939	+ 1.186
Karup	+ 437	— 515	— 1.238	— 1.523	— 1.130	— 369	+ 523	+ 1.254	+ 1.624	+ 1.400

con ánimo de corregir los errores que presentaba la aplicación de los métodos extranjeros. A más del ya empleado para comprobar la ondu-

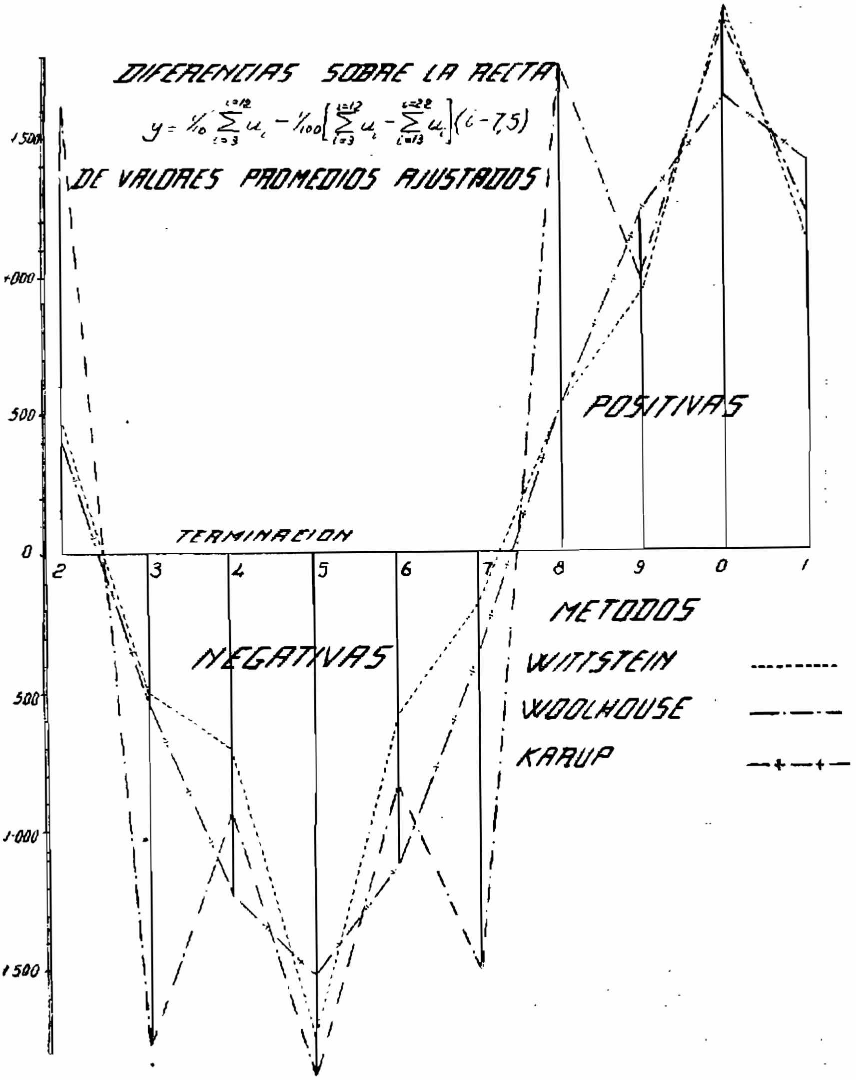


Fig. 36.

lación de ajuste en las clases quinquenales, citaré algunos con los que obtuvimos mejores resultados.

Uno de ellos consiste en hacer pasar parábolas de grado diez (siendo reciso introducir unas modificaciones en la fórmula de interpolación de Newton para su fácil aplicación) por los extremos de las ordenadas representativas de edades, que se diferencian en diez años.

Estas diez parabolas, son:

1.ª parabola pasa por las edades	0, 10, 20, 30	100
2.ª » » » » »	1, 11, 21, 31	101
3.ª » » » » »	2, 12, 22, 32	102
....	....	....
10.ª » » » » »	9, 19, 29, 39	109

Es representado esquemáticamente este método por la figura 37.

Después se obtiene la parábola media de las diez indicadas. Con este métodọ se puede ajustar desde la edad 9 hasta la 100.

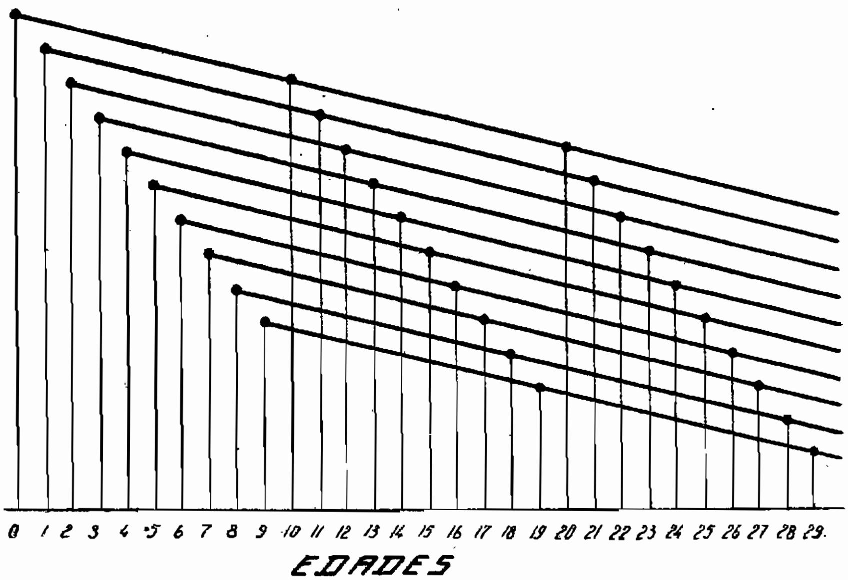


FIG. 37.

El método no presenta ondulación; a más de mantener la influencia que pudieran tener las diferencias finitas, que los métodos extranjeros desprecian. Siendo la parábola de grado diez no quedan, prácticamente, puntos después del 109, por el que pueda pasar la curva.

Otro método fué el de hacer grupos de diez en diez edades y, en cada uno, se ajusta una recta (fig. 38); cumpliendo la condición de que, a más de mantener una suma de cifras ajustadas igual a la de observación, sea, también, constante la suma en los dos grupos de cinco edades componentes.

Después se unen por parábolas de grado diez los valores resultantes del primer ajustamiento de igual terminación y, por último, se obtiene la parábola promedia.

Presenta este método cierta simplicidad de cálculo, pues las ordenadas del primer ajustamiento en los puntos centrales de cada grupo

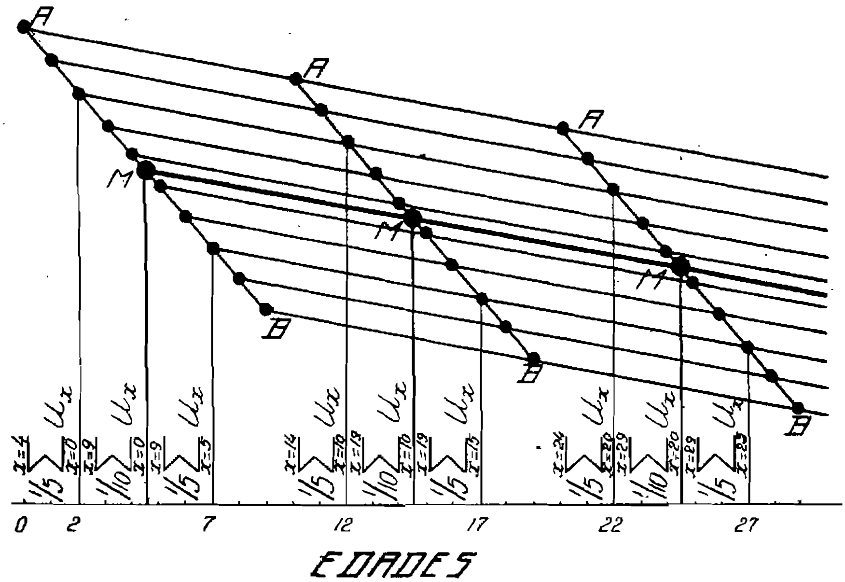


FIG. 38.

de cinco edades, vale $^1/_5$ de la suma y el coeficiente angular se calcula rápidamente, así como el valor central del grupo de diez edades.

De más sencillo cálculo que el anterior, fué una modificación del mismo. Obsérvese que la parábola que pasa por los puntos M es sensiblemente equidistante de cada dos de las diez que se han utilizado, luego si se hiciera pasar una parábola de grado diez por estos puntos centrales, nos encontraríamos con una curva cuyos valores no son promedios de las otras diez, sino equidistantes o equinormales. Los puntos M son de fácil cálculo, pues son $^1/_{10}$ de la suma de cada grupo de observación, y la determinación de la parábola de grado diez que pasa por los once puntos, es de gran sencillez, debida a la forma parabólica que dimos a la fórmula de Newton.

Una modificación de este último método nos produjo el de más fácil cálculo de cuantos métodos de gradación empleamos. Esta modificación consiste en unir cada tres puntos M consecutivos por parábolas de segundo grado.

Los resultados de este método han sido llevados al gráfico (fig. 40) y designado por el número 6 de gradación.

Citaré, por último, otro método; basado en la bisectriz de los pares de rectas que unen terminaciones iguales (fig. 39).

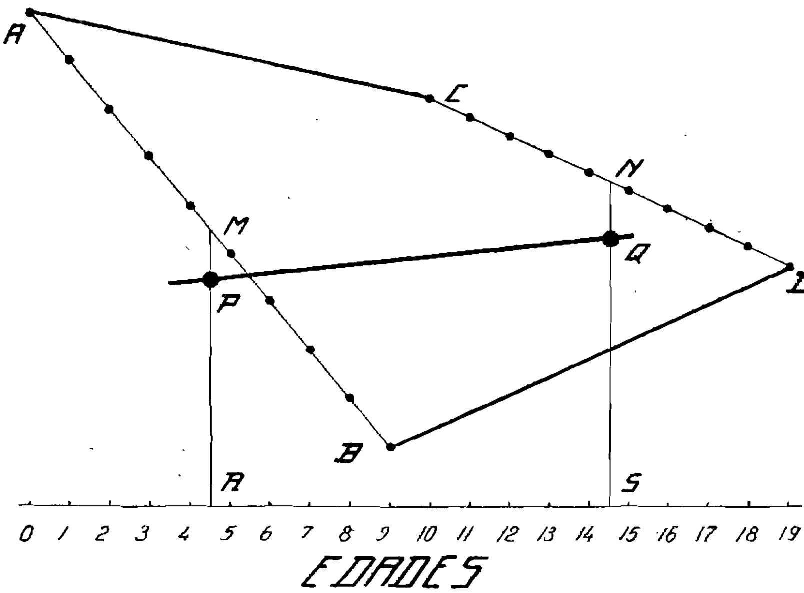
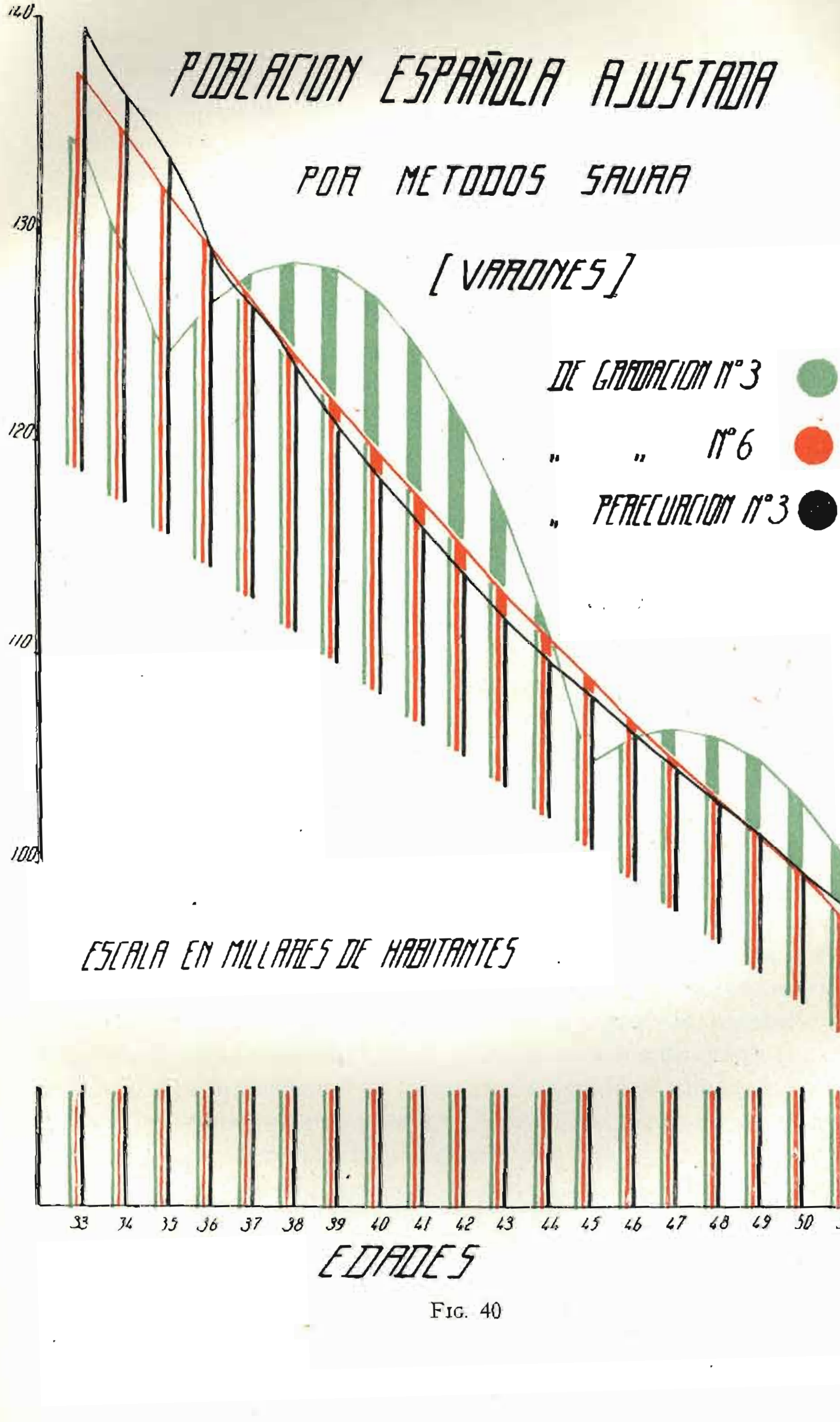


FIG. 39.

Calculadas las dos rectas que mantienen la constancia en dos grupos de diez edades, se obtiene el valor que limita la bisectriz de cada par de rectas A C, B D en las ordenadas centrales de cada grupo de diez; dándonos los puntos P y Q. De análoga forma los otros cinco puntos. Después, sobre cada ordenada central, se calcula el promedio de los diez valores (cinco de las rectas anteriores y cinco de las posteriores) que las bisectrices determinan y, por último, por estos puntos promedios se hace pasar una parábola de grado diez.

Este método, que elimina mejor que los otros los errores ya citados, es de cálculo tan laborioso que lo hace, a nuestro parecer, inaceptable.

Cualquiera de los métodos de gradación extranjeros es también válido para el ajuste de la población española, con sólo introducir la variación del intervalo; pasando de cinco a diez edades. Pero en estas



POBLACION ESPAÑOLA AJUSTADA
POR METODOS SAVAR
[VARONES]
DE GRADACION N° 3
" " N° 6
" PERECUACION N° 3
ESCALA EN MILLARES DE HABITANTES
140
130
120
110
100
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
EDADES

Fig. 40

condiciones, a más de aumentar el trabajo de cálculo, limita extraordinariamente el número de edades susceptibles de ser ajustadas.

Método de perecuación del conferenciante.—Los métodos de gradación empleados en el ajuste de la población sólo dan valores discontinuos, consistiendo su finalidad, como en toda gradación, en sustituir unos valores de observación por otros calculados que se aceptan como más reales que los primeros pero sin formar parte de ley o curva alguna.

Habiendo estimado las ventajas de las funciones en el cálculo de tasas, se orientó la investigación a la obtención de aquéllas, y, por tanto, fueron abandonados los métodos de gradación, recurriendo a los de perecuación.

Aplicados los métodos corrientes de perecuación a la serie de observación expresada por ordenadas, representativas de cada edad, sus resultados, en cuanto a discontinuidad, son los mismos que los de gradación, por lo que, en realidad, sólo dos métodos de los conocidos en la perecuación son posible emplear: uno el de Cantelli y otro el de Gini. El de Cantelli, ya indiqué la dificultad de solucionar el sistema de ecuaciones que emplea, y el de Gini me obligaba a variar el intervalo de cinco a diez edades, produciendo con ello una limitación—lo mismo que en los de gradación—en el número de edades susceptibles de ser ajustadas; por otro lado, no resuelve, más que en parte, el problema, ya que este método no obtiene una función sino trozos de funciones que, después de introducida la variación del intervalo, comprenden sólo diez edades.

Por todo lo anterior trabajé en la formación de un método que resolviera prácticamente el problema. Las dificultades que en primer lugar se presentaban eran las de que usando parámetros que tienen veinte cifras enteras, y que, debido a la precisión necesaria en los trabajos para obtener tasas con seis cifras exactas, es necesario usar veinticinco decimales, empleando en total números de cuarenta y cinco órdenes. Si a la dificultad de trabajo en máquina se une la de resolver por los procedimientos usuales un sistema de ecuaciones, con tales parámetros el problema prácticamente era insoluble.

Como quiera que la función parabólica es la más usada en ajustes de población, fué también aceptada para la solución del problema; influyendo en nuestra decisión, no sólo la costumbre sino también la extraordinaria flexibilidad de esta clase de funciones que la hace adaptable a cualquier serie. Cualidad digna de tenerse en cuenta para la curva de mortalidad por sus varios máximos y mínimos.

Como resultante de todo lo anterior, el sistema a resolver es

$$\sum_{i=x_0}^{i=x_0+h-1} u_i = \int_{x_0}^{x_0+h} P(x)\,dx$$

$$\sum_{i=x_0+h}^{i=x_0+2h-1} u_i = \int_{x_0+h}^{x_0+2h} P(x)\,dx$$

. . . .

$$\sum_{i=x_0+nh}^{i=x_0+(n+1)h-1} u_i = \int_{x_0+nh}^{x_0+(n+1)h} P(x)\,dx$$

En que x_0 es la abcisa del origen de observación, h es el intervalo de perecuación o número de cuantías de observación que se consideran agrupadas en cada ecuación, y u_i son los valores de observación.

Si la función perecuatriz usada es:

$$P(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_{nh}\, x^{nh}$$

el sistema anterior queda transformado en:

$$\sum_{i=x_0}^{i=x_0+h-1} u_i = \left| a_0 x + \frac{a_1}{2} x^2 + \frac{a_2}{3} x^3 + \cdots + \frac{a_{nh}}{nh+1} x^{nh+1} \right|_{x_0}^{x_0+h}$$

$$\sum_{i=x_0+h}^{i=x_0+2h-1} u_i = \left| a_0 x + \frac{a_1}{2} x^2 + \frac{a_2}{3} x^3 + \cdots + \frac{a_{nh}}{nh+1} x^{nh+1} \right|_{x_0+h}^{x_0+2h}$$

. . . .

$$\sum_{i=x_0+nh}^{i=x_0+(n+1)h-1} u_i = \left| a_0 x + \frac{a_1}{2} x^2 + \frac{a_2}{3} x^3 + \cdots + \frac{a_{nh}}{nh+1} x^{nh+1} \right|_{x_0+nh}^{x_0+(n+1)h}$$

Hagamos

$$a_0 = a'_0 \;,,\; \frac{a_1}{2} = a'_1 \;,,\; \frac{a_2}{3} = a'_2 \cdots \frac{a_{nh}}{nh+1} = a'_{nh}$$

con lo que el sistema tomará la forma

$$\sum_{i=x_0}^{i=x_0+h-1} u_i = a'_0[(x_0+h)-x_0] + a'_1[(x_0+h)^2-x_0^2] + \cdots + a'_{nh}\left[(x_0+h)^{nh+1}-x_0^{nh+1}\right]$$

$$\sum_{i=x_0+h}^{i=x_0+2h-1} u_i = a'_0[(x_0+2h)-(x_0+h)] + a'_1[(x_0+2h)^2-(x_0+h)^2] + \cdots$$

$$+ a'_{nh}\left[(x_0+2h)^{nh+1}-(x_0+h)^{nh+1}\right]$$

. . . .

$$\sum_{i=x_0+nh}^{i=x_0+(n+1)h-1} u_i = a'_0[(x_0+(n+1)h)-(x_0+nh)]+a'_1[(x_0+(n+1)h)^2-(x_0+nh)^2]+\cdots$$
$$+a'_{nh}\left[(x_0+(n+1)h)^{nh+1}-(x_0+nh)^{nh+1}\right]$$

Restemos de cada ecuación de este sistema la que le precede, con cuyos resultados formaremos un nuevo sistema y, recordando que

$$\sum_{i=x_0+h}^{i=x_0+2h-1} u_i - \sum_{i=x_0}^{i=x_0+h-1} u_i = \Delta \sum_{i=x_0}^{i=x_0+h-1} u_i$$

....

$$\sum_{i=x_0+nh}^{i=x_0+(n+1)h-1} u_i - \sum_{i=x_0+(n-1)h}^{i=x_0+nh-1} u_i = \Delta \sum_{i=x_0+(n-1)h}^{i=x_0+nh-1} u_i$$

$$(x_0+h)-x_0=\Delta x_0 \;,,\; (x_0+h)^2-x_0^2=\Delta x_0^2 \ldots (x_0+h)^{nh+1}-x_0^{nh+1}=\Delta x_0^{nh+1}$$

$$(x_0+2h)-(x_0+h)=\Delta(x_0+h) \;,,\; (x_0+2h)^2-(x_0+h)^2=\Delta(x_0+h)^2 \ldots (x_0+2h)^{nh+1}-$$
$$-(x_0+h)^{nh+1}=\Delta(x_0+h)^{nh+1}$$

....

$$(x_0+(n+1)h)-(x_0+nh)=\Delta(x_0+nh) \;,,\; (x_0+(n+1)h)^2-(x_0+nh)^2=$$
$$=\Delta(x_0+nh)^2 \ldots (x_0+(n+1)h)^{nh+1}-(x_0+nh)^{nh+1}=\Delta(x_0+nh)^{nh+1}$$

el sistema habrá quedado reducido a

$$\Delta \sum_{i=x_0}^{i=x_0+h-1} u\,i = a'_0[\Delta(x_0+h)-\Delta x_0]+a'_1[\Delta(x_0+h)^2-\Delta x_0^2]+\cdots$$
$$+a'_{nh}[\Delta(x_0+h)^{nh+1}-\Delta x_0^{nh+1}$$

$$\Delta \sum_{i=x_0+h}^{i=x_0+2h-1} u\,i = a'_0[\Delta(x_0+2h)-\Delta(x_0+h)]+a'_1[\Delta(x_0+2h)^2-\Delta(x_0+h)^2]+\cdots$$
$$+a'_{nh}[\Delta(x_0+2h)^{nh+1}-\Delta(x_0+h)^{nh+1}]$$

....

$$\Delta \sum_{i=x_0+(n-1)h}^{i=x_0+nh-1} u\,i = a'_0[\Delta(x_0+nh)-\Delta(x_0+(n-1)h)]+a'_1[\Delta(x_0+nh)^2-$$
$$-\Delta(x_0+(n-1)h)^2]+\cdots+a'_{nh}[\Delta(x_0+nh)^{nh+1}-\Delta(x_0+(n-1)h)^{nh+1}]$$

Según la propiedad que tienen las progresiones aritméticas de ser nulas las diferencias finitas de orden $n+1$ de la serie de potencias n^{esima} de sus términos, serán cero las

$$\Delta(x_0+h)-\Delta x_0=\Delta^2 x_0 \;,,\; \Delta(x_0+2h)-\Delta(x_0+h)=\Delta^2(x_0+h) \ldots$$
$$\Delta(x_0+nh)-\Delta(x_0+(n-1)h)=\Delta^2(x_0+(n-1)h)$$

con lo que el sistema se reducirá en una ecuación por la resta efectuada, y en un término, su segundo miembro, por ser nulas las diferencias finitas que multiplican al valor a determinar a'_0. Queda, pues, el sistema en la fase

$$\Delta \sum_{i=x_0}^{i=x_0+h-1} u_i = a'_1 \Delta^2 x_0^2 + a'_2 \Delta^2 x_0^3 + \cdots + a'_{nh} \Delta^2 x_0^{nh+1}$$

$$\Delta \sum_{i=x_0+h}^{i=x_0+2h-1} u_i = a'_1 \Delta^2 (x_0+h)^2 + a'_2 \Delta^2 (x_0+h)^3 + \cdots + a'_{nh} \Delta^2 (x_0+h)^{nh+1}$$

....

$$\Delta \sum_{i=x_0+(n-1)h}^{i=x_0+nh-1} u_i = a'_1 \Delta^2 (x_0+(n-1)h)^2 + a'_2 \Delta^2 (x_0+(n-1)h)^3 + \cdots$$
$$+ a'_{nh} \Delta^2 (x_0+(n-1)h)^{nh+1}$$

Restando nuevamente de cada ecuación la que le precede y siendo nulas las

$$\Delta^2 (x_0+h)^2 - \Delta^2 x_0^2 = \Delta^3 x_0^2 \text{ ,, } \Delta^2 (x_0+2h)^2 - \Delta^2 (x_0+h)^2 =$$
$$= \Delta^3 (x_0+h)^2 \text{ ,, } \cdots \Delta^2 (x_0+(n-1)h)^2 - \Delta^2 (x_0+(n-2)h)^2 = \Delta^3 (x_0+(n-2)h)^2$$

obtendremos el siguiente sistema, que tiene otra ecuación menos y otro término menos

$$\Delta^2 \sum_{i=x_0}^{i=x_0+h-1} u_i = a'_2 \Delta^3 x_0^3 + a'_3 \Delta^3 x_0^4 + \cdots + a'_{nh} \Delta^3 x_0^{nh+1}$$

$$\Delta^2 \sum_{i=x_0+h}^{i=x_0+2h-1} u_i = a'_2 \Delta^3 (x_0+h)^3 + a'_3 \Delta^3 (x_0+h)^4 + \cdots + a'_{nh} \Delta^3 (x_0+h)^{nh+1}$$

....

$$\Delta^2 \sum_{i=x_0+(n-2)h}^{i=x_0+(n-1)h-1} u_i = a'_2 \Delta^3 (x_0+(n-2)h)^3 + a'_3 \Delta^3 (x_0+(n-2)h)^4 + \cdots$$
$$+ a'_{nh} \Delta^3 (x_0+(n-2)h)^{nh+1}$$

Continuado el proceso de restar en cada nuevo sistema de cada ecuación la que le precede y anulándose las diferencias finitas que afectan al primer término de cada ecuación, llegaremos, por último, a la ecuación

$$\Delta^{nh} \sum_{i=x_0}^{i=x_0+h-1} u_i = a'_{nh} \Delta^{nh+1} x_0^{nh+1}$$

y de cuya igualdad obtendremos

$$a'_{nh} = \frac{1}{\Delta^{nh+1}\, x_0^{nh+1}} \cdot \Delta^{nh} \sum_{i=x_0}^{i=x_0+h-1}$$

Este valor, sustituído en la primera ecuación del penúltimo sistema,

$$\Delta^{nh-1} \sum_{i=x_0}^{i=x_0+h-1} u_i = a'_{nh-1}\, \Delta^{nh}\, x_0^{nh} + a'_{nh}\, \Delta^{nh}\, x_0^{nh+1}$$

dará

$$a'_{nh-1} = \frac{1}{\Delta^{nh}\, x_0^{nh}} \cdot \Delta^{nh-1} \sum_{t=x_0}^{i=x_0+h-1} u_i - a'_{nh} \frac{\Delta^{nh}\, x_0^{nh+1}}{\Delta^{nh}\, x_0^{nh}}$$

y sucesivamente tendremos todos los parámetros

$$a'_{nh-2}\,,\ a'_{nh-3}\,,\ \cdots a'_3\,,\ a'_2\,,\ a'_1\,,\ a'_0$$

En el sistema, tal como ha sido resuelto, sólo es preciso conocer las sucesivas diferencias finitas del grupo de perecuación

$$\sum_{i=x_0}^{i=x_0+h-1} u_i$$

ya que, tanto las diferencias finitas del intervalo de perecuación como sus cocientes, son números independientes de la serie de observación y cuyos valores, para los intervalos de uso más corriente y para grados de la función más empleados, serán dados en las *Tablas Españolas de Mortalidad.*

Aplicado mi método de perecuación a un ejemplo de serie de población, permitirá ver la sencillez de trabajo del mismo.

Por tratarse de un ejemplo, fijemos como límites de la serie las edades 7 y 26 años. Límites entre los que se da el mínimo de las tasas de muerte.

El intervalo de perecuación que emplearemos es $h = 5$ y, por empezar la serie en 7 años, será $x_0 = 7$ y el sistema primario de ecuaciones

$$\sum_{i=7}^{i=11} u_i = \int_7^{12} P(x)\,dx$$

$$\sum_{i=12}^{i=16} u_i = \int_{12}^{17} P(x)\,dx$$

$$\sum_{i=17}^{i=21} u_i = \int_{17}^{22} P(x)\,dx$$

$$\sum_{i=22}^{i=26} u_i = \int_{22}^{27} P(x)\,dx$$

y las que nos dan los coeficientes son

$$a'_3 = \frac{a_3}{4} = \frac{1}{\Delta^4 x_0^4} \cdot \Delta^3 \sum_{i=7}^{i=11} u_i$$

$$a'_2 = \frac{a_2}{3} = \frac{1}{\Delta^3 x_0^3} \cdot \Delta^2 \sum_{i=7}^{i=11} u_i - \frac{\Delta^3 x_0^4}{\Delta^3 x_0^3} \cdot a'_3$$

$$a'_1 = \frac{a_1}{2} = \frac{1}{\Delta^2 x_0^2} \cdot \Delta \sum_{i=7}^{i=11} u_i - \frac{\Delta^2 x_0^3}{\Delta^2 x_0^2} \cdot a'_2 - \frac{\Delta^2 x_0^4}{\Delta^2 x_0^2} \cdot a'_3$$

$$a'_0 = a_0 = \frac{1}{\Delta x_0} \cdot \sum_{i=7}^{i=11} u_i - \frac{\Delta x_0^2}{\Delta x_0} \cdot a'_1 - \frac{\Delta x_0^3}{\Delta x_0} \cdot a'_2 - \frac{\Delta x_0^4}{\Delta x_0} \cdot a'_3$$

La serie de los *factores de transformación*, o relaciones entre diferencias finitas de potencias del intervalo, toma para una aproximación de 15 cifras decimales los valores

$$\frac{1}{\Delta^4 x_0^4} = f_3 = 66 \cdot 666 \cdot 666 \cdot 667 \times 10^{-15}$$

$$\frac{1}{\Delta^3 x_0^3} = f_2 = 1 \cdot 333 \cdot 333 \cdot 333 \cdot 333 \times 10^{-15}$$

$$\frac{\Delta^3 x_0^4}{\Delta^3 x_0^3} = f_2^4 = 30$$

$$\frac{1}{\Delta^2 x_0^2} = f_1 = 2 \times 10^{-2}$$

$$\frac{\Delta^2 x_0^3}{\Delta^2 x_0^2} = f_1^3 = 15$$

$$\frac{\Delta^2 x_0^4}{\Delta^2 x_0^2} = f_1^4 = 175$$

$$\frac{1}{\Delta x_0} = f_0 = 2 \times 10^{-1}$$

$$\frac{\Delta x_0^2}{\Delta x_0} = f_0^2 = 5$$

$$\frac{\Delta x_0^3}{\Delta x_0} = f_0^3 = 25$$

$$\frac{\Delta x_0^4}{\Delta x_0} = f_0^4 = 125$$

Veamos los cálculos a realizar para perecuar la parábola de tercer grado

$$P(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3$$

en la serie de población.

En primer lugar, hemos de obtener de la serie de observación las

$$\sum_{i=7}^{i=11} u_i = 1 \cdot 176 \cdot 025$$

$$\sum_{i=12}^{i=16} u_i = 1 \cdot 112 \cdot 519$$

$$\sum_{i=17}^{i=21} u_i = 938 \cdot 740$$

$$\sum_{i=22}^{i=26} u_i = 834 \cdot 958$$

y, seguidamente, el cuadro de las diferentes finitas de esta serie

Σu_i	$\Delta \Sigma u_i$	$\Delta^2 \Sigma u_i$	$\Delta^3 \Sigma u_i$
1 · 176 · 025			
	— 63 · 506		
1 · 112 · 519		— 110 · 273	
	— 173 · 779		180 · 270
938 · 740		69 · 997	
	— 103 · 782		
834 · 958			

y, finalmente, los coeficientes

$$\Delta^3 \sum_{i=7}^{i=11} u_i \times f_3 = 180 \cdot 270 \times 66 \cdot 666 \cdot 666 \cdot 667 \times 10^{-15} \qquad = \; \ldots\ldots\ldots\ldots\ldots$$

$$a'_3 \qquad = \qquad 12{,}018$$

$$\Delta^2 \sum_{i=7}^{i=11} u_i \times f_2 = -110 \cdot 273 \times 1 \cdot 333 \cdot 333 \cdot 333 \cdot 333 \cdot \times 10^{-15} \; = \; \ldots\ldots\ldots\ldots\ldots$$

$$-a'_3 \times f_2^4 = -12{,}018 \times 30 \qquad = \; \ldots\ldots\ldots\ldots\ldots$$

$$a'_2 \qquad = \qquad -507{,}570 \cdot 666 \cdot 666 \cdot 666 \cdot 667$$

$$\Delta \sum_{i=7}^{i=11} u_i \times f_1 = -63 \cdot 506 \times 2 \times 10^{-2} \quad = \ldots\ldots\ldots$$

$$-a'_2 \times f_1^3 = 507{,}570 \cdot 666 \cdot 666 \cdot 666 \cdot 667 \times 15 \quad = \ldots\ldots\ldots$$

$$-a'_3 \times f_1^4 = -12{,}018 \times 175 \quad = \ldots\ldots\ldots$$

$$a'_1 = 4 \cdot 240{,}29$$

$$\sum_{i=7}^{i=11} u_i \times f_0 = 1 \cdot 176 \cdot 025 \times 2 \times 10^{-1} \quad = \ldots\ldots\ldots$$

$$-a'_1 \times f_0^2 = -4 \cdot 240{,}29 \times 5 \quad = \ldots\ldots\ldots$$

$$-a'_2 \times f_0^3 = 507{,}570 \cdot 666 \cdot 666 \cdot 666 \cdot 667 \times 25 \quad = \ldots\ldots\ldots$$

$$-a'_3 \times f_0^4 = -12{,}018 \times 125 \quad = \ldots\ldots\ldots$$

$$a'_0 = 225 \cdot 190{,}566 \cdot 666 \cdot 666 \cdot 666 \cdot 667$$

Siendo

$$\int_0^x P(x)\,dx = 225 \cdot 190{,}566 \cdot 666 \cdot 666 \cdot 666 \cdot 667\, x + 4 \cdot 240{,}29\, x^2 - 507{,}570 \cdot 666 \cdot 666 \cdot 666 \cdot 667\, x^3 + 12{,}018\, x^4$$

y la función perecuada

$$P(x) = 225 \cdot 190{,}566 \cdot 666 \cdot 666 \cdot 666 \cdot 667 + 8 \cdot 480{,}58\, x - 1 \cdot 522{,}712\, x^2 + 48{,}072\, x^3$$

En este ejemplo, y para mayor sencillez en los cálculos, se ha efectuado un cambio de origen en el segundo miembro equivalente a siete unidades, con lo que, haciendo $x = 0$ tanto en la perecuatriz como en las integrales, se obtendrán los valores correspondientes a 7 años; entre ellos el ajustado.

Funciones de población, mortalidad y de tasas de muerte.—Calculada la perecuatriz

$$P(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_{nh} x^{nh}$$

esta es la *función de población* y su

$$\int_x^{x+1} P(x)\,dx = \Big| a'_0 x + a'_1 x^2 + a'_2 x^3 + \cdots + a'_{nh} x^{nh+1} \Big|_x^{x+1}$$

expresa el valor ajustado del número de habitantes de edad comprendida entre x y $x + 1$ años, y también, como es lógico, el número de habitantes comprendidos entre dos límites cualquiera de edad, mayor o menor que un año.

Volviendo al ejemplo anteriormente resuelto la

$$\int_{x}^{x+1} P(x)\,dx$$

resuelve el ajuste de la población y nos da el número de habitantes de cada edad, entre los 7 y los 26 años.

De igual forma, y por el mismo procedimiento, se pueden ajustar las defunciones ocurridas durante un año, de las edades comprendidas entre 7 y 26 años, con datos de observación sacados del resumen anual del *Movimiento Natural de la Población*.

La perecuatriz correspondiente a mortalidad o *función de mortalidad* toma el siguiente valor en el ejemplo en curso:

$$D(x) = 1 \cdot 575{,}116 \cdot 666 \cdot 666 \cdot 666 \cdot 667 - 358{,}810\,x + 45{,}322\,x^2 - 1{,}452\,x^3$$

y lo mismo que en la *función de población* la

$$\int_{x}^{x+1} D(x)\,dx$$

nos expresa los muertos ocurridos durante un año, ajustados entre las edades 7 y 26 años.

Las funciones de población y mortalidad no juegan aquí más papel que el de auxiliar para la obtención de otra *función*: la de *tasas de muerte* y también para la *función de tasas de supervivencia*.

Recordemos que, como final de la conferencia de ayer, acepté como expresión para la tasa de muerte la fórmula

$$T_m(x) = \frac{\frac{1}{2}\int_{x}^{x+2} D(x)\,dx}{\int_{x}^{x+1} P(x)\,dx}$$

Siendo parabólicas y del mismo grado las perecuatrices

$$D(x) \qquad y \qquad P(x)$$

sus integrales entre los límites x, $x + 2$ para mortalidad y x, $x + 1$ para población son también funciones parabólicas y del mismo grado que aquéllas. El cociente, por tanto, es también una parabólica y su grado vendrá definido por la aproximación con que se quieran obtener los resultados, pudiendo, en consecuencia, ser igual, mayor o menor que el de las perecuatrices empleadas.

La continuación del ejemplo ya indicado aclarará lo anterior. Si en

$$\tfrac{1}{2}\int_{x}^{x+2} D(x)\,dx = \tfrac{1}{2}\Big|\, b'_0 x + b'_1 x^2 + b'_2 x^3 + b'_3 x^4 \,\Big|_{x}^{x+2}$$

ponemos límites y ordenamos, tendremos

$$\tfrac{1}{2}\int_{x}^{x+2} D(x)dx = [b'_0 + 2b'_1 + 4b'_2 + 8b'_3] + [2b'_1 + 6b'_2 + 16b'_3]x + [3b'_2 + 12b'_3]x^2 + 4b'_3 x^3$$

o sea, en el ejemplo después de sustituir parámetros

$$\tfrac{1}{2}\int_{x}^{x+2} D(x)\,dx = 1 \cdot 273{,}232 - 273{,}974\,x + 40{,}966\,x^2 - 1{,}452\,x^3$$

Para el denominador de la *función de tasas de muerte* se obtiene del mismo modo

$$\int_{x}^{x+1} P(x)\,dx = \Big|\, a'_0 x + a'_1 x^2 + a'_2 x^3 + a'_3 x^4 \,\Big|_{x}^{x+1}$$

que, después de poner límites y ordenar, queda

$$\int_{x}^{x+1} P(x)\,dx = [a'_0 + a'_1 + a'_2 + a'_3] + [2a'_1 + 3a'_2 + 4a'_3]\,x + [3a'_2 + 6a'_3]\,x^2 + 4a'_3 x^3$$

y que, en el ejemplo, toma el valor

$$\int_{x}^{x+1} P(x)\,dx = 228 \cdot 935{,}304 + 7 \cdot 005{,}94\,x - 1 \cdot 450{,}604\,x^2 + 48{,}072\,x^3$$

y, como consecuencia,

$$T_m(x) = \frac{1 \cdot 273{,}232 - 273{,}974\,x + 40{,}966\,x^2 - 1{,}452\,x^3}{228 \cdot 935{,}304 + 7 \cdot 005{,}94\,x - 1 \cdot 450{,}604\,x^2 + 48{,}072\,x^3}$$

Si esta división de polinomios la efectuamos en el mismo orden en que están escritos, el número de términos del cociente será, en general, ilimitado, y el grado del mismo crecerá a medida que aquel número lo haga. El resto, en cada división parcial, nos marcará el error cometido, de no seguir la división, y la *precisión* fijada al cociente nos dirá cuál ha de ser el término en que podemos detener la división.

Cualquiera que sea el número de términos que obtengamos en el cociente, lo cierto es que

$$T_m(x) = M_0 + M_1\,x + M_2\,x^2 + \cdots + M_n\,x^n$$

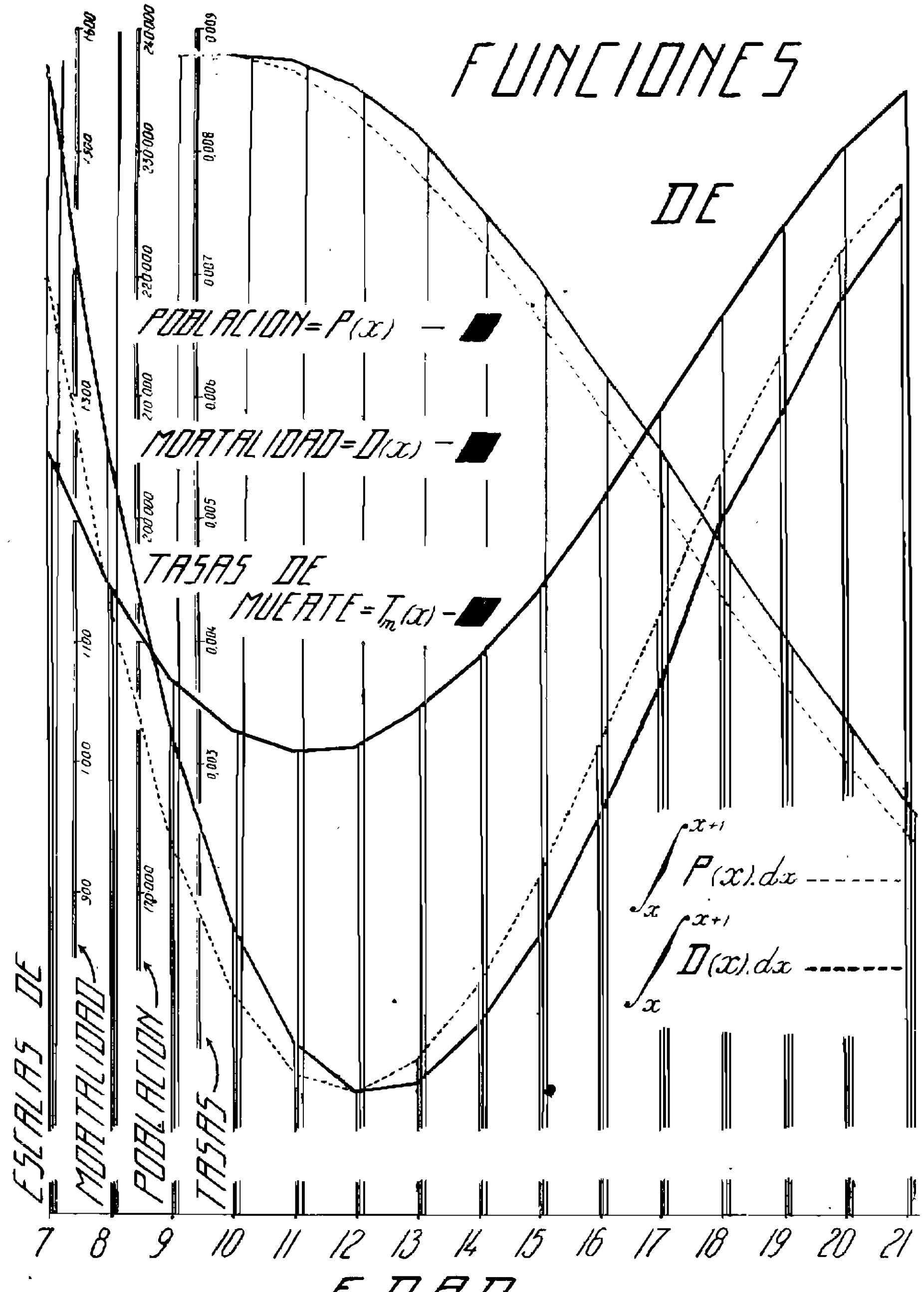

FUNCIONES DE
POBLACION = P(x)
MORTALIDAD = D(x)
TASAS DE MUERTE = $T_m(x)$
$\int_x^{x+1} P(x).dx$
$\int_x^{x+1} D(x).dx$
ESCALAS DE
MORTALIDAD
POBLACION
TASAS
900
1000
1100
1300
1500
1600
170 000
200 000
210 000
220 000
230 000
240 000
0,003
0,004
0,005
0,006
0,007
0,008
0,009
7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
EDAD

FIG. 41

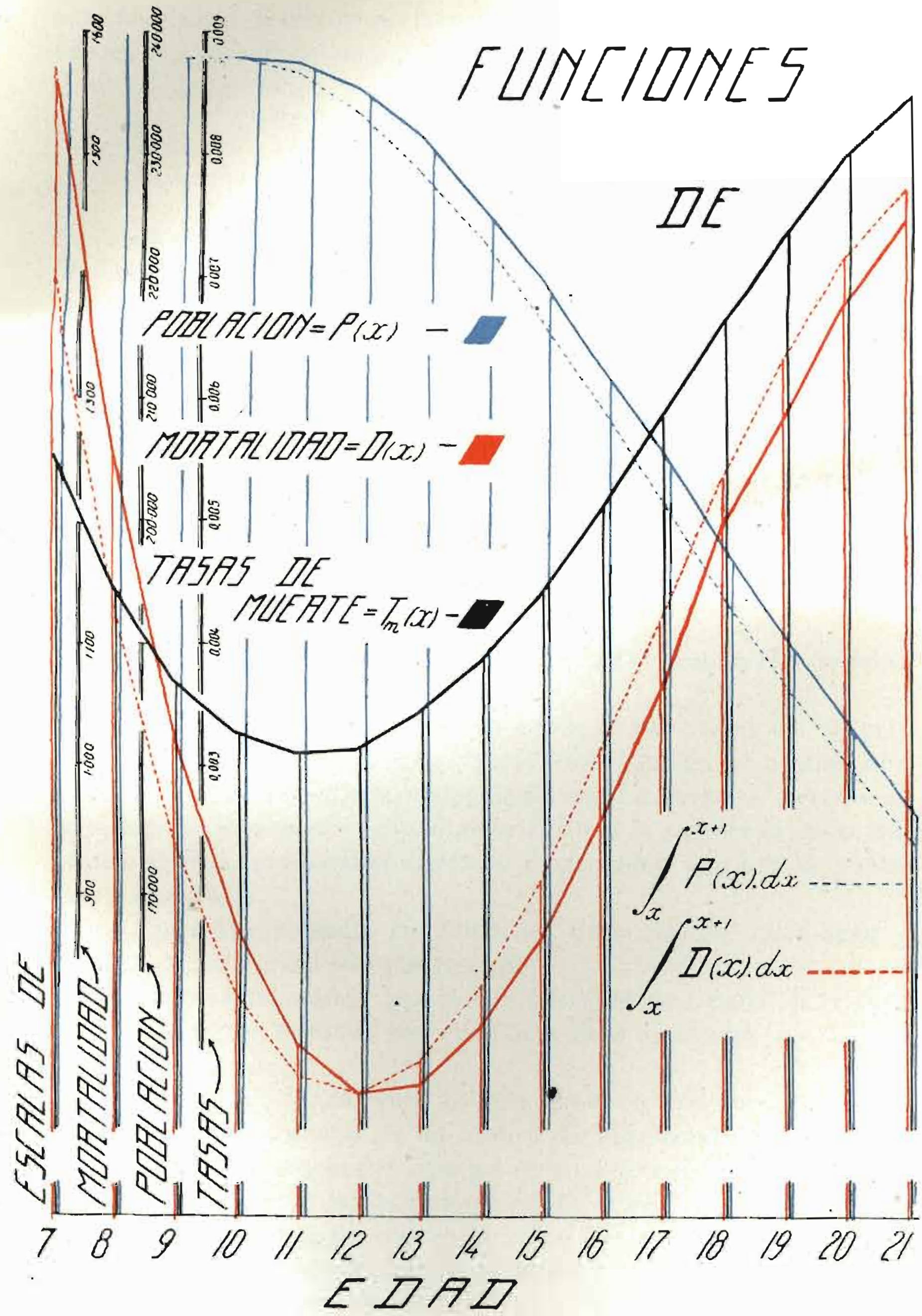

FUNCIONES
DE
POBLACION = P(x)
MORTALIDAD = D(x)
TASAS DE MUERTE = $T_m(x)$
$\int_x^{x+1} P(x).dx$
$\int_x^{x+1} D(x).dx$
ESCALAS DE
MORTALIDAD
POBLACION
TASAS
1.400
1.300
1.200
1.100
1.000
900
240.000
230.000
220.000
210.000
200.000
170.000
0,009
0,008
0,007
0,006
0,005
0,004
0,003
7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
EDAD

FIG. 41

función de tasas de muerte que nos da la probabilidad de morir entre las edades x y $x + 1$ durante un año, de los habitantes que, en principio de observación, tienen x años.

Si x toma el valor $x + \Delta x$ la

$$T_m (x + \Delta x) = M_0 + M_1 (x + \Delta x) + M_2 (x + \Delta x)^2 + \cdots + M_n (x + \Delta x)^n$$

expresa la probabilidad que de morir con edades $x + \Delta x$ a $x + \Delta x + 1$ durante un año tienen los habitantes que en principio de observación hubieran cumplido $x + \Delta x$ años. Cualquiera que sea el punto escogido de la curva de tasas de muerte, siempre nos dará la función $T_m (x)$ la probabilidad que de morir en el plazo de un año tienen los habitantes que, en principio de observación, habían cumplido la edad indicada por la abcisa.

La complementaria de $T_m (x)$ nos da la *función de tasas de supervivencia*

$$T_s (x) = 1 - T_m (x) = S_0 + S_1 x + S_2 x^2 + \cdots + S_n x^n =$$
$$1 - M_0 - M_1 x - M_2 x^2 - M_3 x^3 - \cdots - M_n x^n$$

siendo el único tipo de tasa de muerte en que son complementarias las tasas de muerte y supervivencia.

En las funciones de tasas de muerte y de supervivencia así obtenidas, no sólo se evitan errores de cálculo, sino también el obtener una curva deformada a fuerza de manipulaciones. *La principal ventaja de mi método de perecuación es la de poder llegar a la función de tasas con un solo ajuste y sin necesidad de obtener ningún valor ajustado de población o mortalidad.*

La figura 41 representa las funciones de población, mortalidad y de tasas de muerte del ejemplo resuelto.

Los valores del gráfico, aun siendo parte de una parábola, vienen unidos por rectas, evitando con ello falsas ideas motivadas por habilidades de dibujante.

Aun cuando lo fundamental ha sido expuesto, me veo obligado a no tratar bastantes puntos de las *Tablas de Mortalidad,* entre los que destacan los grupos censales de edad de error compensado, los coeficientes de simpatía en las terminaciones de edad, la precisión en los cálculos, funciones de supervivencia, tasas instantáneas, etc., pero que en unión de la ampliación de todo lo expuesto, serán dados a conocer en las *Tablas Españolas de Mortalidad.*

Antes de terminar he de expresar mi gratitud a cuantos estadísticos con sus consejos y palabras de aliento han contribuído a lograr las *Tablas Españolas de Mortalidad*. A mis jefes y compañeros, siempre animándome a continuar mis trabajos, y al catedrático de Estadística matemática D. Olegario Fernández Baños, a quien debo, no sólo mis conocimientos, sino también el interés que supo despertar en mí por estos problemas de la investigación, les corresponde no pequeña parte de lo que de útil tenga lo expuesto, y, por último, al Catedrático D. Antonio Lasheras-Sanz, ilustre Presidente del Instituto de Actuarios Españoles, debo la invitación a ocupar esta tribuna, lo que ha posibilitado el dar a conocer este trabajo a la Asamblea, que con tanta benevolencia escuchó mis palabras y de la que guardaré un imperecedero recuerdo.