Trabajos de Colaboración

Ampliación del campo de utilidad de las funciones estadísticas de Quiquet (*)

por Antonio LASHERAS-SANZ

I

Introducción

El estudio cuantitativo de los fenómenos que se presentan como pluralidad o masa de cosas susceptibles de variar sin regla dotada del más absoluto rigor, lo que se llama *colectivo, población* o *universo*, constituye en líneas generales, como sabemos, el objeto de la Estadística.

Cada caso particular de los que integran el colectivo, determinado por una o varias circunstancias simultáneas características de naturaleza *cualitativa* —denominadas *atributos*— o *cuantitativa* —*variables*— aun cuando unas y otras puedan ser tratadas con significación matemática, como variables in-

---

(*) Trabajo publicado en una primera parte en el 6.º Suplemento doctrinal, correspondiente al año 1954, XV de su publicación (2.ª época) del *Boletín de Estadística*, del Instituto Nacional de Estadística de España; en el número 1 de la *Revista Internacional de Estadística y Actuariado de la Seguridad Social* (que fue editada por la "Asociación Internacional para la Seguridad Social" —O. I.-T.— Ginebra, Suiza), y dicha parte reproducida y en otro número completada, de los *Anales del Instituto Actuarial Argentino*, que se reproduce aquí a petición de miembros de este Instituto de Actuarios Españoles, para conocimiento de todos ellos, con algunas ligeras modificaciones de redacción.

dependientes (aun en el caso de concurrencia de varias circunstancias características simultáneas, que pueden ser homogéneas o mixtas, en el que cada conjunción se tomará como una sola variable de composición compleja), constituye una *unidad estadística*. Y un conjunto de unidades estadísticas homogéneas con respecto a una misma característica, forman una clase que recibe el nombre de *cantidad* o *dato estadístico*.

Varias cantidades estadísticas concernientes a un mismo fenómeno y referidas a una sucesión del tiempo, constituyen una *serie;* y varias cantidades estadísticas expresivas de variedades graduadas de una misma especie, según una escala de medida previamente establecida conforme la graduación de que sea susceptible la característica —o conjunto unitario complejo de características—, constituyen una *seriación*, aunque comúnmente llamada también *serie de frecuencias*, y las de tiempo, *series cronológicas*. Si la clasificación no procede de una variedad gradual, sino de la ordenación, con respecto a un cierto criterio, de elementos discontinuos, se tiene una *tabla, cuadro* o *estado demostrativo*, llamados también, de ordinario, *serie estadística;* como también se suele llamar "tablas" a series y seriaciones, que es el caso de las "tablas de supervivencia o/y de mortalidad".

La trayectoria que sigue un fenómeno a través de los sucesivos términos de la serie que lo refleje cuantitativamente, puede ser regular, según una forma adecuada a la naturaleza del mismo, pero lo más corriente es que sea irregular. Estas irregularidades provendrán de la propia naturaleza del fenómeno y/o de los errores, sistemáticos y/o accidentales, propios de la observación estadística, de los cuales, los primeros, por su sistematismo, pueden quedar neutralizados, y los segundos, corregidos, lo que nos dejará la expresión numérica y gráfica del mismo, con la forma regular o irregular que por su propia índole corresponda a su trayectoria primaria. La representación gráfica de la trayectoria bruta del fenómeno nos revelará visiblemente esas irregularidades de todo género —sobre todo las propiamente naturales y las de los errores accidentales—, y también, aunque difusamente, se podrá intuir la que corresponda como trayectoria regular. Es preciso, pues, que, por algún procedimiento, obtengamos la serie correspondiente a la línea perfectamente definida que pueda reemplazar a la quebrada que representa a la serie directa e inmediatamente resultante de la observación, por otra que presente una absoluta regularidad en la sucesión de sus puntos.

Esta operación es a la que se llama en Estadística suavizado de curvas, o también *perecuación* (de "perequazione", en italiano); *graduación* (del inglés, "graduation"); *ajuste* (del francés, "ajustement"), pero que, en sí, constituyen una *compensación* (como significa en alemán "ausgleichung"). Lo que en realidad constituye es, conforme a lo que significa el vocablo italiano, un reparto equitativo de las diferencias entre los términos consecutivos de la serie observada, en relación con lo que deben ser sus homólogos de la serie correcta. En español se ha adoptado la denominación de *ajuste.*

Esta operación puede practicarse por:

a) Procedimientos puramente mecánicos, como el método gráfico o de Sprague, y los promedios móviles repetidos dos o tres veces para conseguir un buen suavizado.

b) Procedimientos mecánico-analítico-matemáticos, como las fórmulas interpolatorias de Newton y Higham, las "diferencias centrales" y las basadas en ellas, de Stirling y Bessel; la fórmula interpolatoria de Lagrange y las de Karup y Woolhouse. Todos estos métodos para la corrección de las series estadísticas, y otros muchos semejantes a ellos, pueden ser considerados como casos particulares de la fórmula o ecuación general:

$$u'_x = c_0 \cdot u_0 + c_1(u_{x-1} + u_{x+1}) + c_2(u_{x-2} + u_{x+2}) + \cdots$$

para lo cual, Landré plantea el problema-condición de la determinación de los coeficientes c de forma que

$$\sqrt{c_0^2 + 2(c_1^2 + c_2^2 + c_3^2 + \ldots)} = \text{mínimum}$$

Además de por esta condición, el sistema de las c queda definido por la de que la ecuación de corrección que se obtenga no altere la forma de valores u_x.

c) Analítico-matemáticos estrictos o simplemente cuantitativos, consistentes en hallar la forma de la función analítica de tipo conocido que mejor interprete la "función empírica" cuyo valor numérico ha sido proporcionado por la observación estadística. El problema general que se plantea en los métodos de esta clase es el de, dados *n* puntos, (y_1, x_1), (y_2, x_2), ..., (y_n, x_n), determinar la curva regular o continua más sencilla, que tenga con la línea quebrada representativa de los valores directamente observados, el mayor número de puntos comunes, dentro de la irregularidad de la una y la regularidad de la otra, quedando los puntos de la segunda no comunes con las de la primera, a uno y otro lado de los de ésta y diferenciándose lo menos posible. Entre ellos podemos citar primeramente los que proporciona de manera directa e inmediata la Geometría analítica, al amparo de la representación gráfica de la curva que interpreta la sucesión de valores obtenidos por la observación para la "función", en relación con los respectivos de la variable independiente, para proceder luego analíticamente y poder calcular los valores "teóricos" de la "función" correspondiente a los respectivos de la variable independiente que, desde un principio, se nos ofrecen con la debida regularidad que los empíricos no presentan.

Entre los métodos particulares que vienen a resolver el problema general expuesto, podemos citar el de "las áreas", de Cantelli, el de los "polinomios ortogonales" conforme a cualquiera de las interpretaciones de Charlier, Tchebycheff, Romanowsky, etc. También pertenecen a este grupo las fórmulas de las llamadas curvas de Pearson, a que luego vamos a referirnos detalladamente por ser una de las razones del presente trabajo.

d) Analítico-matemático-cualitativas, que son las que, en lugar de corregir la serie observada traduciendo los resultados experimentales, proceden a determinar los coeficientes de una relación algebraica o ecuación, que nos reproduzca cada término de la serie en función de la correspondiente variable característica, de forma que de una manera general podamos representarla por una

$$f(a, b, c, \dots, k, x)$$

donde *a, b, c, ..., k* son parámetros dentro del intervalo definido por toda la extensión de la serie; y una vez adoptada la forma de la función *f*, proceder con estas constantes tratándolas como incógnitas, de forma que sirvan para representar la serie observada tan exactamente como sea posible.

Según Wirtinger, estos medios de corrección se basan esencialmente en la integral:

$$y = m_0 + m_1x + m_2x^2 + \dots + m_{n-1}x^{n-1}$$

de la ecuación a diferencias finitas para la cual $\Delta^n y = 0$.

Casos particulares que resuelven esta cuestión son las funciones de Dormoy, Gompertz, Makeham, Lazarus, Risser y otros, pero quien enfocó la solución del problema con toda generalidad fué Quiquet constituyendo un sistema de curvas tan general que las de Pearson no son más que casos particulares del mismo, la demostración de lo cual constituye el tema de este trabajo.

Ya nos hemos ocupado de él en ocasiones anteriores, en cada una de las cuales hemos ido adicionando algo que nuestra observación o la ajena nos habían hecho apreciar que era preciso tener en cuenta. Entre esas observaciones hay dos que son con las que completamos hoy nuestra exposición. Una de ellas que no habíamos resaltado antes por creer que no era necesario, pues entendíamos que se justificaba por sí misma, pero que, por lo visto, no es así, y la otra, que constituye una solución empírica para la cual, sin que varíe el resultado a que habíamos llegado, creemos haber encontrado una justificación científica que exponemos en su lugar oportuno.

II

Las fórmulas o curvas de Pearson

Partamos de la probabilidad p de que se realice un cierto hecho, y su contraria q. La probabilidad de que ese hecho se realice una, dos, ..., hasta n veces, vendrá expresada, como sabemos, por los términos del desarrollo de $(p + q)^n$, y si tenemos N clasificaciones, los términos de $N(p + q)^n$ darán la frecuencia de distribución de las N clases en $p + 1$ grupos. Pero la serie binomial no representa todas las probabilidades que pueden presentarse, siendo preciso acudir a las series hipergeométricas. Así, la probabilidad de que resulten $r, r - 1, \ldots, 0$ bolas negras cuando de una urna que contiene np negras y nq blancas, si se efectúan r extracciones, viene representada por los términos sucesivos de la serie:

$$\frac{np(np-1)\ldots(np-r+1)}{n(n-1)\ldots(n-r+1)} = 1 + \frac{rnq}{np-r+1} + \frac{r(r-1)}{1\cdot 2}\cdot\frac{nq(nq-1)}{(np-r+1)(np-r+2)} + \ldots$$

Haciendo ahora, en general:

$$y_x = \frac{np(np-1)\ldots(np-r+1)}{n(n-1)\ldots(n-r+1)}\cdot \frac{r(r-1)\ldots(r-x+2)}{(x-1)!}\cdot \frac{nq(nq-1)\ldots(nq-x+2)}{(np-r+1)(np-r+2)\ldots(np-r+x-1)}$$

lo que nos permite obtener:

$$\Delta y_x = y_{x+1} - y_x = y_x \frac{r - x + 1}{x} \cdot \frac{nq - x + 1}{np - r + x} - 1 =$$

$$= y_x \frac{(r + 1)(nq + 1) - x(n + 2)}{x(np - r + x)}$$

por ser $p + q = 1$, e

$$y_{x+\frac{1}{2}} = \tfrac{1}{2}(y_x + y_{x+1}) =$$

$$= \tfrac{1}{2} y_x \cdot \frac{(r + 1)(nq + 1) - x[2(r + 1) + n(q - p)] + 2x^2}{x(np - r + x)}$$

por lo que:

$$\frac{\Delta y_x}{y_{x+\frac{1}{2}}} = \frac{2[(r + 1)(nq + 1) - x(n + 2)]}{(r + 1)(nq + 1) - x[2(r + 1) + n(q - p)] + 2x^2}$$

la cual, dividiendo por $-(n + 2)$ y haciendo

$$k = \frac{2(r + 1)(nq + 1)}{n + 2}$$

$$a = \frac{(r + 1)(nq + 1)}{n + 2}$$

$$b = \frac{2(r + 1) + n(q - p)}{n + 2}$$

y

$$c = \frac{2}{n + 2}$$

puede ser sintetizada, en el límite, de la forma:

$$\frac{1}{y_x} \cdot \frac{dy_x}{dx} = \frac{x + k}{a + bx + cx^2 + \ldots}$$

que es la fórmula de que partió Pearson limitando el polinomio en x al segundo grado, en la que, operando, hace:

$$(a + bx + cx^2) \frac{dy_x}{dx} = y_x(k + x)$$

que luego multiplica por x^n e integra, obteniendo:

$$\int (a + bx + cx^2) \frac{dy_x}{dx} dx = \int y_x(k + x) dx$$

Efectuando la integración del primer miembro por partes, tratando a $\frac{dy_x}{dx}$ como una de las partes, y el segundo miembro, por descomposición, admitiendo que, si en los extremos de la curva, el término en $x^n(a + bx + cx^2)$ se desvanece, llega a la expresión aproximada:

$$- \int nax^{n-1}y_x dx - \int (n + 1) bx^n y_x dx -$$
$$- \int (n + 2) cx^{n+1} y_x dx - \ldots =$$
$$= \int y_x x^{n+1} dx + \int y_x ax^n dx$$

que no es otra cosa que una sucesión de momentos de la forma:

$$\mu'_n = \int y_x x^n dx$$

Esto permite determinar los valores de los parámetros k, a, b, c y de $\frac{1}{y_x} \cdot \frac{dy_x}{dx}$ en función de los momentos.

Luego vuelve a la expresión de que parte y la expone como sigue:

$$\frac{d \log_e y_x}{dx} = \frac{x + k}{a + bx + cx^2} = \frac{x + k}{c(x - x_1)(x - x_2)}$$

a fin de aplicar el criterio general de integración de funciones racionales, teniendo en cuenta los casos que pueden darse en las expresiones de las respectivas raíces:

$$\frac{1}{2c}\left(-b \pm \sqrt{b - 4ac}\right)$$

con lo que se obtienen las expresiones de los tres principales tipos de curvas del sistema de Pearson:

1.º) $$y_0\left(1 + \frac{x}{k_1}\right)^{vk_1}\left(1 - \frac{x}{k_2}\right)^{vk_2}$$

2.º) $$y_0\left(1 + \frac{x^2}{k^2}\right)^{-h} e^{-v \operatorname{arc} \operatorname{tg} \frac{x}{k}}$$

3.º) $$y_0(x - k)^{s_2} x^{s_1}$$

Pero las fórmulas de Pearson ofrecen grandes molestias operatorias, ya que, aparte la laboriosidad del cálculo, presentan inconvenientes como el de que, una vez determinado el *criterio*, dado que los elementos que han servido para ello están afectados de errores y desviaciones, que son los que se trata de evitar con la sustitución de la huella bruta por otra que presente una total regularidad en la sucesión de sus elementos, suele suceder muy frecuentemente que el tipo de curva que resulta aconsejada no sea el de la que realmente interpreta al fenómeno.

III

Nuestra apreciación experimental

Nuestras meditaciones y ensayos animados del propósito de salvar tales inconvenientes, nos hicieron pensar en la posibilidad de reemplazar la serie directamente observada por su correspondiente acumulativa, tomando como primer elemento de ésta el relativo a la edad más joven de la serie bruta; como segundo término de la curva acumulativa, la suma de los dos primeros de la serie original, y así sucesivamente hasta llegar a la edad última de la serie original, para la que corresponde, en la serie acumulativa, un valor igual al de la suma total del colectivo. Esto se nos ocurrió operando con las series de distribución por edades de los componentes de los colectivos laborales, y luego hemos tenido ocasión de extenderlo a colectivos correspondientes a seguros de riesgos no personales.

Procediendo de esta manera, la serie acumulativa ofrece análogas características a las de las series de sobrevivencia, sólo que diferenciándose de éstas en que, en tanto que aquéllas son acumulativas, éstas son desacumulativas.

Se nos ocurrió, pues, interpretar la curva acumulativa mediante fórmulas análogas a las utilizadas para la interpretación de las curvas de sobrevivencia, si bien los valores obtenidos para los parámetros, en las curvas acumulativas, resultaban de carácter inverso a los correspondientes para una curva similar desacumulativa. Hecho así el ajuste de la curva acumulativa y diferenciando luego en campo finito, obteníamos como resultado la curva o serie ajustada sustituta de la observada. La diferencia entre la curva desacumulativa y la acumulativa radica en que la trayectoria del proceso correspondiente sigue la ley natural de evolución de la escala a que, de menor a mayor, siguen los valores de la "variable independiente".

Procediendo empíricamente de esta forma, con ocasión de una visita a Madrid del eminente matemático francés, miembro de Honor de aquel Instituto de Actuarios, Profesor Mr. Frechet, con motivo de unas conferencias pronunciadas por él en el Instituto de Investigaciones Estadisticas, del Consejo Superior de Investigaciones Científicas español, y dado que éramos portadores de unos trabajos realizados con alumnos nuestros sobre la materia, al explicar nuestra manera de proceder, nos preguntó si nos habiamos preocupado de averiguar la explicación científica de todo aquello. Y como no lo habíamos hecho hasta entonces, decidimos ocuparnos de ello, habiendo conseguido lo que sigue a continuación.

Si en lugar de proceder en forma tan subjetiva como lo hizo Pearson, volvemos a la expresión inicial, tenemos que:

$$\frac{d \log_e y_x}{dx} = \frac{x + k}{a + bx + cx^2 + \ldots} = a' + b'x + c'x^2 + \ldots$$

donde hay que tener en cuenta:

1.° Que $\dfrac{d \log_e y_x}{dx}$ es una función $F(x)$, la que, a su vez, es un número obtenido de la observación estadistica, relacionado a la expresión cuantitativa de un "atributo" o de una "variable": u_x.

2.° Que es $x > 1$, por lo que irán creciendo sus potencias sucesivas.

3.° Que, por ser u_x un número que estadísticamente hacemos equivalente a $F(x)$, el polinomio entero en x que desarrolla en serie esta función, por el método de los coeficientes indeterminados, tiene sus términos decrecientes, por lo que, al ser crecientes las sucesivas potencias de x, sus respectivos coeficientes serán funciones decrecientes de esa misma variable y, por tanto, habrá un término, en x^h, que teóricamente tiende a 0, por lo que podemos afirmar que la serie es convergente, y que su límite de convergencia es u_x.

4.° Ahora bien; sabemos que una función continua de una variable no admite más que un solo desarrollo en serie convergente ordenada según las potencias enteras y crecientes de su variable, y que el desarrollo de $F(x)$, conforme a la fórmula de Mac-Laurin, es:

$$F(x) = F(0) + xF'(0) + \frac{x^2}{2!} F''(0) + \frac{x^3}{3!} F'''(0) + \ldots$$

por lo que

$$a' = \varphi(x^0) = F(0)$$
$$b' = \varphi(x^1) = F'(0)$$
$$c' = \varphi(x^2) = \frac{1}{2!} F''(0), \ldots$$

cosa que podemos comprobar derivando sucesivamente $(x + k)/(a + bx + cx^2 + \ldots)$ e igualando a cero esas derivadas, con lo que obtendremos valores numéricos iguales a los que resulten partiendo de

$$x + k = (a' + b'x + c'x^2 + \ldots) \cdot (a + bx + cx^2 + \ldots)$$

e igualando los coeficientes de las mismas potencias de x en ambos miembros, lo que nos proporcionará las siguientes ecuaciones de condición:

$$a'a - a = 0$$
$$a'b - b'a - = 0$$
$$a'c + b'b + c'a + 0$$

y desde el tercer coeficiente de la serie en adelante, la ecuación de recurrencia

$$a'm + b' + c'r = 0$$

en las que m, n, r son los coeficientes respectivos de tres términos consecutivos del polinomio cociente.

Dicho esto, integraremos y pasaremos del logaritmo al número, con lo que resultará:

$$y_x = e^{A+Bx+Cx^2+Dx^3+\ldots}$$

Por cierto que, el número de términos del polinomio que figura como exponente del número e, será ilimitado. Sin embargo, es limitado el número de términos de la serie estadística que se considere, lo que impondrá una limitación en el número de términos del polinomio $A + Bx + Cx^2 + \ldots$, por lo que, por lo menos, tendrá que limitarse el número de términos del polinomio, haciéndolo coincidir con el de términos de la serie a ajustar, lo que, por otra parte, tampoco satisface al ajuste, porque, con ello, se nos reproducirían rigurosamente los términos de la serie observada, en su aspecto bruto, y no conseguiríamos regularización ninguna. Habrá, pues, que limitar dicho número de términos a aquel que nos aconsejen las propias características de la curva observada. De todas formas, una función con cinco o siete parámetros cuando más, nos proporcionará resultados muy aceptables. Sin embargo, es preciso determinar la forma que deberá tener la expresión explícita de la función que haya de adoptarse, la cual, al limitarse en ella el número de términos, plantea el problema de la forma del término complementario o resto.

Para tratar de encontrarla, observemos que la función L(x) es:

$$L(x) = A + Bx + Cx^2 + Dx^3 + \dots + Px^n + Qx^{n+1} + Sx^{n+2} + \dots = L(0) + xL'(0) + \frac{x^2}{2!} L''(0) + \frac{x^3}{3!} L'''(0) + \dots + \frac{x^n}{n!} L^{(n)}(0) + \frac{x^{n+1}}{(n+1)!} L^{(n+1)}(0) + \frac{x^{n+2}}{(n+2)!} L^{(n+2)}(0) + \dots$$

donde podemos hacer:

$$\frac{x^{n+1}}{(n+1)!} L^{(n+1)}(0) + \frac{x^{n+2}}{(n+2)!} L^{(n+2)}(0) + \frac{x^{n+3}}{(n+3)!} L^{(n+3)} + \dots = \frac{x^{n+1}}{(n+1)!} L^{(n+1)}(0) \cdot \left[1 + \frac{x}{n+2} \frac{L^{(n+2)}(0)}{L^{(n+1)}(0)} + \frac{x^2}{(n+2)(n+3)} \frac{L^{(n+3)}(0)}{L^{(n+1)}(0)} + \dots\right] \simeq$$

$$\simeq \frac{x^{n+1}}{(n+1)!} L^{(n+1)}(0) \left[1 + x \log_e r + x^2 \frac{\log^2{}_e r}{2!} + \dots\right] =$$

$$= \frac{x^{n+1}}{(n+1)!} \cdot L^{(n+1)}(0) \cdot r^x$$

puesto que al ser tratados los coeficientes como incógnitas, para que desempeñen su misión de parámetros para toda la extensión de la serie, los valores numéricos que resulten para ellos serán los necesarios para que la serie a ajustar sea reproducida por cálculo con toda la regularidad que su estructura teórica requiera. Incluso, por esta misma razón, podremos tomar como expresión del término complementario:

$$\frac{L^{(n+1)}(0)}{(n+1)!} \cdot r^x$$

Y, entonces, por las razones de tipo empírico o práctico que hemos aducido antes, limitar la forma de la función de ajuste a una del tipo siguiente:

$$y_x = e^{A+Bx+Cx^2+Dr^x} = k \cdot s_1^x \cdot s_2^{x^2} \cdot g^{r^x} \ (^*)$$

Esta función, y_x, si la serie observada responde directamente a ella, es que se trata de una serie desacumulativa, como la de sobrevivencia, por ejemplo; y si la serie directamente observada es de la forma de las típicas de distribución de frecuencias, representa la serie resultante de la acumulación sucesiva de los términos de aquélla. En este caso, para proceder al ajuste de una curva de las de distribución de frecuencias, cualquiera que sea su forma, se formará la acumulativa correspondiente, se ajustará ésta conforme a la función obtenida y, después, se diferenciará en campo finito que es lo que homólogamente corresponde a la expresión inicial:

$$\frac{dy_x}{dx} = \frac{y_x(x+k)}{a+bx+cx^2+\dots}$$

(*) Con una función de semejante tipo, conteniendo más parámetros exponenciales en *s*, hemos logrado un buen ajuste de una serie l_x de sobrevivencia, recogiéndola completa desde la edad cero a la de 106 años, ambas inclusive.

IV

Encuadramiento en el sistema de funciones de Quiquet

La función que hemos obtenido coincide con la llamada "segunda de Makeham". Esta, como la primera de este mismo actuario inglés; la de Gompertz, anterior a ellas y base de las mismas, y las demás que hemos citado en d) de I, que fueron establecidas directamente por sus autores, quedaron luego recogidas como casos particulares del sistema que estableció el actuario francés Quiquet basándose en la propiedad ofrecida por las fórmulas que acabamos de repetir su cita, de que la probabilidad de que más de una cabeza vivan conjuntamente al término de un cierto tiempo dado, pueda ser reemplazada por la probabilidad de que, respectivamente, una sola cabeza de edad promedia u otras tantas cabezas de una misma edad vivan al término de dicho tiempo —propiedad conocida actuarialmente por *ley del envejecimiento uniforme*—.

Dicha propiedad podemos expresarla de la siguiente forma:

$$\frac{f(a+t)}{f(a)} \; \frac{f(b+t)}{f(b)} \; \frac{f(c+t)}{f(c)} \cdots \frac{f(h+t)}{f(h)} =$$
$$= G(\alpha, \beta, \gamma, \ldots, \theta, t)$$

igualdad cuyo primer miembro contiene N factores, y el segundo es una función de $n + 1$ variables independientes.

Obteniendo sus derivadas logarítmicas sucesivas hasta la de orden $n + 1$, se establece un sistema de ecuaciones cuyos primeros miembros son $n + 1$ funciones de N variables independientes; y los segundos, otras tantas funciones de n variables independientes, después de anular el valor de t, porque, si cada igualdad del sistema ha de ser generalmente cierta, tiene que serlo para todos los valores de *t*, incluso el *cero*. Por consiguiente, ha de existir entre los segundos miembros de las ecuaciones de este sistema una relación independiente de las variables, y por consiguiente, entre los primeros miembros de las mismas ecuaciones.

El wronskiano cuya anulación condiciona la existencia de esta relación, proporciona la ecuación diferencial, lineal, homogénea, de coeficientes constantes y sin segundo miembro, de la forma:

$$A_0 y_x + a_1 y_x' + A_2 y_x'' + \dots + A_n y_x^{(n)} = 0$$

cuya ecuación característica de la integración es:

$$A_0 + A_1 r + A_2 r^2 + \dots + A_n r^n = 0$$

y la integral consiguiente:

$$y_x = c_1 e^{r_1 x} + c_2 e^{r_2 x} + \dots + c_n e^{r_n x}$$

que, en caso de raíces múltiples, es:

$$y_x = e^{r_1 x}\,\varphi_1(x) + e^{r_2 x}\,\varphi_2(x) + \dots + e^{r_j x}\,\varphi_j(x)$$

donde $\varphi(x)$ es un polinomio entero en x, de un grado inferior en una unidad al de la multiplicidad de las raíces, r (aquí r tiene significación distinta a la que se le ha atribuido en las últimas fórmulas de III).

Los diferentes tipos de leyes que pueden resultar de esta expresión general dependen del valor de n y de las raíces r que resulten en consecuencia. Por ello, si $n = 0$, será $A_0 = 0$ e $y_x = 0$, por lo que, integrando, se obtiene $f(x) = e$.

Cuando $n = 1$, tendremos $A_0 + A_1 r = 0$, y en el caso de que sea $r = 0$, será $y_x = \xi$, e integrando, se llega a $f(x) = e^{\alpha+\beta x+\gamma x^2}$. Y si es $r \neq 0$, la función resultante será $y_x = e^{\alpha+\beta x+\gamma e}$, que es la primera fórmula de Makeham, en la que cabe considerar el caso particular de $\beta = 0$, que proporciona la expresión $y_x = e^{\alpha+\gamma e}$, que es la de Gompertz.

Para $n = 2$, será $A_0 + A_1 r + A_2 r^2 = 0$, por lo que habrá dos valores de r, pudiendo darse los casos de que sean

$r_1 = r_2 = 0$; $r_1 = 0$ y $r_2 \neq 0$; $(r_1 \neq r_2) \neq 0$; $(r_1 = r_2) \neq 0$. Y si nos detenemos en el caso segundo, la ley que obtendremos será $y_x = e^{\alpha+\beta x+\gamma x^2+\delta e^{rx}}$, que es la segunda de Makeham que es con la que coincide una de las anteriormente obtenidas en III. Ahora bien, si nos detenemos en el caso cuatro, la ley que resultará es $y_x = e^{\alpha+\beta x e^{rx}}$, que responde a una primera acepción del concepto de término complementario o resto, conforme lo hemos considerado en III también.

Si tomamos $n = 3$, la ecuación característica de la integración será $A_0 + A_1 r + A_2 r^2 + A_3 r^3 = 0$, por lo que hay que considerar todos los casos similares que análogamente a cuando $n = 2$, pueden darse respecto a los tres valores de r. Ateniéndonos al en que sean $(r_1 = r_2 = r_3) \neq 0$, la función resultante es $y_x = e^{\alpha+\beta x+\gamma x^2 e^{rx}}$, que responde a lo que acabamos de denominar una primera concepción del término complementario tal y como lo hemos considerado en III.

Si la forma de la función fuese $y_x = {}^{\alpha+\beta x+\gamma x^2+\delta x^3 e^{rx}}$, se trata indudablemente de que es $n = 5$ y, de los valores de las cinco raíces de la ecuación características de la integración, $r_1 = 0$, $(r_2 = r_3 = r_4 = r_5) \neq 0$.

Así sucesivamente podríamos ir considerando casos y más casos semejantes a los dos que en III nos han resultado para las dos concepciones que hemos atribuido a la interpretación del término complementario o resto, pero como ya hemos hecho constar anteriormente, podemos detenernos empíricamente en funciones que comprendan cinco parámetros, pues la práctica nos pone de relieve que, con ellas, pueden conseguirse ajustes completamente satisfactorios.

La circunstancia de haber coincidido con algunos de los tipos de leyes matemáticas de las del sistema de Quiquet, nos permite intuir que cualquier otro tipo de función, de las que constituyen este sistema, puede ser aplicable a la serie o curva integral. Dependerá, en todo caso, de la contextura e incidencias típicas de la curva y, éstas, a su vez, de las de la serie o curva de frecuencias de la que aquélla sea su integral.

La adopción de una u otra función expresiva de la curva integral del fenómeno observado se hará aplicando el criterio de Quiquet para tal fin, debiendo tener en cuenta que, en los precedentes razonamientos, nos hemos desenvuelto en ámbito infinitesimal, pero que las series o que los mismos son aplicables son numéricas y, por ello, pertenecientes al campo discreto, únicamente expresables en términos funcionales continuos formularia o formalísticamente, por equiparación. No obstante, los anteriores razonamientos tienen su paralelismo con otros desarrollados en ámbito finito, al amparo de las diferencias finitas. Y sabido es que casi todas las fórmulas y problemas del cálculo infinitesimal encuentran sus analogías, bajo formas más generales, en el cálculo de las diferencias finitas.

Para ello, con los términos de la serie observada, formaremos la curva de acumulación gradual, con lo que obtendremos, si procedemos a acumular en el sentido del menor al mayor valor de la variable de clase, una curva del tipo de la logística; y si procedemos al revés, una curva desacumulativa del orden de la de supervivencia. En cualquiera de los dos casos, la expresión de la función representativa será la misma, variando únicamente los valores numéricos y los signos que puedan corresponder a los parámetros.

Una vez formada la curva acumulativa, se tomarán los logaritmos de sus términos y se hallarán las sucesivas diferencias finitas, hasta un cierto orden. Si las diferencias segundas son nulas, o pueden ser tenidas por tales, la integración en campo finito nos conducirá a $f(x) = e^{\alpha+\beta x}$.

Si son iguales entre sí,

$$\Delta^2 \log_e f(x) = \ldots = \Delta^2 \log_e f(x+h) = \ldots$$

la función será de la forma

$$f(x) = e^{\alpha+\beta x+\gamma x^2}$$

Cuando se dé

$$\frac{\Delta^2 \log_e f(x+1)}{\Delta^2 \log_e f(x)} = \frac{\Delta^2 \log_e f(x+2)}{\Delta^2 \log_e f(x+1)} = \ldots = e$$

la función será de la forma

$$f(x) = e^{\alpha+\beta x+c\gamma^{x}}$$

En caso de que sea $\Delta^4 \log f(x) = 0$, resultará

$$f(x) = e^{\alpha+\beta x+\gamma x^2+\theta x^3}$$

De producirse que

$$\frac{\Delta^3 \log_e f(x+1)}{\Delta^3 \log_e f(x)} = \frac{\Delta^3 \log_e f(x+2)}{\Delta^3 \log_e f(x+1)} = \ldots = \gamma$$

se tendrá

$$f(x) = e^{\alpha+\beta x+\gamma x^2+\theta\gamma^{x}}$$

Y así sucesivamente para todos los restantes casos imaginables.

Una vez ajustada la curva acumulativa por la función que haya sido procedente, se determinarán las diferencias de primer orden entre los términos de la serie ajustada y obtendremos la curva ajustada del fenómeno observado.

V

Algunos casos prácticos

Para terminar expondremos, siquiera sea someramente, algunos casos prácticos ante los que nos hemos encontrado.

1.º Se trata del ajuste de la serie formada por la distribución, según sus edades, del colectivo que en 1949 constituían los empleados y obreros afectos a los trabajos siderometalúrgicos de la "Empresa Nacional Bazán de Construccio Navales Militares".

El ajuste se efectuó con toda escrupulosidad por las dos fórmulas de Makeham, expuestas en cuanto antecede.

Para la primera, los valores obtenidos para los parámetros fueron:

$$
\begin{aligned}
c &= + \ 0{,}8794498 \\
\log g &= - \ 18{,}801098 \\
\log s &= + \ 0{,}0036741 \\
\log k &= + \ 3{,}8536242
\end{aligned}
$$

Para la segunda fórmula los valores obtenidos para los parámetros fueron:

$$
\begin{aligned}
c &= + \ 0{,}7694834 \\
\log g &= - \ 152{,}45004 \\
\log s_2 &= + \ 0{,}0395715 \\
\log s_1 &= - \ 0{,}0002931 \\
\log k &= + \ 2{,}7607522
\end{aligned}
$$

La serie original, así como su acumulada y las respectivas de éstas, ajustadas por ambas fórmulas de Makeham, figuran en el cuadro I que está al final, así como, en el gráfico I, las curvas relativas a la serie original y a sus dos respectivos ajustes por las fórmulas citadas.

2.° Este se refiere a la Población laboral española en 1944. Los datos procedentes de observación estaban agrupados por grupos de edades, lo cual ya constituía un principio de curva acumulativa. Se rectificó la distribución por grupos homogéneos de edades y se procedió a la acumulación completa. Tal Población estaba separada por sexos y, para ambos, se procedió al ajuste por la segunda fórmula de Makeham. Los datos y resultados figuran en los cuadros II y III, y en el gráfico II. El trazo grueso de éste, para representar la Población laboral total, se obtuvo simplemente por la suma de los resultados parciales correspondientes a cada sexo.

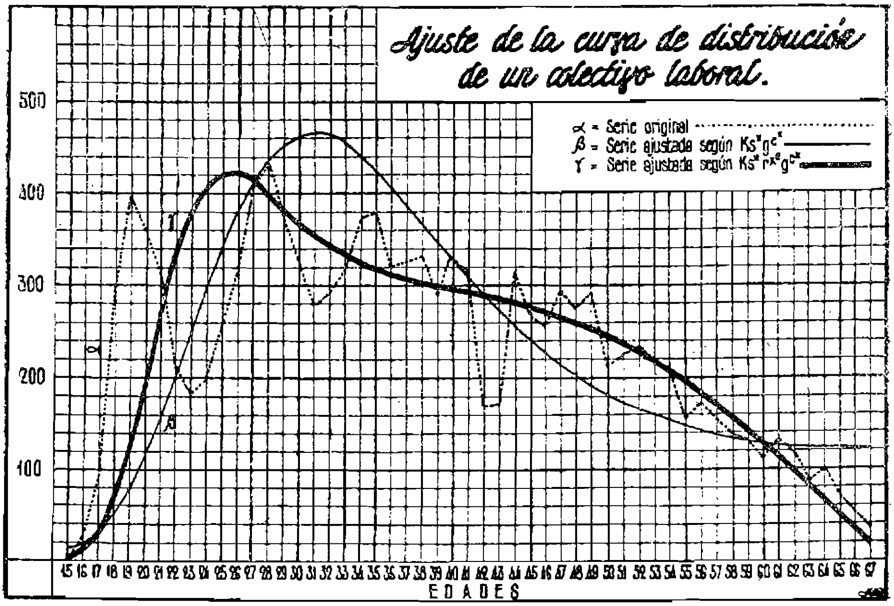
3.° Recientemente, hemos tanteado el ajuste de la curva de los tantos anuales de morir en estado de casado, que adopta también una curva de distribución de frecuencias y hemos podido comprobar que, al aplicar la fórmula segunda de Makeham, ésta no era conveniente, y sí la primera, porque el parámetro en x^2, nos tomó un valor que, por ser tan próximo a *cero*, equivalía a la unidad su potencia x^2.

Cuadro I

Curva de distribución de un colectivo laboral

Edades	Serie sencilla			Serie acumulada		
	Original	Ajustada 1ª	Ajustada 2ª	Original	Ajustada 1ª	Ajustada 2ª
15	2	15	2	2	15	2
16	30	17	8	32	32	10
17	96	31	28	128	63	38
18	287	51	66	415	114	104
19	396	79	125	811	193	229
20	359	114	198	1.170	307	427
21	308	154	270	1.478	461	697
22	211	199	333	1.689	660	1.030
23	181	249	380	1.870	909	1.410
24	199	294	408	2.069	1.203	1.818
25	250	338	420	2.319	1.451	2.238
26	315	377	420	2.634	1.918	2.658
27	404	409	412	3.038	2.327	3.070
28	430	435	398	3.468	2.762	3.468
29	378	452	385	3.846	3.214	3.853
30	324	461	370	4.170	3.675	4.223
31	277	465	356	4.447	4.140	4.579
32	291	462	344	4.738	4.602	4.923
33	318	454	334	5.056	5.056	5.257
34	373	441	326	5.429	5.497	5.583
35	379	426	318	5.808	5.923	5.901
36	320	408	312	6.128	6.331	6.213
37	324	390	308	6.452	6.721	6.521
38	331	369	303	6.783	7.090	6.824
39	287	349	300	7.070	7.439	7.124
40	330	329	296	7.400	7.768	7.420
41	313	309	293	7.713	8.077	7.713
42	268	291	289	7.981	8.368	8.002
43	271	273	286	8.252	8.641	8.288
44	316	257	282	8.568	8.898	8.570
45	267	242	277	8.835	9.140	8.847
46	254	227	272	9.089	9.367	9.119
47	293	214	267	9.382	9.581	9.386
48	273	203	260	9.655	9.784	9.646
49	292	192	254	9.947	9.976	9.900
50	211	182	246	10.158	10.158	10.146
51	224	173	237	10.382	10.331	10.183
52	231	166	228	10.613	10.497	10.611
53	217	159	218	10.830	10.656	10.829
54	207	153	208	11.037	10.809	11.037
55	154	147	196	11.191	10.956	11.233
56	170	142	184	11.361	11.098	11.417
57	153	139	172	11.514	11.237	11.589
58	179	134	157	11.651	11.371	11.746
59	133	131	144	11.784	11.502	11.890
60	111	129	130	11.895	11.631	12.020
61	131	126	114	12.026	11.757	12.134
62	119	125	99	12.145	11.882	12.233
63	84	123	83	12.229	12.005	12.316
64	97	121	67	12.326	12.126	12.383
65	71	120	50	12.397	12.246	12.433
66	52	120	34	12.449	12.366	12.467
67	35	118	17	12.484	12.484	12.484

Grafico I

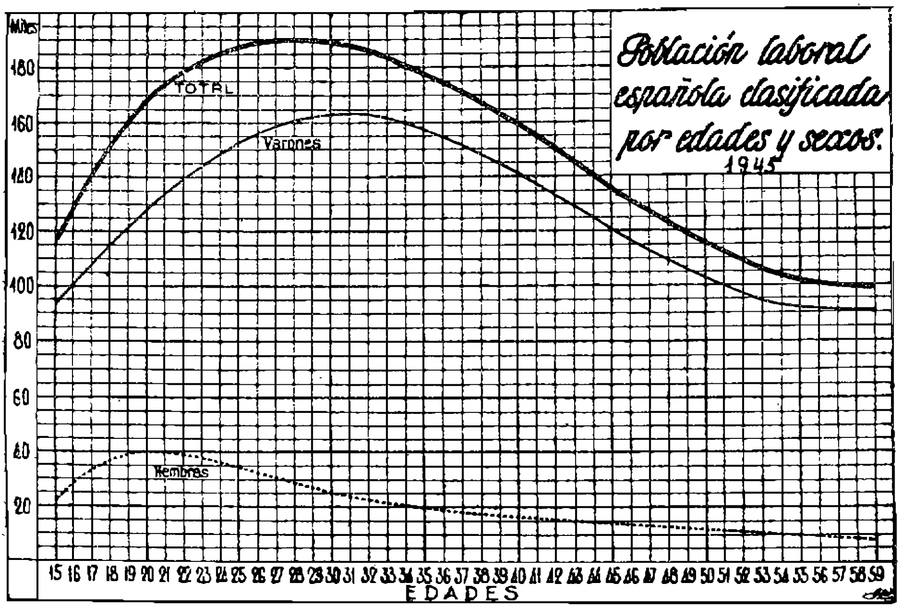


Cuadro II

Serie de varones que integraban la población laboral española, computada al 30 de junio de 1945 sobre el censo de 1940

Edad	Valores de la serie		Edad	Valores de la serie	
	Acumulados	Singulares		Acumulados	Singulares
14	718.969	—	37	4.019.410	152.831
15	812.091	93.122	38	4.169.078	149.668
16	912.628	100.537	39	4.314.013	144.935
17	1.020.153	107.525	40	4.457.138	143.157
18	1.134.652	114.499	41	4.594.376	137.851
19	1.255.948	121.296	42	4.728.736	134.460
20	1.393.519	127.571	43	4.858.297	129.564
21	1.516.842	133.323	44	4.984.125	125.828
22	1.656.353	139.511	45	5.106.178	122.053
23	1.800.910	144.557	46	5.223.556	117.378
24	1.949.751	148.841	47	5.336.369	112.823
25	2.102.289	152.538	48	5.446.483	119.114
26	2.258.653	156.364	49	5.549.600	106.517
27	3.418.000	159.347	50	5.556.263	103.263
28	2.578.313	150.313	51	5.758.053	101.790
29	2.741.078	162.765	52	5.854.316	96.263
30	2.904.065	162.987	53	5.949.200	94.884
31	3.067.875	163.810	54	6.042.563	93.363
32	3.230.754	162.879	55	6.535.329	92.766
33	3.392.589	161.835	56	6.227.211	91.882
34	3.553.587	160.998	57	6.318.771	91.560
35	3.711.097	157.510	58	6.410.161	91.390
36	3.866.579	155.482	59	6.501.360	91.199

Grafico II



Cuadro III

Serie de hembras que integraban la población laboral española, computada al 30 de junio de 1945 sobre el censo de 1940

Edad	Valores de la serie		Edad	Valores de la serie	
	Acumulados	Singulares		Acumulados	Singulares
14	39.387	—	40	769.424	17.052
15	62.076	22.689	41	787.304	16.550
16	90.463	28.387	42	803.335	16.031
17	123.672	33.109	43	818.951	15.616
18	160.473	36.801	44	834.164	15.213
19	199.530	39.057	45	848.864	14.700
20	239.605	40.075	46	863.205	14.341
21	279.354	39.749	47	877.030	13.825
22	318.082	38.928	48	890.438	13.408
23	355.760	37.678	49	903.300	12.862
24	391.760	36.000	50	915.880	12.580
25	426.179	34.419	51	927.896	12.016
26	459.034	32.855	52	939.415	11.519
27	489.359	30.325	53	950.429	11.014
28	517.904	28.545	54	960.924	10.495
29	544.798	26.894	55	970.887	9.963
30	570.231	25.433	56	980.301	9.414
31	594.319	24.088	57	989.155	8.854
32	617.221	22.902	58	997.434	8.279
33	639.077	21.849	59	1.005.075	7.642
34	659.987	20.917	60	1.012.213	7.137
35	680.172	20.185	61	1.018.688	6.475
36	699.427	19.255	62	1.024.545	5.857
37	718.114	18.687	63	1.029.766	5.221
38	736.196	18.082	64	1.034.345	4.579
39	753.702	17.506	65	1.038.273	3.928