Notas sobre cálculo de población probable

ING. FRANCISCO MARDONES O.

Voy a ocuparme ahora en un tema que no he visto tratado antes: «población probable de una localidad en un futuro próximo». Este problema se presenta a menudo en estudios técnicos, particularmente en investigaciones de orden social, económico o de otros aspectos de la ingeniería; y por lo mismo que el cálculo correspondiente se efectúa con alguna frecuencia, no han de encontrarse en esta disertación muchas cosas nuevas; pero sí expuestas de una manera nueva. *Non nove, sed nove.*

1. Demográficamente hablando, el desarrollo de la población del mundo, desligado como lo está, de otros planetas en que exista la vida humana – si los hay – se encuentra regulado solamente por dos factores: la natalidad y la mortalidad. Cualesquiera otros hechos o circunstancias, sean de orden biológico o del orden sociológico, actúan a través de aquéllos.

2. Si nos concretamos a una región determinada del mundo – un continente, un grupo de países, o un país solamente – tendremos que reconocer que en la aceleración o retardo del desarrollo de su población interviene otro factor: el movimiento migratorio, que produce la incorporación de habitantes provenientes de otras regiones, o la extracción de habitantes de la primera que pasan a radicarse en estas últimas.

3. Si nos preocupamos ahora de una región determinada de un país, como una provincia, un departamento, una comuna o una ciudad, por ejemplo, es a todas luces evidente que un nuevo factor introduce su influencia en la marcha del fenómeno a que nos referimos: el movimiento migratorio interno; es decir, el que se produce entre aquella determinada región y el resto del país.

4. Las informaciones disponibles acerca de la manera cómo se ha desarrollado la población, en cualquiera de los tres casos, se encuentran en los censos periódicos.

Carecemos de datos precisos, aun hoy día, con respecto a la población mundial. Sin embargo, las estimaciones para épocas posteriores al siglo XVII, aunque no pueden considerarse rigurosamente exactas, nos dan una noción suficientemente satisfactoria acerca de las tasas de crecimiento total, y aún en los diversos continentes. Las cifras generalmente aceptadas para los años que se indican, son las siguientes:

Población mundial

Años	Millones de habitantes	Crecimiento tasa media anual %	En el período acumulativo medio anual %
1650	545	..	..
1700	..	..	..
1750	728	0,23	0,29
1800	906	0,49	0,44
1850	1,171	0,58	0,51
1900	1,608	0,75	0,63
1940	2,171	0,88	0,75

Estas cifras ponen de manifiesto la tendencia ascendente y acelerada del desarrollo de la población mundial. No hay, pues, indicios de que tienda a estabilizarse o, en otras palabras, que tienda a hacerse asintótica con respecto al atributo tiempo.

Si admitimos una población total de 2,250 millones solamente para el próximo año 1950 (*), resulta que la tasa acumulativa media anual del crecimiento de la población entre el tiempo transcurrido desde el año 1750, vísperas de la revolución industrial hasta ahora, habría sido de 0,55%; cifra bastante elevada si se la compara

(*) El «World Almanac» de 1949 da las siguientes cifras:

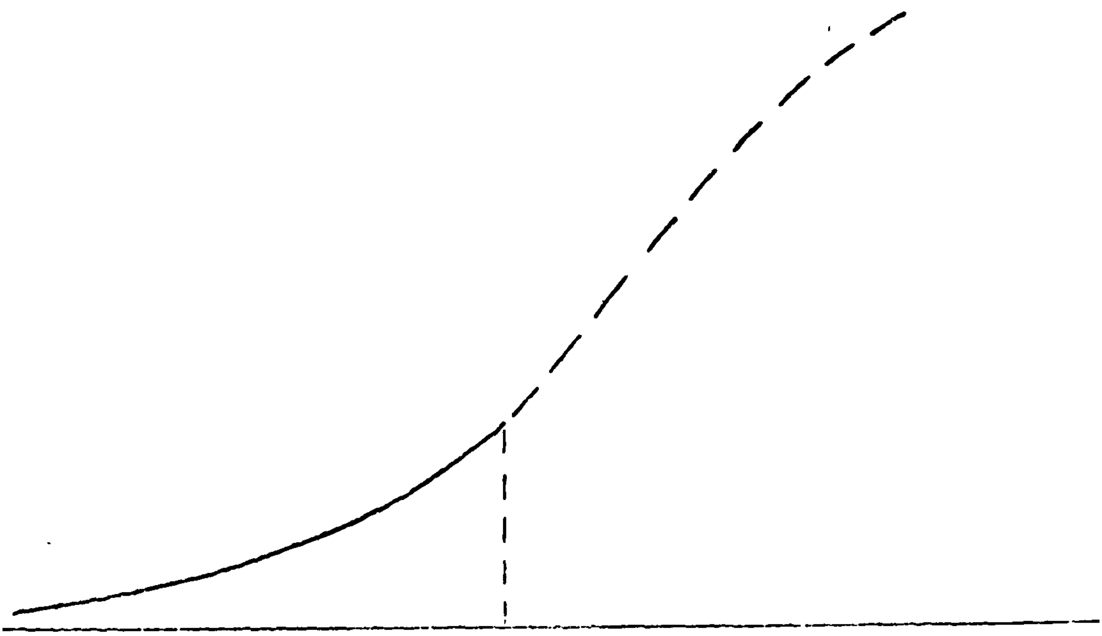
Continentes:	Kms.^{2}	Población
Africa	32 537 747	175 869 488
Asia (a)	25 105 705	1 178 341 250
Europa (a)	4 999 482	380 960 346
URSS. (b)	21 647 017	211 384 985
Norte América (c)	23 382 569	192 656 859
Sud América	17 858 903	101 399 661
Australia y Oceanía (d)	9 440 582	23 796 858
Regiones Antárticas	16 069 709	(?)
Total (e)	151 041 715	2 264 409 447

a) Sin incluir parte de URSS.; *b)* en Europa y Asia; *c)* Incluye México y Naciones del Caribe; *d)* Incluye Australia y Nueva Zelandia, posesiones holandesas, inglesas y francesas y territorios bajo fideicomiso de EE. UU. En muchos casos la población exacta es desconocida y las cifras son más o menos inciertas y a veces hipotéticas. (*e*) El Servicio de Costas y Geodesia de EE. UU. estima en 510 061 110 km^2. el área superficial del planeta, de los cuales 361 121 712 km^2. están ocupados por las aguas y 148 939 398 km^2. por la tierra.

con la que resulta para el millón de años transcurridos desde la primera infracción a una norma de que nos informa la tradición paradisíaca, cifra que apenas excede de dos milésimos del uno por ciento.

Si suponemos ahora que la media acumulativa anual de 0,75% que acusa la penúltima década se mantuviera durante doscientos años, resultaría que en el año 2100 la población mundial se aproximaría a los 7,200 millones.

Los progresos futuros de la técnica ¿podrán ser capaces para asegurar una producción alimenticia que cuantitativa y cualitativamente permita mantener tal población sobre la tierra en condiciones satisfactorias de bienestar humano? Hasta ahora nada nos autoriza para dudar de la eficiencia de la técnica dirigida a tal fin: en otros términos nada nos autoriza para creer que la velocidad $\frac{dP}{dt}$, creciente hasta



ahora, pueda llegar muy pronto a un valor máximo a partir del cual empezara a decrecer, sin que por esto la población tienda a disminuir. Si tal cosa sucediera la representación gráfica del desarrollo de la población a lo largo del tiempo sería la indicada por la figura 1, en la cual las ordenadas son proporcionales a las cifras de población y las abcisas proporcionales a los tiempos respectivos.

Esta curva concuerda perfectamente con el concepto de que la población de nuestro planeta no puede crecer indefinidamente, sino que, por el contrario, sólo podrá llegar a un valor máximo compatible con las condiciones del medio en que vive.

5. Busquemos la ecuación de semejante curva sobre la base de las mismas hipótesis empleadas por quienes la establecieron.

Consideremos una población cerrada; esto es, no sujeta a movimientos migratorios, como sería el caso típico de la población mundial. Las variaciones de esta población dependerían exclusivamente de causas constantes que refundiremos en la expresión: fuerza reproductiva de la especie humana; y de causas variables que frenan la marcha expansionista de ésta, sea por aumento de la mortalidad, sea por atenuación de la natalidad, o por ambas cosas a la vez.

Si consideramos variaciones de la población en intervalos infinitamente pequeños de tiempo, el crecimiento de aquélla sería:

$$dP = Pqdt \tag{1}$$

El factor q es la diferencia entre los coeficientes demográficos de natalidad N y de mortalidad M, ambos dependientes de P y t, dado que la natalidad bruta decrece y la mortalidad crece con el aumento de población. Si designamos por:

n el coeficiente intrínseco de natalidad;
m el coeficiente intrínseco de mortalidad;
η el coeficiente restrictivo de natalidad; y
μ el coeficiente aumentativo de mortalidad;

y admitimos una relación lineal entre estos coeficientes y los coeficientes demográficos de natalidad y mortalidad, tendremos:

$$\begin{aligned} M &= m + \mu P \\ N &= n - \eta P \end{aligned}$$

de modo que:

$$q = N - M = (n - m) - (\eta + \mu)P = \lambda - rP$$

y por consiguiente:

$$\frac{dP}{dt} = P(\lambda - rP) \tag{2}$$

separando las variables y multiplicando por λ, obtenemos:

$$\lambda dt = \frac{\lambda dP}{P(\lambda - rP)} = \frac{\lambda dP + rPdP - rP\,dP}{P(\lambda - rP)} = \frac{dP}{P} + \frac{dP}{\dfrac{\lambda - rP}{r}}$$

integrando:

$$\lambda t = \lg P - \lg \frac{\lambda - rP}{r} + \lg b \tag{3}$$

haciendo en (3) $t = 0$, para el cual $P = P_0$, se obtiene:

$$\lg b = -\lg P_0 + \lg \frac{\lambda - rP_0}{r} = \lg \frac{\lambda - rP_0}{rP_0}$$

$$b = \frac{\lambda - rP_0}{rP_0} \tag{4}$$

Y por lo tanto, introduciendo en (3) el valor de lgb,

$$\lambda t = \lg\left(P\frac{\lambda - rP_0}{r}\right) - \lg\left(P_0\frac{\lambda - rP}{r}\right) \tag{5}$$

Verhulst, profesor de la Escuela Militar de Bruselas, dió a esta ecuación el nombre de función logística (1838) y de logística a la curva representada por ella.

Se ignora el por qué de esta denominación que no tiene nexo distintivo con la etimología del vocablo: Este viene del griego *logisticos*, de *logizesthai* (λογιχή), calcular, que a su vez deriva de *logos* (λογος), razón, estudio. Así, pues, la curva de que se trata es tan logística como cualquiera otra resultante de una teoría o de un razonamiento.

El nombre elegido para esta curva ha tenido el inconveniente de que algunos autores la consideren como la curva típica, o la única lógica para ajustar analíticamente el desenvolvimiento de la población de un país cualesquiera. Veremos, más adelante, que tal interpretación es enteramente injustificada.

Continuamos. Si en las expresiones (4) y (5) hacemos $\frac{\lambda}{r} = K$, tendremos:

$$\lambda t = \lg \frac{P(K - P_0)}{P_0(K - P)}$$

y de aquí:

$$P = \frac{KP_0}{P_0 + \frac{K - P_0}{P_0} e^{-\lambda t}} = \frac{K}{1 + be^{-\lambda t}} \tag{6}$$

El análisis de la ecuación (6) nos proporciona los siguientes pares de valores:

$$t = -\infty; \qquad t = 0; \qquad t = \infty$$

$$P = 0 \quad : \qquad P = \frac{K}{1 + b} \qquad P = K$$

Por otra parte, las tres primeras derivadas de P con respecto a t son:

$$f'(P) = tg\alpha = Kb\frac{\lambda e^{\lambda t}}{(e^{\lambda t} + b)^2} \tag{7}$$

$$f''(P) = Kb\frac{\lambda^2 (b - e^{\lambda t})\, e^{\lambda t}}{(e^{\lambda t} + b)^3} \tag{8}$$

$$f'''(P) = Kb\frac{\lambda^3\left[(e^{\lambda t}-2b)^2-3b^2\right]e^{\lambda t}}{(e^{\lambda t}+b)^4} \qquad (9)$$

Por lo tanto, la logística es asintótica con respecto al eje negativo de los tiempos, ($P \rightarrow O$ cuando $t \rightarrow -\infty$), y con respecto al eje positivo de los tiempos, ($P \rightarrow K$ cuando $t \rightarrow \infty$).

Si hacemos $f''(P) = O$ y designamos por t' el valor de la abcisa para la cual se satisface esta condición, obtendremos:

$$t' = \frac{\lg b}{\lambda \lg e} \qquad (10)$$

Introduciendo este valor en (6) y en (9) obtenemos respectivamente:

$$P = \frac{K}{2} \text{ y } f'''(P) = -\frac{K\lambda^3}{8}$$

Luego, la curva presenta un punto de inflexión de coordenadas:

$$t' = \frac{\lg b}{\lambda \lg e} \text{ y } P' = \frac{K}{2}$$

El valor de $tg\alpha$ en el punto de inflexión se obtiene, asimismo, introduciendo ese valor de t' en $f'(P)$. Resulta:

$$tg\alpha = \frac{K}{4}\lambda$$

La subtangente m, y la subnormal n correspondientes a este punto de inflexión son, respectivamente,

$$m\,P = \frac{dt}{dP} = \frac{2}{\lambda} \text{ y } n = P\frac{dP}{dt} = \frac{K^2}{8}\lambda$$

En suma, pues, la forma de la curva logística presenta un crecimiento $\frac{dP}{dt}$ cada vez más acentuado, desde el punto inicial hasta alcanzar su valor máximo $\frac{K\lambda}{4}$ en el punto de inflexión de coordenadas $t' = \frac{\lg b}{\lambda \lg e}$ y $P' = \frac{K}{2}$. A partir de este punto la velocidad de crecimiento $\frac{dP}{dt}$ se reduce progresivamente hasta anularse, sin que las ordenadas dejen de seguir creciendo, hasta alcanzar un valor máximo finito K en un tiempo infinito ($t = \infty$).

Como se ha visto, la curva se ha deducido sobre la base de tratarse de una población cerrada, y de que en este medio aislado los coeficientes intrínsecos de natalidad (n) y de mortalidad (m) serían constantes a través del tiempo, y de que el propio aumento de la población habría de producir una fase de saturación gradual en la cual las condiciones económicas y otros factores ligados al medio como la insuficiencia de la producción alimenticia, tendrían por consecuencia una declinación de la natalidad (ηP) y un aumento de la mortalidad (μP), proporcionales a la población.

Supongamos que el factor ($\eta + \mu$) fuere despreciable porque fueran pequeñas la restricción de natalidad y el aumento de la mortalidad.

Esta suposición no está lejos de la verdad puesto que si bien las estadísticas acusan alguna restricción de la natalidad en muchos países principalmente en los estratos superiores en cultura, también demuestra una apreciable reducción de los coeficientes de mortalidad, debido, sin duda alguna, a los progresos alcanzados en la salubridad pública y a los no menos importantes conquistados por la ciencia médica. El efecto de esta suposición es el de reducir la ecuación (2) a la siguiente:

$$\frac{dP}{P} = \lambda dt$$

de donde:

$$\lg P = \lambda t + \lg A$$
$$p = Ae^{\lambda t} = Aq^{t}$$

Es decir que en tal caso la población seguiría la conocida ley exponencial.

6. Volvamos a considerar la función (6). Su valor recíproco es:

$$\frac{1}{P} = \frac{1 + be^{-\lambda t}}{K}$$

Tomemos valores de t crecientes en progresión aritmética de razón *m* y escribamos:

$$\frac{1}{P_0} = \frac{1 + be^{-\lambda t_0}}{K}$$

$$\frac{1}{P_1} = \frac{1 + be^{-\lambda(t_0 + m)}}{K}$$

$$\frac{1}{P_2} = \frac{1 + be^{-\lambda(t_0 + 2m)}}{K}$$

Formemos, ahora, las primeras diferencias:

$$\Delta_1 = \frac{e^{-\lambda m} - 1}{K} be^{-\lambda t_0}$$

$$\Delta_2 = \frac{e^{-\lambda m} - 1}{K} be^{-\lambda(t_0 + m)}$$

$$\Delta_3 = \frac{e^{-\lambda m} - 1}{K} be^{-\lambda(t_0 + 2m)}$$

De ellas obtenemos:

$$\frac{\Delta_2}{\Delta_1} = \frac{\Delta_3}{\Delta_2} = \ldots\ldots = e^{-\lambda m}$$

Luego, las diferencias de primer orden de los valores recíprocos de la función logística están en progresión geométrica de razón: $e^{-\lambda m}$.

7. Diversos estadísticos que han estudiado el tema de que tratamos han llegado a establecer, por caminos diversos, aunque similares, formas distintas de esta ecuación que sólo difieren en la contextura de la constante de integración.

a) Si en (6) y (10) hacemos $\lambda = \frac{1}{\alpha}$, tendremos:

$$b = e^{\frac{t'}{\alpha}}$$

$$P = \frac{K}{1 + e^{\frac{t' - t}{\alpha}}} \tag{11}$$

que es la forma conocida como la ecuación de Verhulst.

b) Si hacemos:

$$b = \frac{1}{c} \text{ y } Kc = \pi$$

Se tendrá $P = \dfrac{\pi}{c + e^{-\lambda t}}$ (12)

que es la fórmula de Pearle y Read.

c) Según (10), $b = e^{\lambda t'}$; y si hacemos $e^{\lambda t'} = qe^{\lambda c}$ y $\lambda = \frac{1}{\alpha}$, y substituímos en (6), tendremos:

$$P = \frac{K}{1 + ge^{\frac{(c - t)}{\alpha}}} \tag{13}$$

que es la forma clásica de Yule.

De (13) podemos pasar a (11) introduciendo en lugar de q su valor $be^{\lambda(t'-c)}$ con $\lambda = \frac{1}{\alpha}$

d) Substituyendo en (6) el valor $b = e^{\lambda t'}$ se tendrá:

$$P = \frac{K}{1 + e^{-\lambda(t-t')}} \qquad (14)$$

que es la forma adoptada por Kostintzin.

Las bases sobre las cuales se ha establecido la ecuación de la logística han sido consideradas como aceptables para que tal ecuación pueda emplearse como la curva de ajustamiento más propia para representar la serie de estadísticas de la población mundial. Pero hemos visto al principiar que el desarrollo de la población mundial continúa con su velocidad de crecimiento en aumento; es decir que ésta no ha alcanzado aún a su valor máximo a partir del cual empezaría a decrecer. En otras palabras, no hay concordancia entre dicha función y la marcha observada del fenómeno que ocupa nuestra atención. Dicho de otra manera, los hechos no se han producido en las condiciones que deben esperarse de las hipótesis referidas.

Es oportuno observar que el ritmo lento del desarrollo de la población mundial en los siglos anteriores al XVII ha sido modificado merced a los progresos incesantes de la técnica aplicada a las explotaciones agropecuarias, a la aplicación de métodos cada vez más eficaces y perfectos en el campo de la salubridad pública, y en no pequeña parte, también, al avance de las ciencias básicas de la medicina. Estos factores se han traducido en una baja apreciable y progresiva del coeficiente de mortalidad. Por el contrario, las estadísticas acusan una declinación del coeficiente de natalidad, sobre todo en los países de más avanzada industrialización, y por desgracia también, en sus estratos de mayor cultura, que por otra parte no son los que más contribuyen a acrecentar la población humana. Sería aventurado formular apreciaciones acerca de la influencia que estos hechos han tenido en el modo como han fluctuado los coeficientes η y μ. Pero es indiscutible, en suma, que el coeficiente restrictivo r de la ecuación (2), no ha adquirido importancia bastante para impedir que la velocidad de crecimiento de la población mundial continúe en su marcha acelerada. Ahora bien, fundándonos en que aún quedan apreciables porciones del planeta sin explotar o sólo rudimentariamente explotadas, y en que no existen signos de haberse agotado la capacidad inventiva del hombre, no es aventurado suponer que ese ritmo ha de continuar por largo tiempo, a pesar de las voces de alarma que periódicamente previenen al mundo contra el *replete terram* del precepto bíblico.

Veremos más adelante que la función logística ha tenido amplia aplicación en el ajustamiento de series cronológicas de poblaciones que no cumplen con las bases de su establecimiento, porque ni son cerradas, lo que implica, por una parte, que los medios de subsistencia de que disponen no están limitados a los que pueden producirse en su propio territorio y, por otra parte, que la población recibe el aporte de individuos foráneos y los entrega a otros centros poblados. La ha tenido, y muy propiamente, en el ajustamiento de series cronológicas de otra índole, como las relativas al desarrollo de organismos unicelulares, al de especies inferiores sometidas a una determinada provisión de alimentos, al ritmo de la producción de ciertos bienes que tienden a colmar la necesidad que están destinados a satisfacer, etc.

8. El desarrollo de la población dentro de una región determinada, como el territorio de una nación, por ejemplo, está sometido a la influencia de factores que actúan en el sentido de acelerar su crecimiento y de otros que tienden a restringirlo. No es del caso ocuparse en esta ocasión de analizar esos factores de orden económico, social o jurídico, ni de valorar la influencia positiva o negativa que cada uno de ellos ejerce en el fenómeno de que se trata. Basta recordar que ellos existen y que, como ya se dijo, unos estimulan y otros restringen ese crecimiento. Esas influencias se reflejan en las cifras de natalidad e inmigración, por una parte, y en las de mortalidad y emigración por otra; todas las cuales participan en los resultados totales de los censos periódicos.

De un modo semejante, el desenvolvimiento o marcha evolutiva de la población dentro de determinada zona de un país está regido por los mismos factores enunciados y, además, como ya se dijo, por los movimientos migratorios internos que desplazan habitantes de unas regiones a otras del mismo país.

El estudio de los métodos estimativos de la población probable en un futuro próximo no necesita distinguir entre región del planeta o región de un país. Nos referiremos, pues, indistintamente, a unos y otros.

9. La constatación de cierta antinomia entre la virtud procreadora de la especie humana y la virtud germinativa de la región en que vive, es de muy remota antigüedad, pues ya aparece enunciada por Botero (1). Cerca de dos siglos más tarde, en 1775, Juan María Ortes (2) admite que el crecimiento natural o vegetativo sigue una progresión geométrica que duplica la población cada 30 años, sino es contrariada por la presión de las subsistencias que crecen más lentamente.

Veinte años después (1798) Malthus (3) en su «Essay on the principle of Population», expuso su teoría sobre el desenvolvimiento de la población, según la cual el instinto fisiológico de procreación ejercido sin cortapisas, haría crecer aquélla en progresión geométrica que la duplicaría cada 25 años; en tanto que la producción de las subsistencias sólo crecería en progresión aritmética. No entra en mi programa analizar los dos factores de la teoría de Malthus, contrariada por la experiencia que ha demostrado que ni la población del mundo ha crecido con la rapidez que él imaginó, ni la producción de subsistencias ha seguido la marcha lenta que él pronosticó. Pero es lo cierto que desde entonces, ha venido aflorando periódicamente la preocupación de individuos y de organismos sociales por las consecuencias de un exceso de habitantes sobre la capacidad del planeta para entregar a la actividad humana los medios de subsistencias que ella le exija.

Poco más de un tercio de siglo después que Malthus, en 1835, Quetelet (1) en su «Física Social», trató de comparar el fenómeno demográfico con un fenómeno físico, asemejándolo a un móvil que se desplazaría en un medio resistente. La propia velo-

---

(1) JUAN BOTERO: «Delle Ragion di Stato», 1583.

(2) JUAN MARÍA ORTES: «Riflessioni sulla populazioni delle nazioni», 1775.

(3) TOMAS ROBERTO MALTHUS.

(1) L. ADOLFO J. QUETELET, matemático y sociólogo belga. Profesor de Matemáticas y Astronomía en la Academia Militar de Bruselas y Presidente de la Comisión Central de Estadística de Bélgica.

cidad de crecimiento de la población provocaría una restricción de ésta que tendría su origen en la deficiencia progresiva de la alimentación. Tal restricción sería proporcional al cuadrado de aquella velocidad.

Muy poco más tarde, Verhulst (2) en sus «Notas sobre la ley que sigue la población en su acrecentamiento», publicadas en «Correspondencia Matemática y Física», se ocupó, según se ha escrito a instancias del propio Quetelet, de la formulación matemática del problema en una forma similar a la que ya hemos expuesto.

Cerca de un siglo después (1920) la función logística de Verhulst, olvidada, fué redescubierta por los señores R. Pearle y L. J. Read (estadísticos norteamercianos), deduciéndola de experiencias de laboratorio sobre un universo de moscas Drosophilas confinadas en un medio aislado; y aplicaron sus deducciones al estudio de la tasa de crecimiento de la población de Estados Unidos entre 1790 y 1920.

Más tarde publicaron su *On the mathematical theory of population* (1933). Pearle y Read formularon ciertas premisas según las cuales las poblaciones se desenvolvían por ciclos de amplitud variable. La tasa de crecimiento absoluto principiaba con un valor pequeñísimo en cada uno de estos ciclos y aumentaba, progresivamente, hasta un máximo que correspondería a la relación óptima entre el territorio y la población. A partir de este valor decrecía hasta alcanzar un período de saturación, en que la población permanecía prácticamente estacionaria. La representación matemática de esa evolución exige una curva convexa con respecto al eje de los tiempos y asintótica con respecto a éste, que pase por un punto de inflexión en el cual inicie el cambio de su convexidad en concavidad y siga, hasta hacerse asintótica con respecto a una paralela al eje de los tiempos, lo que limitaría superiormente el crecimiento de la población.

Sobre la base de estos postulados redescubrieron, como lo hemos dicho, la logística de Verhulst. Posteriormente generalizaron su concepción en dos aspectos diferentes. Considerando la posibilidad de una sucesión de ciclos, admitieron que la asíntota inferior fuese paralela al eje de los tiempos con una ordenada finita que correspondería aproximadamente al límite superior de población en ciclo precedente.

De aquí la expresión:

$$P = d + \frac{K}{1 + be^{-\lambda t}} \tag{15}$$

contándose ahora K a partir de la asíntota de ordenada d;

En seguida consideraron el caso más general en que el exponente del término exponencial fuese substituído por una expresión polimonial en t.

Entonces:

$$P = d + \frac{k}{1 + be^{-(a_1 t + a_2 t^2 + \ldots\ldots a_n t^n)}}$$

De este modo se puede conseguir una mejor adaptabilidad a los datos de la observación; a expensas, es cierto, de una mayor dificultad de cálculo por la intro-

(2) Pedro Francisco Verhulst, matemático belga. Profesor de Matemáticas en la Academia Militar de Bruselas.

ducción de nuevas constantes, las cuales, por lo demás, no acusan relación con la biometría ni con los movimientos migratorios.

Puesto que nuestro intento no va más allá de estudiar métodos de cálculo adecuados para obtener cifras de población en futuros más o menos próximos, y de ningún modo a buscar expresiones matemáticas descriptivas del fenómeno a lo largo de grandes extensiones de tiempo, no es del caso entrar en mayores consideraciones sobre la función logística y su verdadero significado.

10. Sabemos que la población no se encuentra uniformemente repartida en el mundo y que la irregularidad de esta distribución está relacionada más o menos directamente con las diferencias de aptitud que presentan las distintas extensiones geográficas para ser habitadas. Factores como el clima, la situación geográfica, el relieve del suelo; la existencia de aguas, de minerales, de tierras aptas para la explotación agropecuaria, hacen que una región sea más ventajosa que otra para que el hombre prefiera radicarse en ella y aplicar allí los métodos que la humanidad ha creado y perfeccionado a través de los tiempos para obtener la mejor utilización de los recursos naturales.

De aquí ha resultado la desuniforme distribución de los habitantes del planeta, caracterizada por el hecho de que poco menos del 17% de la población mundial reside en la zona europea, que cubre poco más del 3% de la extensión de tierra firme.

Ninguno de los Estados europeos, ni la Europa en su conjunto, dispone de la extensión y recursos naturales indispensables para sostener la población que su elevado progreso cultural e industrial ha radicado dentro de sus límites territoriales. Y es de notar, por otra parte, que cada una de las agrupaciones de población está dominada por un concepto nacionalista que importaría un verdadero aislamiento con respecto a las vecinas, si la necesidad de intercambiar productos no les indujera a convivir con las otras.

Más todavía, pueden señalarse tres regiones geográficas suficientemente extensas y dotadas de todos los elementos necesarios para que su población pudiera sostenerse independientemente de cualquiera otra región: Norte América, Sud América y Rusia. Sin embargo, ni la mejor dotada de estas tres regiones de diversa intensidad de población y de distinto grado de industrialización, vive o puede vivir aislada del resto del mundo.

Las circunstancias que señalamos, provocan desplazamientos de población de unas regiones a otras, más o menos continuos o esporádicos, que influyen considerablemente en el desarrollo de las poblaciones afectadas. Esta influencia no se puede medir sólo por el número de personas que entran y salen del territorio en determinados períodos de tiempo; pues en ella participa muy particularmente la composición de dichos conjuntos de personas, según el sexo y la edad de éstas, y su afinidad étnica con la población a que se incorporan. Se ha tratado de avaluar la contribución cuantitativa de la corriente migratoria descomponiéndola en dos términos; el uno el aporte directo; estaría representado por el saldo migratorio positivo, y el otro, la contribución indirecta, consistiría en el acrecentamiento de la natalidad producido en el país por la corriente inmigratoria. Así, por ejemplo, sobre la base de que las tasas anuales de crecimiento inmigratorio y natural mantuviesen entre sí la misma relación que el exceso de inmigrados sobre emigrados guarda con el excedente de nacimientos sobre defunciones, se ha estimado que el aumento total de 200 millones de habitantes verificado en América durante el siglo 1840...1940, se descompone en cifras redondas, como sigue (*):

Aumento natural, independiente de la inmigración...	71%
Aumento natural, proveniente de la inmigración....	13 »
Aumento por saldo migratorio....................	16 »
Total	100%

Sería aventurado formular hipótesis aceptables para introducir en fórmulas, coeficientes que tradujeran la influencia del movimiento migratorio en el desarrollo de la población. Esa influencia no obedece a causas que actúan de un modo continuo a través del tiempo ni varían según funciones determinadas de la población.

Una consecuencia directa e inmediata de lo que hemos dicho es que ninguna curva que se base en la facultad procreadora de la especie es aplicable por este sólo hecho al desarrollo de la población de una región del planeta, de un país o de una parte de él. Podrá serlo, sin embargo, porque su forma concuerde con la tendencia general que manifieste la serie cronológica de los datos disponibles, y por tal razón esta clase de curvas podrá ser utilizada como cualesquiera otras para la descripción del importante fenómeno social de que tratamos.

11. Alguna vez he leído que si se pudiera disponer de un aparato que registre la curva ininterrumpida de las variaciones de la población dentro de una zona cualquiera a través del tiempo, se obtendría una gráfica semejante a la que dibujan otros aparatos inscriptores. Así como un termógrafo, por ejemplo, registra las variaciones de la temperatura en función del tiempo, inscribiendo el número de grados celsius del ambiente en cada instante, nuestro hipotético inscriptor registraría a su turno, el número de habitantes existentes en cada instante. El problema consiste, entonces, en encontrar la función que substituya esta gráfica y que pueda admitirse como descriptiva del fenómeno registrada por ella.

Esta curva nos permitirá conocer por interpolación el valor probable de la población en años intermedios, y por extrapolación, el valor probable de la población en años anteriores o posteriores, más o menos próximos a los extremos de la serie. No se trata, como se ve, de pronosticar un hecho cierto sino que de anunciar un hecho que será probable, siempre que se mantengan las condiciones que produjeron aquella gráfica y, por consiguiente, aquella función matemática.

Los datos disponibles son los que porporcionan los censos sucesivos, cada uno de los cuales nos suministra un sólo punto de aquella gráfica. Entre todos ellos forman una sucesión cronológica de valores observados que representan la verdad estadística con relativa aproximación, porque no existen censos rigurosamente exactos. Todos ellos adolecen de errores voluntarios o involuntarios; aunque estos errores puedan considerarse despreciables en los países de gran cultura en los cuales reina una alta disciplina civil.

Si se hace la representación geométrica de tal sucesión o serie estadística tomando sus valores como ordenadas y los tiempos como abscisas, y unimos uno a uno

(*) MORTARA. «Revista de Economía y Estadística». Año IV-1942, Buenos Aires.

los puntos superiores sucesivos, obtendremos una poligonal. Esta representación gráfica de los datos en que resaltan las variaciones absolutas o diferencias entre los valores de la serie en los años sucesivos, permite una rápida apreciación de los caracteres generales de la tendencia. Esta apreciación puede afinarse considerablemente substituyendo dicha poligonal por una curva continua que pase por nuevos puntos obtenidos por el método de los promedios móviles. El juicio que nos formemos acerca de la forma de dicha curva nos servirá de guía para elegir la clase o tipo de función más adecuada para representar la serie de que se trata. La operación que consiste en buscar la ecuación de la línea continua que pase por dichos puntos o determine la función que se satisfaga con todos los pares de valores sucesivos, constituye una interpolación propiamente dicha. Muchas líneas pueden cumplir en cada caso con tal condición, y la indeterminación se evita mediante el empleo de polinomios de interpolación. Estos polinomios pueden considerarse como expresiones matemáticas descriptivas de los datos experimentales y tienen el carácter de leyes empíricas de estos últimos.

La determinación de los parámetros de los polinomios de interpolación requiere gran trabajo operatorio; tanto más grande cuanto mayor es el número de datos de la serie, puesto que el número de sus términos excede en uno al número de puntos disponibles de aquélla. Por otra parte, estos datos, como ya lo hemos dicho, adolecen de errores, que ni siquiera obedecen a las leyes del azar. Así, pues, no hay razones valederas para preferir un polinomio de interpolación a una curva de ajustamiento como función descriptiva del fenómeno de que se trata. Obsérvese que el primero describe la condición en que se presentan los datos experimentales, mientras que la segunda describe la tendencia del fenómeno de que ellos son expresiones sucesivas a lo largo del atributo tiempo. La curva de ajustamiento debe acercarse en lo posible a los puntos experimentales y para obtener tal resultado se la somete a la condición de hacer mínima la suma de los cuadrados de las diferencias entre sus ordenadas y las correspondientes de la sucesión experimental, lo que implica hacer nula la suma de las mencionadas diferencias.

La condición indicada puede cumplirse por curvas de distintos tipos o familias, y entre dos de distinto tipo, la que con mayor probabilidad represente dicha tendencia será la que produzca para los cuadrados de las diferencias la suma mínima..Esta afirmación no obliga a extremar la prolijidad en la búsqueda más allá de límites que tomen en consideración que los resultados de los censos no son rigurosamente exactos, como ya se ha dicho. Así, en el caso de que tratamos bastará comparar la media aritmética de los valores absolutos de dichas diferencias con el valor medio de las ordenadas experimentales, lo que equivale a comparar la suma absoluta de dichas diferencias con la suma total de las ordenadas.

Por lo general esta comparación se expresa en tanto por ciento:

$$V = 100 \frac{\Sigma |(Z_i - P_i)|}{\Sigma P_i}$$

12. Apliquémonos a buscar la curva de ajustamiento más adecuada para describir el desarrollo de la población de Chile.

La tabla I contiene las series cronológicas por considerar:

TABLA I.—POBLACION DE CHILE

Años	Miles de habitantes							
c1	c2	h	c3	h	c4	h	*ca*	h
1870	....	....	....		1944		2087	
1875	2076	....	2219					
1880	....	....	....		2282	17,3	2342	12,2
1885	2507	20,8	2472	11,4				
1890	....	....	....		2600	13,9	2593	10,7
1895	2696	7,5	2720	10,0				
1900	....	....	....		2907	11,8	2911	12,2
1907	3231	16,5	3210	15,0				
1910	....	....	....		3340	14,9	3312	13,8
1920	3732	11,9	3710	11,9	3732	11,7	3710	12,0
1930	4287	14,9	4287	15,6	4287	14,9	4287	15,6
1940	5024	17,1	4887	13,9	5024	17,1	4887	13,9
	23552		23505		26116		26128	
Promedio		13,9		13		13,7		12,9

Las cifras incluídas en las diversas columnas de esta tabla tienen el siguiente origen:

c2: los resultados de los censos de 1875 adelante, como aparecen en la «Estadística Chilena» de octubre de 1941 (pág. 450), publicada por la Dirección General del ramo.

c3: los datos de la columna 2 modificados con las correcciones que se expresan:

a) Las indicadas por el ex Director don Roberto Vergara H. en su estudio «Los Censos de Población de Chile» presentado al VIII Congreso Científico Pan Americano en cuanto se refiere a los censos de 1875 hasta el de 1920. La mayor parte de estas correcciones tienen su origen en la necesidad de hacer homogénea la serie, como son las que provienen de los cambios territoriales producidos y del hecho de no haber comprendido los censos anteriores a 1907 la población indígena radicada en la parte sur del país. Las demás provienen de la verificación de errores tanto positivos como negativos, de no mucha importancia en relación con la población total y con el grado de aproximación a la verdad de los censos.

b) La que afecta al censo de 1940 está justificada por la información publicada por el ex Director don Emilio Rodríguez Mendoza en «El Mercurio», de Santiago, del día 17 de marzo de 1949. Según esta información el Director propuso introducir un aumento de 5% sobre el resultado final de dicho censo; le fué aceptado un 3%, que en definitiva se redujo a 2,8% (sic.).

c4: la serie de cifras de población en los años terminados en cero, constatadas por los censos o calculadas por interpolación geométrica entre los datos corregidos de los censos vecinos; tomados de c3.

Se observará que en el lapso considerado disponemos de tres censos, los primeros que se efectuaron en años terminados en cinco y los tres últimos en años terminados en cero; de modo que para formar la serie de datos equiespaciados, podríamos haber elegido una u otra terminación. Se prefirió la última en atención a los antecedentes que se expresan en seguida:

La ley de 8 de julio de 1853 dispone que el Censo General de la República se formará cada diez años y como el primero se efectuó en 1855 se debería deducir que los censos deben efectuarse en todos los años terminados en 5. En la práctica sólo fué así hasta 1895 y el intervalo decenal de la serie se interrumpió con la postergación del censo de 1905 al año 1907 y se restableció, en seguida, con los censos de los años 1920, 30 y 40; por lo que imaginamos que esta serie no se ha de interrumpir nuevamente. Esta hipótesis está, además, abonada por resoluciones internacionales que recomiendan unas el levantamiento decenal de los censos de población, y otras particularmente efectuar el censo en todas las Repúblicas Americanas el año 1950.

Las tasas h de crecimiento medio anual en los períodos sucesivos oscilan, en todas estas series, alrededor de una media poco diferente de 13°/₀₀. Si, por otra parte, se colocan estas series en una gráfica se observa que todas ellas oscilan hacia uno y otro lado de una tendencia creciente acelerada, lo que indica como satisfactorias para describir el desarrollo de la población chilena, aquellas curvas en que la relación $\frac{dP}{dt}$ sea también creciente; es decir, curvas como la exponencial: $P = P_0q^t$, o la función de Gompertz: $P = Ka^{b^t}$

13. Se observará que escribimos la función exponencial $P = Aq^t$. Expliquemos por qué. Si consideramos variaciones de población en intervalos infinitamente pequeños de tiempo, el acrecentamiento sería:

$$dP = P\alpha dt$$

de donde el acrecentamiento relativo:

$$\frac{dP}{P} = \alpha dt$$

Integrando:

$$\text{Ln}\, P = \alpha t + \text{Ln}\, A$$

y por consiguiente:

$$P = Ae^{\alpha t}$$

El coeficiente α ha recibido el nombre de «coeficiente de procreación biológica pura».

Ahora bien, cuando se trata de adquirir conocimientos acerca de la población en años intermedios en que no ha habido censos o en años anteriores o posteriores a los que son límites de la serie, nos basamos, como ya se ha dicho, en los resultados obtenidos en los censos efectuados. Cada uno de éstos proporciona, salvo errores inevitables, una cifra que debería ser la del censo anterior incrementada en las diferencias producidas en el intervalo de tiempo, entre nacimientos y defunciones (crecimiento natural) y entre pasajeros entrados y salidos del territorio (saldo migratorio). El coeficiente q tiene, pues, un significado distinto del que se obtuvo al considerar variaciones de población a intervalos infinitamente pequeños. Es, pues, más realista o conforme con la efectividad de los hechos, adoptar, en lugar del coeficiente e^{α}, el coeficiente $q = 1 + r$, en que r es la tasa de crecimiento acumulativo periódico.

La comparación entre las ecuaciones $P_i = Aq^t$ y $P_i = Ae^{\alpha t}$ nos demuestra que:

$$q = 1 + r = e^{\alpha}$$

De donde:

$$\alpha = Lq = L(1 + r) = r - \frac{r^2}{2} + \frac{r^3}{3} - \frac{r^4}{4} - \ldots\ldots$$

Resulta que el coeficiente α diferirá tanto menos de r cuanto menor sea el valor de este último.

En suma, pues, dentro de los valores que tiene r en el caso que estudiamos, la cuestión examinada es más bien del orden de la lógica en la elección de los coeficientes que de exactitud matemática.

14. Elegida la función $P = Aq^t$, corresponde determinar los valores de los coeficientes A y q. El método más adecuado para obtenerlos es el de transformar aquella ecuación en la de una recta mediante la anamorfosis logarítmica. Así tendremos:

$$\log P_i = \lg A + t_i \lg q$$

Aplicando el método de los cuadrados menores dispondremos de la ecuación de condición:

$$\Sigma (\lg A + t_i \lg q - \lg P_i)^2 = \text{mínima}$$

y por lo tanto, las ecuaciones normales serán:

$$2\,\Sigma (\lg A + t_i \lg q - \lg P_i) = 0$$
$$2\,\Sigma t_i (\lg A + t_i \lg q - \lg P_i) = 0$$

De aquí:

$$\lg q = \frac{n \Sigma t_i \lg P_i - \Sigma t_i \Sigma \lg P_i}{n \Sigma t_i^2 - (\Sigma ti)^2} \qquad (16)$$

$$\lg A = \frac{\Sigma t_i^2 \Sigma \lg P_i - \Sigma t_i \Sigma t_i \lg P_i}{n \Sigma t_i^2 - (\Sigma ti)^2} \qquad (17)$$

Encontrados los valores de los coeficientes q y A, podemos calcular los valores sucesivos Z_i y comparar la serie analítica con la experimental. Encontraremos siempre que la suma de las diferencias $Z_i - P_i$ no es nula, lo que proviene (aparte de las epqueñas diferencias que tienen su origen en la aproximación con que se calculan A y q), del hecho de que la condición de hacer mínima la suma de los cuadrados entre las diferencias de los logaritmos no hace mínima la suma de los cuadrados de las diferencias entre los números correspondientes.

Esto indica que los parámetros A y q encontrados deberían considerarse sólo como una primera aproximación que puede ser o no suficiente según el valor que resulte para $v = 100 \frac{\Sigma|(Z_i - P_i)|}{\Sigma P_i}$ y para $\Sigma Z_i - \Sigma P_i$.

15. Si fuere necesario, se pueden calcular las correcciones α y β.

La nueva ecuación será:

$$P_i = (A + \alpha)(q + \beta)^t \tag{18}$$

$$LP_i = L(A + \alpha) + tL(q + \beta)$$

$$LP_i = LA\left(1 + \frac{\alpha}{A}\right) + t\,Lq\left(1 + \frac{\beta}{q}\right)$$

Desarrollando en serie y despreciando los términos de grado superior al primero:

$$LP_i = LA + \frac{\alpha}{A} + tLq + t\frac{\beta}{q}$$

$$LP_i = LZ_i + \frac{\alpha}{A} + t\frac{\beta}{q}$$

La ecuación de condición será ahora:

$$\sum\left[LZ_i + \frac{\alpha}{A} + t_i\frac{\beta}{q} - LP_i\right]^2 = \text{mínima}$$

Por consiguiente, las ecuaciones normales serán:

$$\frac{n}{A^2}\alpha + \frac{\beta}{qA}\sum t_i = \frac{1}{A}\left[\Sigma\,\mathrm{lg}P_i - \Sigma\,\mathrm{lg}Z_i\right]$$

$$\frac{1}{Aq}\alpha\sum t_i + \frac{1}{q^2}\beta\sum t_i^2 = \frac{1}{q}\left[\Sigma\, t_i\, LP_i - \Sigma\, t_i\, LZ_i\right]$$

De aquí:

$$\alpha = \frac{A\left[\Sigma\, t_i^2\,(\Sigma\, LP_i - \Sigma\, LZ_i) - (\Sigma\, t_i LP_i - \Sigma t_i\, LZ_i)\right]}{n\,\Sigma\, t_i^2 - (\Sigma\, t_i)^2}$$

$$\beta = \frac{n(\Sigma t_i LP_i - \Sigma t_i LZ_i) - \Sigma t_i(\Sigma LP_i - \Sigma LZ_i)}{n\Sigma t_i^2 - (\Sigma t_i)^2}$$

16. Se puede también proceder como sigue:

Escribamos la ecuación (18) así:

$$P_i = (A+\alpha)\cdot q^t\left(1+\frac{\beta}{q}\right)^t$$

Desarrollando en serie y despreciando los términos de grado superior al primero:

$$P_i = (A+\alpha)\, q^{t_i}\left(1+t_i\frac{\beta}{q}\right)$$

$$P_i = Aq^{t_i} + \alpha q^{t_i} + \frac{At_i q^{t_i}}{q}\beta = Z_i + \alpha q^{t_i} + \frac{A}{q}t_i q^{t_i}\beta$$

La ecuación de condición será:

$$\sum\left[\alpha q^{t_i} + \frac{A}{q}t_i q^{t_i}\beta + Z_i - P_i\right]^2 = \text{mínima.}$$

Hagamos $\frac{A}{q} = Q$ y $P_i - Z_i = X_i$

$$\Sigma[\alpha q^{t_i} + Q\beta t_i q^{t_i} - X_i]^2 = \text{mínima}$$

Por consiguiente:

$$\alpha\Sigma q^{2t_i} + Q\beta\Sigma q^{t_i}t_i = \Sigma q^{t_i}X_i$$

$$\alpha\Sigma t_i q^{t_i} + Q\beta\Sigma t_i^2 q^{2t_i} = \Sigma t_i X_i q^{t_i}$$

De donde:

$$\alpha = \frac{\Sigma q^{t_i}(X_i\cdot\Sigma t_i^2 q^{2t_i} - \Sigma t_i q^{t_i}\Sigma t_i X_i q^{t_i}}{\Sigma q^{2t_i}\Sigma q_i^2\cdot q^{2t_i} - (\Sigma t_i q^{2t_i})^2}$$

$$\beta = \frac{q}{A}\,\frac{\Sigma t_i X_i q^{t_i}\Sigma q^{2t_i} - \Sigma q^{t_i}(X_i\Sigma t_i q^{2t_i}}{\Sigma q^{2t_i}\Sigma t_i^2 q^{2t_i} - (\Sigma t_{iq}^{2t_i})^2}$$

17. Si se prefiere no despreciar el término en $\alpha\,\beta$ se procederá como sigue:

$$P_i = Aq^t\left(1+\frac{\alpha}{A}\right)\left(1+\frac{\beta}{q}t\right) = VZ_i(1+Bt)$$

Por consiguiente, la ecuación de condición será:

$$\Sigma\left[VZ_i(1+Bt)-P_i\right]^2=\text{min.}$$

Y las ecuaciones normales:

$$V\Sigma Z_i^2(1+Bt)^2=\Sigma Z_iP_i(1+Bt)\ldots \tag{19}$$

$$V^2\Sigma Z_i^2t(1+Bt)=V\Sigma Z_iP_it\ldots \tag{20}$$

De las cuales se obtiene:

$$\alpha=A\frac{\Sigma t_iZ_i(\Sigma t_iZ_i-\Sigma t_iZ_iP_i)-\Sigma t_i^2Z_i^2(\Sigma Z_i^2-\Sigma Z_iP_i)}{\Sigma Z_i^2\Sigma Z_i^2P_i^2-(\Sigma t_iZ_i^2)^2}$$

$$\beta=q\frac{\Sigma X_i^2\Sigma t_iZ_iP_i-\Sigma t_iZ_i^2\Sigma Z_iP_i}{\Sigma t_i^2Z_i^2t_i\Sigma Z_iP_i-\Sigma t_iZ_i^2\Sigma t_iZ_iP_i}$$

18. Volvamos a las cuatro series cronológicas de la población de Chile. El cálculo de los coeficientes A y q, para cada una de estas series, con las expresiones (16) y (17), nos conduce a las siguientes ecuaciones:

Serie C.2) $Z_i = 2\,114 \cdot 1{,}013\,133^t$
Serie C.3) $Z_i = 2\,179 \cdot 1{,}012\,269^t$
Serie C.4) $Z_i = 1\,971 \cdot 1{,}013\,21^t$
Serie C.5) $Z_i = 2\,056 \cdot 1{,}012\,13^t$

Los valores de Z_i así como las diferencias $Z_i - P_i$ se consignan en la tabla II para las series C.2 y C.3, y en la tabla III para las series C.4 y C.5. En estas mismas tablas se indican los valores de $v = 100\dfrac{\Sigma\,|(Z_i - P_i)|}{\Sigma P_i}$

Ninguno de ellos alcanza al 2%, lo que indica, desde luego, lo innecesario que sería ajustar cualquiera de estas series con curvas de otro tipo. Por otra aparte, las diferencias algebraicas $\Sigma Z_i - \Sigma P_i$, son tan pequeñas que es también innecesario corregir los coeficientes A y q encontrados.

19. Sin embargo, por simple tanteo hemos obtenido la ecuación

$$Z_i = 2\,061 \times 1{,}012\,136^t$$

que proporciona los valores que se encuentran en la tercera columna de la tabla IV. Si bien se conserva el mismo el valor de V = 1,27, la curva se ha acercado a los datos experimentales puesto que la diferencia $\Sigma Z_i - \Sigma P_i$ ha descendido de −75 a −8.

En la séptima columna de la misma tabla IV se encuentran los valores obtenidos con la ecuación de Gompertz:

$$Z_i = Ka^{b^t}$$

El método de ajustamiento con esta función y el cálculo de los coeficientes respectivos es el siguiente.

20. La anamorfosis logarítmica nos permite transformar la ecuación de Gompertz en la recta:

$$\lg P_i = \lg K + b^t \lg a$$

Elijamos tres puntos equiespaciados, y sea m la distancia entre ordenadas, y escribamos:

$$\lg P_1 = \lg K + b^{t_1} \lg a$$
$$\lg P_2 = \lg K + b^{t_1 + m} \lg a$$
$$\lg P_3 = \lg K + b^{t_1 + 2m} \lg a$$

De estas tres relaciones deducimos:

$$\Delta_1 = \lg P_2 - \lg P_1 = b^{t_1} (b^m - 1) \lg a$$

$$\Delta_2 = \lg P_3 - \lg P_2 = b^{t_1 + m} (b^m - 1) \lg a$$

$$\frac{\Delta_2}{\Delta_1} = b^m$$

$$\lg a = \frac{\Delta_1}{b^{t_1} (b^m - 1)}$$

$$\lg K = \lg P_1 - b^{t_1} \lg a$$

Sean los tres pares de valores obtenidos como medias geométricas entre los de abcisas vecinas:

$P_i =$	2 211 ,	3 105 ,	4 577
$t =$	5	35	65

Con estos valores obtenemos:

$$b^m = 1{,}164\,079$$

$\lg b = 0\,002\,189$	;	$b = 1{,}005\,053$
$\lg a = 0\,885\,318$	;	$a = 7{,}679$
$\lg K = 2{,}434\,191$	;	$K = 271{,}763$

(1) Definir características de la curva.

(2) Tomemos nota de que las diferencias de primer orden de los logaritmos de las ordenadas equiespaciadas de la curva de Gompertz, están en progresión geométrica de razón b^m.

Se observa que esta curva baja considerablemente la $\Sigma \mid (Z_i - P_i) \mid$, pero hace crecer la diferencia $\Sigma Z_i - \Sigma P_i$, de 8 a 44, lo que no la recomienda como aceptable. Pero si afinamos los coeficientes por tanteo, o por aproximaciones sucesivas, llegamos a los valores:

$$\begin{aligned} \lg b &= 0{,}002166 \quad ; \quad b = 1{,}005 \\ \lg a &= 0{,}885361 \quad ; \quad a = 7{,}68 \\ \lg k &= \qquad\qquad\quad\;\; k = 272{,}863 \end{aligned}$$

Ahora los valores de Z_i que se insertan en la 7.ª columna de la Tabla IV nos dan:

$$\Sigma Z_i - \Sigma P_i + 1$$

$$\Sigma (\mid Z_i - P_i \mid) = 121 \qquad\qquad 100 \frac{\Sigma (\mid Z_i - P_i \mid)}{\Sigma P_i} = 046$$

lo que nos dice que la función de Gompertz con los coeficientes indicados describe mejor que la exponencial el desarrollo de la población de Chile representado por la serie C.5 de la Tabla I.

TABLA II.—POBLACION DE CHILE: SERIES C.2 Y C.3

Años		Serie C.2 Miles de habitantes			Serie C.3. Miles de habitantes		
	t	P_i	Z_i	$Z_i - P_i$	P_i	Z_i	$Z_i - P_i$
1875	0	2 076	2 114	+ 38	2 219	2 179	− 40
1885	10	2 507	2 408	− 99	2 472	2 461	− 11
1895	20	2 696	2 744	+ 48	2 720	2 780	+ 60
1907	32	3 231	3 209	− 22	3 210	3 218	+ 8
1920	45	3 732	3 802	+ 70	3 710	3 771	+ 61
1930	55	4 287	4 332	+ 45	4 287	4 260	− 27
1940	65	5 024	4 936	− 88	4 887	4 813	− 74
Σ		23 553	23 545	410	23 505	23 482	281
$V = 100 \frac{\Sigma \mid (Z_i - P_i) \mid}{\Sigma P_i}$				1,74		1,24	1,24
1950			5 624				
1960			6 408				
1970			7 302				
$Z_i = 2\ 114 \times 1{,}013133^t$					$Z_i = 2\ 179 \times 1{,}012269^t$		
$\lg A = 3{,}325105$					$\lg A = 3{,}338257$		
$\lg q = 0{,}005667$					$\lg q = 0{,}005295$		

TABLA III.—POBLACION DE CHILE; SERIES C.4 Y C.5

Años	t	Serie C.4. Miles de habitantes P_i	Z_i	$Z_i - P_i$	Serie C.5. Miles de habitantes P_i	Z_i	$Z_i - P_i$
1870	0	1 944	1 971	+ 27	2 087	2 056	− 31
1880	10	2 282	2 248	− 34	2 342	2 319	− 23
1890	20	2 600	2 563	− 37	2 593	2 616	+ 23
1900	30	2 907	2 923	+ 16	2 911	2 952	+ 41
1910	40	3 340	3 333	− 7	3 312	3 330	+ 18
1920	50	3 732	3 800	+ 68	3 709	3 757	+ 48
1930	60	4 287	4 333	+ 46	4 287	4 240	− 47
1940	70	5 024	4 941	− 83	4 887	4 783	− 104
		26 116	26 112	318	26 128	26 053	335
$V = 100 \frac{\Sigma \mid (Z_i - P_i) \mid}{\Sigma P_i}$				1,22			1,27
1950	80		5 625			5 356	
1960	90		6 413			6 034	
1970	100		7 313			6 840	

$Z_i = 1\ 971 \times 1{,}01321^t$
$\lg A = 3{,}294087$
$\lg 1{,}01321 = 0{,}0057$

$Z_i = 2\ 056 \times 1{,}0121^t$
$\lg A = 3{,}312023$
$\lg 012{,}134 = 0{,}00523$

TABLA IV.—POBLACION DE CHILE: SERIE C.5

Años	t	Serie C.5. Miles de habitantes P_i	P_i	$Z_i - P_i$	Serie C.5. Miles de habitantes P_i	Z_i	$Z_i - P_i$
1870	0	2 087	2 061	− 26	2 087	2 096	+ 9
1880	10	2 342	2 325	− 17	2 342	2 326	− 16
1890	20	2 593	2 623	+ 30	2 593	2 595	+ 2
1900	30	2 911	2 960	+ 49	2 911	2 912	+ 1
1910	40	3 312	3 339	+ 27	3 312	3 287	− 25
1920	50	3 709	3 767	+ 58	3 709	3 733	+ 24
1930	60	4 287	4 250	− 37	4 287	4 268	− 19
1940	70	4 887	4 795	− 92	4 887	4 912	+ 25
Σ		26 128	26 120	336	26 128	26 129	121

Años		Serie C.5. Miles de habitantes			Serie C.5. Miles de habitantes		
	t	P_i	Z_i	$Z_i - P_i$	P_i	Z_i	$Z_i - P_i$
$V = 100 \dfrac{\Sigma \lvert (Z_i - P_i) \rvert}{\Sigma P_i}$				1,27			0,46
1950	80		5 410			5 694	
1960	90		6 104			6 652	
1970	100		6 886			7 832	
$Z_i = 2061 \times 1{,}012136^t$ $\lg A = 3{,}314078$ $\lg q \quad 0{,}005239$					$Z_i = Ka^{b^t}$ $k = 272{,}863$ $a = 7{,}68$ $b = 1{,}005$		

19. En conclusión, si se parte de las cifras oficiales de los censos efectuados entre 1875 y 1940 (Series C.2 y C.4), se puede concluir que la tasa de crecimiento acumulativo media anual ha sido de 13°/₀₀. Si se basa el cálculo en las cifras corregidas de aquellos censos, C.3 y C.5, la conclusión es que la tasa de crecimiento acumulativo media anual ha sido prácticamente de 12°/₀₀.

Puesto que las correcciones aplicadas son razonables y necesarias para ajustarse a la verdad, nos parece recomendable adoptar como desarrollo de la población de Chile desde 1870 adelante, la serie C.5.

Ahora bien, es posible que la tasa 12°/₀₀ resultante no varíe apreciablemente en un lapso relativamente grande; esto es que la velocidad de crecimiento $\frac{dP}{dt}$ continúe todavía aumentando; pero no es discutible, puesto que la población de ningún país puede crecer indefinidamente, que dicha velocidad ha de llegar alguna vez a un valor máximo a partir del cual el aumento de población entre censos empieza a disminuir. Sobre ningún indicio podemos basarnos para predecir esa época, que solamente esperamos.

Estaría demás agregar que cualquiera que sea la serie que se adopte, los cálculos deben rehacerse con posterioridad a cada nuevo censo, a fin de que la base de previsión esté siempre ajustada a la realidad más cercana.

La extrapolación para los años 1950 y 1960 se encuentra consignada en las mismas tablas II a IV. Estos valores son cifras previstas de población futura, tanto más próximas a ser realidades, cuanto menos se alteren las condiciones bajo el influjo de las cuales se ha producido el desarrollo de la población en los años precedentes. Aun cuando los efectos de los cambios de circunstancias no se hacen sentir bruscamente, y ni siquiera en lapso muy breve, la condición indicada limita la verosimilitud de la extrapolación a plazos no muy prolongados; digamos del orden de los 25 a 30 años. Para plazos más largos habría que ejercitar la virtud de la prudencia, que en el caso de que se trata puede influir positiva o negativamente.

Esta virtud debe ejercitarse de consuno con la paciencia en la búsqueda de los factores que han de intervenir para modificar el desarrollo de la población en el futuro, y la adopción consciente, es decir sin subestimaciones ni exageraciones derivadas de ideas preconcebidas.

La extrapolación para el año 1950 con la serie C.5 de la 7.ª columna de la Tabla IV da una población probable de 5 694. La estadística chilena da la cifra de 5 709 para junio de este año, calculada por medio de la expresión:

$$P = P_0 (1+KN) - D$$

en la cual:

P_0 es el resultado del último censo;

N el número de nacimientos inscritos a partir del último censo (inscripción ordinaria que no comprende los inscritos mayores de 2 años de edad);

D el número de defunciones inscritas (sin incluir los nacidos muertos); y

K un coeficiente de mayoración que antes fué igual a 1,2% y ahora es 2,4%.

La expresión indicada no toma en cuenta el saldo migratorio.

La falta de aproximación a la verdad de este método empírico puede apreciarse fácilmente mediante el examen de las siguientes cifras:

En el tiempo transcurrido desde el último censo hasta el 30 de junio próximo pasado, se ha tenido una inscripción ordinaria de 1 498 318 nacidos vivos y se han registrado 851 883 defunciones; de modo que el saldo o crecimiento natural fué de 646 435 habitantes. El término aditivo que origina el coeficiente K es de 35 960 habitantes, con lo cual la población calculada se lleva a 5 709 000 habitantes en números redondos.

Las inscripciones extraordinarias, es decir de mayores de 2 años de edad alcanzaron en el mismo período de tiempo a 317 993 habitantes; de modo que queda una diferencia de 317 993 − 35 960 = 282 033 habitantes no incluídos en la población calculada.

Tampoco está incluído el saldo migratorio que en el lapso considerado fué de + 21 823 personas.

Nos parece muy justificado preconizar el abandono de la citada expresión empírica y su reemplazo por la siguiente:

$$P_i = P_0 + N - D + N_e + Sm$$

que toma en consideración valores ciertos y registrados por la estadística, como lo son el número N de nacimientos inscritos en el registro ordinario, el número D de defunciones registradas, el número N_e de inscripciones extraordinarias de nacimientos y el número Sm, del saldo migratorio positivo o negativo.

Para comparar las cifras 5 694 que da la función de ajustamiento de Gompertz con la de 5 709 que indica como población calculada para junio próximo pasado, deberíamos corregir ésta, adicionando 282 y 22 y restándole el término correctivo 137 en el N.° 12.

Así tendríamos:

$$5\,709 - 282 - 22 + 137 = 5\,869$$

de modo que ya en junio de 1949 tendríamos una población superior en más de 175 mil habitantes a la prevista para fines de 1950.

Si se examina la estadística demográfica de los últimos años se observa una ligera atenuación de coeficiente de natalidad y un decrecimiento importante del coeficiente de mortalidad, Se apreciará lo dicho anteriormente con el examen de las siguientes cifras:

	NATALIDAD		MORTALIDAD		CRECIMIENTO VEGETATIVO	
AÑOS	Coeficiente medio %	N.°	Coeficiente medio %	N.°	Coeficiente medio %	Absoluto
1933 - 1940	33,7	1 247 625	24,9	909 692	9,4	446 182
1941 - 1948	33,3	1 421 891	18,6	801 443	13,9	620 448
Diferencias	— 0,4	+ 174 266	— 6,3	— 108 049	+ 4,4	—+174 266

Los coeficientes demográficos de natalidad y mortalidad están afectados por el error voluntario positivo del último censo y por la omisión de los nacidos después de los dos años de efectuado el censo respectivo. No podemos apreciar la importancia de esta omisión en la población total que es uno de los factores que interviene en el cálculo de los expresados coeficientes, pues conocemos sólo las cifras de inscritos con más de dos años de edad; pero no sabemos cuanto de estos nacieron dos años después de efectuado el censo.

Si E es aquel error voluntario e i la omisión aludida los coeficientes demográficos de natalidad y mortalidad deberían multiplicarse por el factor $\left(1 + \frac{E - i}{P}\right)$ en que el segundo término podrá ser positivo o negativo. Es obvio que el coeficiente de crecimiento vegetativo está exento de esta corrección.

20. Tomemos ahora como ejemplo la estadística de los censos decenales efectuados en EE. UU. desde 1790 hasta 1940. La estadística que utilizamos comprende sólo la población continental y la sucesión es la que se reproduce en C.3 de la Tabla V.

Puesto que la tasa media anual del crecimiento de la población en los períodos sucesivos acusa una atenuación después del año 1910, aparece justificado elegir la logística como curva de ajustamiento.

Para encontrar los coeficientes K, b y λ procederemos como sigue: elegimos tres puntos de abcisas t_1, t_2 y t_3 equidistantes y que tienen por ordenadas P_1, P_2 y P_3, y sea m la distancia entre aquellas ordenadas. Escribamos ahora las siguientes relaciones:

$$\frac{1}{P_1} = \frac{1+be^{-\lambda t_1}}{K} = \frac{e^{\lambda t_1} + b}{Ket_1}$$

$$\frac{1}{P_2} = \frac{1 + be^{-\lambda(t_1+m)}}{X} = \frac{e^{\lambda(t_1+m)} + b}{Ke^{\lambda(t_1+m)}}$$

$$\frac{1}{P_3} = \frac{1 + be^{-\lambda(t_1+2m)}}{K} = \frac{e^{\lambda(t_1+2m)}}{Ke^{\lambda(t_1+2m)}}$$

De estos valores recíprocos de los puntos elegidos, obtenemos:

$$\Delta_1 = \frac{1}{P_2} - \frac{1}{P_1} = \frac{b}{K} \quad \frac{e^{\lambda m} - 1}{e^{\lambda t_1} \cdot e^{\lambda m}}$$

$$\Delta_2 = \frac{1}{P_3} - \frac{1}{P_2} = \frac{b}{K} \quad \frac{e^{\lambda m} - 1}{e^{\lambda t_1} \cdot e^{\lambda m}}$$

$$\Delta_1 - \Delta_2 = \frac{P_1 - P_2}{P_1 P_2} - \frac{P_2 - P_3}{P_2 P_3} = \frac{(e^{\lambda m} - 1)^2}{e^{\lambda t_1} \cdot e^{2\lambda m}}$$

Finalmente:

$$R_1 = \frac{\Delta_1}{\Delta_2} = \frac{(P_1 - P_2)}{(P_2 - P_3)} = e^{\lambda m}$$

de donde:

$$\lambda = \frac{\lg R}{m \lg e}$$

$$Q = \frac{\Delta_1^2}{\Delta_1 - \Delta_2} = \frac{b}{Ke^{\lambda t_1}}$$

$$\Delta_3 = \frac{1}{P_1} - Q = \frac{e^{\lambda t} + b}{Ke^{\lambda t_1}} - \frac{b}{Ke^{\lambda t_1}} = \frac{1}{K}$$

Luego:

$$K = \frac{1}{\Delta_3} \qquad y \qquad b = \frac{Qe^{\lambda t_1}}{\Delta_3}$$

Sean entonces, las tres ordenadas con sus respectivas abcisas:

$P_1 = 3929$	$P_2 = 34\ 280$	$P_3 = 131\ 669$
$t_1 = 0$	$t_2 = 75$	$t_3 = 150$

La equidistancia entre las ordenadas es igual a 75, y se advierte que para el año 1865, al cual corresponde esta abcisa, se ha tomado como ordenada la cifra 34 820, que es la media geométrica entre las ordenadas de abcisas 70 y 80 correspondientes a los años 1860 y 1870, respectivamente.

Con dichos valores se ha obtenido la ecuación:

$$Z_i = \frac{184693}{1 + 46{,}005\, e^{-0{,}0315885\, t}} \tag{22}$$

Los valores sucesivos de Z_i se encuentran inscritos en la columna 5 de la Tabla V y las diferencias $Z_i - P_i$ en la columna 6.

TABLA V.—POBLACION CONTINENTAL DE ESTADOS UNIDOS

Años C1	t C2	P_i (miles) C3	C4	Z_i (miles) C5	$Z_i - P_i$ C6	Z_i C7	$Z_i - P_i$ C8
1790........	0	3 929		3 929	0	3 930	1
1800........	10	5 308	35,4	5 347	+ 39	5 353	+ 45
1810........	20	7 240	36,4	7 255	+ 15	7 267	+ 27
1820........	30	9 638	33,1	9 806	+ 168	9 839	+ 201
1830........	40	12 866	33,5	13 189	+ 323	13 246	+ 380
1840........	50	17 070	32,7	17 621	+ 551	17 711	+ 641
1850........	60	23 192	35,9	23 400	+ 208	23 480	+ 288
1860........	70	31 443	35,6	30 574	— 869	30 778	— 665
1870........	80	38 558	22,6	39 503	+ 945	39 785	+1 227
1880........	90	50 156	30,1	50 189	+ 33	50 563	+ 407
1890........	100	62 948	25,5	62 522	— 426	62 990	+ 42
1900........	110	75 995	20,1	76 169	+ 174	76 725	+ 730
1910........	120	91 972	21,0	90 586	—1 386	91 210	— 762
1920........	130	105 711	14,9	105 103	— 608	105 754	+ 43
1930........	140	122 775	16,1	119 179	—3 596	119 648	—3 127
1940........	150	131 669	7,2	131 669	0	132 307	+ 638
Σ		790 470		786 041	9 431	790 586	9 224

$$V_i = 100\,\frac{\Sigma\,|(Z_i - P_i)|}{\Sigma\, P_i} \qquad 1{,}19 \qquad 1{,}16$$

1945	155	138 042
1950	160	143 356
1955	165	148 222
1960	170	152 632

$$Z_i = \frac{184693}{1 + 46{,}005\, e^{-0{,}0315885\, t}} \qquad Z_i' = \frac{184700}{1 + 46\, e^{-0{,}0317\, t}}$$

La suma de estas diferencias $|(Z_i - P_i)|$ representa sólo un 1,18% de la suma de P_i, coeficiente que según dijimos, podemos tomar como medida de la variabilidad de los valores calculados con respecto a los experimentales.

Observemos, desde luego, que dicho valor no es ajeno a la elección que hemos hecho de los puntos P_1, P_2 y P_3.

Algunos autores, con el propósito de encontrar mayores probabilidades de obtener una función de ajustamiento que evite afinar los cálculos posteriormente, prefieren tomar como ordenada en cada abcisa su media geométrica con las dos ordenadas vecinas; pero el método no ofrece garantías suficientes de que podrá evitarse una corrección posterior de los parámetros.

21. Si estimamos aquel valor demasiado alto para considerar satisfactoria la función de ajustamiento resultante, habría que introducir términos de corrección de los coeficientes K, b y λ;

La nueva ecuación será:

$$P_i = \frac{K + \alpha}{1 + (b + \beta)\, e^{-(\lambda + \gamma)t}} \tag{23}$$

Desarrollando en serie el término $e^{-\partial t}$ y despreciando los términos de grado superior al primero, tendremos:

$$K - P_i\,(1 + be^{-\lambda t}) = \beta\, P_i e^{-\lambda t} + P_i bte^{-\lambda t} - \alpha \tag{24}$$

Desígnemos por Z_i los valores sucesivos de la función primitiva y substituyamos en (24) $(1 + be^{-\lambda t})$ por $\frac{K}{Z_i}$. Tendremos:

$$K\,(Z_i - P_i) = \beta\, P_i\, Z_i\, e^{-\lambda t} + \gamma\, P_i\, Z_i\, bte^{-\lambda t} - \alpha\, Z_i$$

El valor de los términos correctivos se encontrará por el método de los cuadrados menores. La condición de hacer mínima la suma de los cuadrados de las diferencias entre los valores observados P_i y los Z_i que proporcione la nueva función será:

$$\Sigma\left[\beta\, P_i\, Z_i\, e^{-\lambda t} + \gamma\, P_i\, Z_i\, bte^{-\lambda t} - \alpha\, P_i - K\,(Z_i - P_i)\right]^2 = \text{mínima}$$

La condición anterior implica que sean nulas las primeras derivadas de esa función con respecto a α, β y γ. Por consiguiente:

$$\Sigma\left[\beta P_i^2\, Z_i^2 e^{-2\lambda t} + \gamma\, btP_i^2 Z_i^2 e^{-2\lambda t} - \alpha\, P_i\, Z_i e^{-\lambda t} - KP_i\, Z_i(Z_i - P_i)e^{-\lambda t} = 0\right.$$

$$\Sigma\left[\beta btP_i^2\, Z_i^2 e^{-2\lambda t} + \gamma b^2 t^2 P_i Z_i e^{-2\lambda t} - \alpha btP_i Z_i^2{}^{-\lambda t} - KbtP_i Z_i(Z - P_i)e^{-\lambda t}\right] = 0$$

$$\Sigma\left[\beta\, P_i\, Z_i^2 e^{-\lambda t} + \gamma\, btP_i Z_i^2 e^{-\lambda t} - \alpha\, K_i^2 - KZ_i\,(Z_i - P_i)\right] = 0$$

Con el propósito de simplificar las escrituras designemos los siguientes términos con las letras que se indican:

$$\Sigma P_i^2 Z_i^2 e^{-2\lambda t} = B \qquad \Sigma P_i Z_i^2 e^{-\lambda t} = H$$

$$b \Sigma t P_i^2 Z_i^2 e^{-2\lambda t} = C \qquad bK \Sigma t P_i Z_i (Z_i - P_i) e^{-\lambda t} = Q$$

$$b^2 \Sigma t^2 P_i^2 Z_i^2 e^{-2\lambda t} = D \qquad K \Sigma_i P_i (P - Z_i) = R$$

$$b \Sigma t P_i Z_i^2 e^{-\lambda t} = E \qquad K \Sigma P_i Z_i (Z_i - P_i) e^{-\lambda t} = S$$

$$\Sigma P_i^2 = F$$

Las ecuaciones normales se escribirán entonces:

$$B\beta + C\gamma - H\alpha - S = 0$$

$$C\beta + D\gamma - E\alpha - Q = 0$$

$$H\beta + E\gamma - F\alpha - R = 0$$

De estas ecuaciones deducimos:

$$\alpha = \frac{S(EC - HD) - C(RG - HQ) + B(RD - EQ)}{B(FD - E^2) + C(FC - EH) - H(EC - HD)}$$

$$\beta = \frac{S(FD - E^2) + C(FQ - RE) - H(EQ - DR)}{B(FD - E^2) + C(FC - EH) - H(EC - HD)}$$

$$\gamma = \frac{S(FC - HE) - B(FQ - RE) - H(RC - HQ)}{B(FD - E^2) + C(FC - EH) - H(EC - HD)}$$

Si se substituyen en lugar de las letras los términos que ellas representan, se podrán calcular los valores de α, β y γ; pero desde luego podrá juzgarse acerca de la fatiga que esta operación ocasiona; fatiga que no aparece suficientemente justificada, en relación con la exactitud relativa de los datos observados, y a la que ya hemos aludido. De aquí es que no se practique habitualmente este afinado de los coeficientes de la función logística sino en los casos de estudios simplemente especulativos. Esta circunstancia ha dado ocasión a que se objete la curva logística, porque la arbitrariedad en la elección de los tres censos es causa de que se obtengan valores distintos para los tres parámetros. Por otra parte, la intervención de sólo tres censos de la serie, implica la prescindencia de toda la información proveniente de los demás, y obligándose a la logística a satisfacerse con ellos, o sea, a aceptarlos como correctos, quedan encubiertos los errores que ellos puedan contener, así como sus oscilaciones con relación a la propia línea de tendencia que se busca para describir el fenómeno.

Por otra parte, la rigidez de las condiciones de simetría de la curva, conduce a que el punto de inflexión aparezca con diversas coordenadas, según cuales sean los tres puntos elegidos y casi siempre en discordancia con las informaciones que proporciona la serie de valores observados.

22. Se podrían encontrar los valores de α, β y γ por un método de aproximaciones sucesivas, como el siguiente:

Busquemos primero el correctivo γ.

La ecuación (6) será ahora la siguiente:

$$P_i = \frac{K}{1 + be^{-(\lambda + \gamma)t}} = \frac{K}{1 + be^{-\lambda t - \gamma t}}$$

Desarrollando en serie el factor $e^{\gamma t}$ y despreciando los términos de grado superior al primero, obtendremos:

$$P_i = Z_i (1 + \gamma t) \tag{25}$$

Aplicando el método de los cuadrados menores, escribiremos la ecuación de condición:

$$\Sigma (Z_i + \gamma t Z_i - P_i)^2 = \text{mínimo}$$

$$\Sigma tZ_i^2 - \gamma \Sigma t^2 Z_i^2 - \Sigma tZ_i P_i = 0$$

Luego:

$$\gamma = \frac{\Sigma tZ_i P_i - \Sigma tZ_i^2}{\Sigma t^2 Z_i^2}$$

Conocido γ tendremos $\lambda' = \lambda + \gamma$ y $Z_i' = \dfrac{K}{1 + be^{-\lambda' t}}$

Buscamos ahora los términos correctivos β y α. La nueva ecuación de la logística será:

$$P_i = \frac{K + \alpha}{1 + (b + \beta) e^{-\lambda' t}} = \frac{K + \alpha}{1 + be^{-\lambda' t} + \beta e^{-\lambda' t}}$$

de donde:

$$K (Z_i' - P_i) = \beta P_i Z_i' e^{-\lambda' t} - Z_i' \alpha$$

Aplicando otra vez el método de los cuadrados menores, escribiremos la ecuación de condición:

$$\Sigma \left[\beta P_i Z_i' e^{-\lambda' t} = \alpha Z_i' - K (Z_i' - P_i')\right]^2 = \text{mínimo}$$

y por consiguiente:

$$\beta \Sigma P_i^2 Z_i'^2 e^{-2\lambda' t} - \alpha \Sigma P_i Z_i'^2 e^{-\lambda' t} - K \Sigma P_i Z_i' e^{-\lambda' t} (Z_i' - P_i') = O$$

$$\beta \Sigma P_i Z_i'^2 e^{-\lambda' t} - \alpha \Sigma Z_i'^2 - K \Sigma Z_i' (Z_i' - P_i) = O$$

Para simplificar la escritura designemos los términos siguientes por las letras que se indican:

$$\Sigma P_i^2 Z_i'^2 e^{-2\lambda'} = A \qquad \Sigma Z_i'^2 = D$$

$$\Sigma P_i Z_i'^2 e^{-\lambda' t} = B \qquad \Sigma Z_i' (Z_i' - P_i) = E$$

$$\Sigma P_i Z_i' e^{-\lambda' t} (Z_i' - P_i') = C$$

Entonces las ecuaciones normales serán:

$$\beta A - \alpha B - BC = O$$

$$\beta B - \alpha D - KE = O$$

De aquí:

$$\alpha = \frac{K(AE - BC)}{AD - B^2}$$

$$\beta = \frac{K(DC - BE)}{AD - B^2}$$

Substituyendo las letras por los términos que ellas representan calcularíamos los valores de los correctivos α y β. La fatiga operatoria ha disminuído; pero se mantiene todavía muy considerable para los cálculos de la práctica; tanto más cuanto mayor sea el número de términos de la serie cronológica de que se trata.

23. Otro método utilizable para el cálculo de los coeficientes de que se trata sería el siguiente, asimilable al empleado por Pearl para calcular los coeficientes de la ecuación (15).

Calculados los parámetros K, b y λ, adoptaríamos $K' = K + \alpha$, determinando α como en los casos anteriores:

$$P_i = \frac{K + \alpha}{1 + be^{-\lambda t}} = Z_i + \frac{\alpha}{1 + be^{-\lambda t}} = Z_i + \alpha \rho$$

De la ecuación de condición:

$$\Sigma (Z_i + \alpha\rho - P_i)^2 = \text{mínimo}$$

de la cual deducimos:

$$\Sigma \rho Z_i + \alpha \Sigma \rho^2 - \Sigma \rho P_i = 0$$

De donde:

$$\alpha = \frac{\Sigma \rho P_i - \Sigma \rho Z_i}{\Sigma \rho^2}$$

Ahora, la anamorfosis logarítmica de la nueva ecuación:

$$P_i = \frac{K'}{1 + b'e^{-\lambda' t}} \tag{26}$$

nos permitirá encontrar directamente los parámetros $b' = b + \beta$ y $\lambda' = \lambda + \gamma$.

En efecto, de la ecuación (26) deducimos:

$$\lg (K' - P_i) - \lg P_i = \lg b' - \lambda' t \lg e$$

Otra vez la aplicación del método de los cuadrados menores nos permite obtener el sistema de ecuaciones normales:

$$\lambda' \lg e \, \Sigma t^2 - \lg b' \, \Sigma t + \Sigma t \lg (K' - P_i) - \Sigma t \lg P_i = 0$$

$$\lambda' \lg e \, \Sigma t - n \lg b' + \Sigma \lg (K' - P_i) - \Sigma \lg P_i = 0$$

De donde obtenemos:

$$\lg b' = \frac{\Sigma t^2 [\Sigma \lg (K' - P_i) - \Sigma \lg P_i] - \Sigma t [\Sigma t \lg (K' - P_i) + \Sigma t \lg P_i]}{n \Sigma t^2 - (\Sigma t)^2},$$

$$\lambda' = \frac{\Sigma t [\Sigma \lg (K' - P_i) - \Sigma \lg P_i] - n [\Sigma t \lg (K' - P_i) + \Sigma t \lg_i]}{\lg e [n \Sigma t^2 - (\Sigma t)^2]}$$

Como se ve, la fatiga operatoria ha disminuído algo más; pero existe ahora mayor base para juzgar de que el cálculo conducirá a una logística de ajustamiento aceptable, en las condiciones actuales del problema.

24. Volvamos a la serie estadística de la población continental de EE. UU. Observamos ya que la medida de la variabilidad adoptada es tan reducida que no aconseja una corrección de los parámetros, tan trabajosa como la que resulta de los métodos matemáticos expuestos. Sin embargo, si observamos que la suma algebraica de las diferencias $Z_i - P_i$ tienen un valor cercano al 50% de la suma absoluta de las mismas, nos tentaríamos a procurar una corrección que acercara más la curva de ajustamiento a los valores observados.

Obtendremos este resultado por simple tanteo, que nos lleva a la ecuación:

$$Z_i' = \frac{184\ 700}{1 + 46e^{-0,0317t}} \tag{27}$$

con lo cual calculamos los valores de la C7 de la Tabla V. Se observará que si bien el coeficiente de variabilidad V ha mejorado relativamente poco, en cambio la diferencia $\Sigma Z_i - \Sigma P_i$ ha bajado de 4 429 a 116, lo que indica que la curva se ha aproximado a los valores experimentales.

W. A. Kostinzin, en su «Biologie Mathematique», París, 1937, señala que la logística

$$P = \frac{197\ 293,000}{1 + e^{-0,03134\ (t-1913,25)}} \tag{28}$$

representa muy bien la población de EE. UU; y agrega una tabla comparativa entre los valores de los censos (1790 - 1930) y los calculados con la expresión anterior. Desgraciadamente los valores calculados difieren bastante de los que proporciona la ecuación (27) y se anotan además, algunas diferencias entre los valores de los censos que figuran en dicha tabla y las cifras oficiales que hemos empleado en este estudio. En todo caso, interesa anotar que la diferencia ($| \Sigma Z_i - \Sigma P_i |$) en la tabla de Kostinzin es de + 4 100, que se eleva a 4 200 en la que se formaría con las cifras verdaderas. Observaremos todavía que la ecuación (28) proporciona por extrapolación la cifra 137 719 para el año 1940 y el censo dió sólo 131 669.

25. Ajustemos ahora algunas poblaciones de la Provincia de Concepción. No poseemos antecedentes para corregir las cifras oficiales de los censos, como en el caso de la población total del país; de modo que tomaremos como base dichas cifras. Las relativas a la provincia colocadas en gráfica nos inducen a adoptar como función de ajustamiento la exponencial:

$$P_i = Aq^t$$

La aplicación del método de los cuadrados menores previa la transformación logarítmica a la recta:

$$\lg Z_i = \lg A + t \lg q$$

nos permite calcular los siguientes valores:

$$\lg A = 2,136865 \quad ; \quad A = 137$$
$$\lg q = 0,005277 \quad ; \quad q = 1,012226$$

Los Z_i correspondientes y sus diferencias con P_i se han insertado en la Tabla VI. La población de la ciudad de Concepción ajustada por medio de una recta nos da:

$$P_i = 17,28 + 1,076t$$

Los valores correspondientes se consignan en la segunda sección de la Tabla VI.

TABLA VI.—POBLACION DE CONCEPCION

Años		Provincia			Ciudad		
		Miles de habitantes			Miles de habitantes		
	t	P_i	Z_i	$Z_i - P_i$	P_i	Z_i	$Z_i - P_i$
1875	0	134,3	137,0	+ 2,7	18,3	17,3	— 1,0
1885	10	163,8	154,8	— 9,0	24,2	28,0	+ 3,8
1895	20	171,2	174,8	+ 3,6	39,8	38,8	— 0,1
1907	32	200,4	202,1	+ 1,7	55,3	51,7	— 3,6
1920	45	229,4	236,8	+ 7,4	64,1	65,7	+ 1,6
1930	55	268,4	267,4	— 1,0	77,6	76,6	— 1,0
1940	65	308,2	302,0	— 6,2	85,8	87,2	+ 1,8
Σ		1 475,7	1 474,9	31,6	365,1	364,8	12,9
$V = 100\,\Sigma \frac{\|(Z_i - P_i)\|}{\Sigma P_i} =$				2,14			3,53
1950	75						
1960	85						
1970	95						
$P_i = 137 \times 1{,}012226^t$					$P_i = 17{,}28 + 1{,}076\,t$		

TABLA VII.—POBLACION DE CONCEPCION

Años		Depto. de Concepcion			Depto. de Talcahuano		
		Miles de habitantes			Miles de habitantes		
	t	P_i	Z_i	$Z_i - P_i$	P_i	Z_i	$Z_i - P_i$
1875	0	42,8	41,2	− 1,6	2,5		
1885	10	52,8	53,1	+ 0,3	5,0		
1895	20	65,0	65,0	0	10,5		
1907	32	80,3	79,3	− 1,0	15,6		
1920	45	90,8	94,8	+ 4,0	22,1		
1930	55	106,1	106,7	+ 0,6	27,6		
1940	65	121,6	118,6	− 3,0	35,8		
Σ		559,4	558,7	10,5	119,0		

Años	Depto de Concepción Miles de habitantes			Depto. de Talcahuano Miles de habitantes		
t	P_i	Z_i	$Z_i - P_i$	P_i	Z_i	$Z_i - P_i$
$V = 100 \frac{\Sigma \mid (Z_i - P_i) \mid}{\Sigma P_i}$			1,88			
1950 75						
1960 85						
1970 95						
$P_i = 41,2 + 1,19\,t$				$P_i =$		

Asimismo, en la Tabla VII se encuentran los valores correspondientes a la población del Departamento de Concepción, ajustado con la función:

$$P_i = 41,2 + 1,19\,t$$

26. La población de la Provincia de Concepción ha crecido en la proporción de 100 a 231 y su relación con la población total del país ha bajado de 6,45% a 6,13%.

Ni estas cifras ni las deducidas de las curvas de ajustamiento que hemos calculado pueden tomarse como guías únicas para prever lo que sucederá en el futuro, aunque no estuviesen afectadas de los errores que han hecho necesarias las correcciones aplicadas al desarrollo de la población del país; puesto que existen claras manifestaciones de que no se mantendrán las condiciones que han influído en el desarrollo de la población local.

El cálculo de la población actual tiene interés para conocer lo que haya sucedido entre el último censo y el momento en que se calcula, e inducir de este conocimiento alguna conclusión respecto a lo que probablemente sucederá en el tiempo venidero próximo.

Conviene, pues, estudiar el método que aplica la Dirección General de Estadística. Esta Dirección emplea para calcular la población de las provincias o de las ciudades en un momento determinado, la expresión:

$$P_i' = \frac{K n_i'}{n_i} P_i$$

en la cual:

P_i' es la población en el instante en que se calcula;

P_i la población en el instante inicial;

n_i el número de nacimientos registrados en los doce meses que preceden al instante inicial;

n_i' el número de nacimientos que preceden al instante en que se calcula; y

K un coeficiente que tiene por valor:

$$K = \frac{P_i'}{\Sigma \frac{n_i'}{n_i} P_i} = \frac{P_i + 1{,}024\ N\text{-}D}{\Sigma \frac{n_i'}{n_i} P_i}$$

El método indicado se basa en aceptar que la natalidad es el mejor índice de movimientos de población regionales, y el menos afectado por causas extraordinarias que no sean regionales. Admite, como se ve, una proporcionalidad K igual para toda provincia o localidad, que no toma en cuenta sino un solo coeficiente demográfico y prescinde de los movimientos migratorios internos.

Antes hemos visto (N.° 19) que para calcular la población total del país era recomendable substituir el método empírico actual por otro que tome en cuenta todos los términos que integran aquella suma algebraica. Análogamente, nos parece lógica igual recomendación para el caso de una provincia, ciudad o unidad de población cualquiera, puesto que para todas ellas se conocen todos los términos demográficos que intervienen. Así, la población actual de una de estas unidades debería ser:

$$P_i' = P_i + n_i - d_i + n_e \pm s_e \pm s_i$$

en la cual:

P_i', población en el momento actual;

P_i, población en el momento inicial;

n_i, el número de nacimientos registrados entre los dos momentos;

d_i, número de defunciones registradas entre los dos momentos;

n_i, el número de inscripciones extraordinarias de nacimientos entre los dos momentos, y que hayan nacido después de 2 años de efectuado el último censo;

s_e, el saldo migratorio externo; y

s_i, el saldo migratorio interno.

Los dos últimos términos ofrecen serias dificultades.

Cuanto al primero, a fin de guardar concordancia con el cálculo de la población total del país deberíamos tomarlo en consideración aceptando alguna hipótesis respecto a su repartición en el país, como sería la de suponer que s_e se distribuya entre las unidades de población en cuotas proporcionales al número de sus habitantes. De este modo:

$$s_e = wP_i'$$

$$\text{siendo} \quad w = \frac{s_e}{P'}$$

Cuanto a los movimientos migratorios internos producidos después de un último censo, no tenemos medio estadístico alguno para conocerlos, y en atención a esta dificultad insuperable, creemos que las Estadísticas tienen que prescindir de él para el cálculo de la población actual.

Pero el investigador que trata de fundamentar alguna resolución en la población futura calculada, debe agotar sus esfuerzos en la búsqueda de este término por medios indirectos

27. Tratándose de la provincia de Concepción y sus principales centros poblados, sería erróneo en esta ocasión basarse exclusivamente en las curvas de ajustamiento que hemos determinado, para formular previsiones acerca de su población en un futuro más o menos alejado. La instalación de una gran industria básica como la Siderúrgica de Huachipato y sus derivados, altera radicalmente las condiciones preexistentes.

Estaría demás dar énfasis a la recomendación de ser extremadamente cuidadosos en la apreciación de la influencia más o menos considerable que pueden ejercer los nuevos factores, puesto que ella no es consecuencia directa y exclusiva de estos últimos, sino que dependen en parte muy considerable de las condiciones locales.

Así, por ejemplo, se incurriría en serio error si se pretendiera medir la influencia que ejercerá la implantación de la Siderúrgica de Huachipato en el desarrollo futuro de la Población de Concepción por la que ejerció el establecimiento de los Altos Hornos de Corral en la localidad de ese nombre y en su vecina principal, la ciudad de Valdivia.

Las condiciones locales de Concepción y Corral son substancialmente distintas desde el punto de vista de la atracción que son capaces de ejercer para que nuevas industrias se radiquen dentro de sus límites jurisdiccionales; y por lo tanto, de provocar un cambio apreciable en el desenvolvimiento de su futura población.

28. Observaciones de la índole de las que anteceden, que se aprecian fácilmente en el caso considerado, deben examinarse en todos los casos en que se trate de avaluar una población futura. La curva de ajustamiento indica sólo una tendencia que se modifica con la introducción de los resultados de los censos sucesivos y que sólo tiene valor de previsión mientras se conserven inalterables las condiciones preexistentes. Cualquier cambio previsto debe ser considerado y estimada su influencia para afinar debidamente la previsión anterior.