NUEVA DETERMINACION DE LA CURVA LOGISTICA DE LA POBLACION DE ESPAÑA

INTRODUCCION

En los veinticinco primeros años del siglo actual se concedió un gran interés al estudio de las leyes fundamentales del crecimiento de las poblaciones, aunque ya, en 1844, el matemático belga Verhulst publicó, en una Memoria de la Real Academia de Bruselas, el trabajo titulado *Recherches Mathématiques sur la loi d'accroissement de la population*, en el que aparece y se usa por primera vez la curva logística de ecuación

$$Y = \frac{m\,k\,e^{at}}{1 + m\,e^{at}} \qquad [1]$$

para representar el crecimiento de una población.

Raymond Pearl y Rowell J. Reed (1) fueron los principales divulgadores de esta teoría al tratar de evaluar las variaciones de la población de los Estados Unidos desde 1870, para lo que utilizaron la curva de ecuación [1]. Se plantearon el siguiente problema: *Dado que un espacio limitado (como la Tierra) habitado por seres vivos de los que se conoce una cierta vida media, su forma de reproducción, sus necesidades vitales, etc., ¿qué sucederá en el transcurso del tiempo respecto al número de individuos que pueblen ese espacio?*

---

(1) PEARL, R. y REED, L. J., *On the rate of growth of the population of the United States since 1790 and its mathematical representation*, "National Academy of Sciences", vol. 6, págs. 275-288, 1920. Id., *On the Mathematical Theory of Population Growth.* "Metron", vol. III, núm. 1, págs. 6-19, 1923.

Para resolver el problema construyeron una teoría matemática del crecimiento de la población basada en los siguientes principios fundamentales:

I. La población crece dentro de un espacio finito y perfectamente definido.

II. El número máximo de seres que pueden habitar ese espacio finito es también finito.

III. El límite inferior que puede tener una población es cero (no se conciben poblaciones con un número negativo de habitantes).

IV. Históricamente es comprobable que cada avance en el nivel cultural ha traído consigo la posibilidad de un crecimiento adicional de población dentro de cualquier área definida. O, con otras palabras, en cada unidad geográfica que ha sido habitada durante un largo período de tiempo se han sucedido períodos de crecimiento de la población, cada uno de ellos superpuesto al anterior, que marcan, aproximadamente, la duración de la época cultural definida.

V. Dentro de cada época cultural o ciclo de crecimiento de la población, la tasa de crecimiento no ha sido constante en el tiempo. Al principio, la población aumenta suavemente, pero la tasa crece de manera constante hasta un cierto valor en que alcanza su máximo. En ese momento (punto de inflexión de la curva de crecimiento) puede considerarse que se encuentra la relación óptima entre el número de habitantes y los recursos del área definida. Después de este punto, la tasa de crecimiento se hace sucesivamente más baja, hasta que finalmente la curva se comporta como una recta horizontal, es decir, es asintótica a una recta paralela al eje de abscisas (tiempo), definida por la propiedad II.

Las tres primeras propiedades, de las cinco formuladas por Pearl y Reed, se pueden aplicar al estudio del crecimiento de cualquier población de seres vivos, mientras que las otras dos, y especialmente la cuarta, sólo son válidas en el caso de poblaciones humanas. De acuerdo con lo expuesto, si llamamos Y a la variable estadística "número de habitantes", ésta habrá de satisfacer las siguientes condiciones:

a) $$0 \leqslant Y \leqslant k,$$

en donde k es el número máximo definido por la propiedad II, y el límite inferior nulo es consecuencia de lo establecido en III;

$$b) \qquad \frac{dY}{dt} = CY\,(k - Y) \quad (C > 0),$$

ya que la forma de la curva, establecida en V, implica que la velocidad de crecimiento debe ser positiva, al variar Y de 0 a k. Como

$$\frac{d^2Y}{dt^2} = Ck - 2CY = C(k - 2Y),$$

el crecimiento es más que proporcional con el tiempo, hasta el instante en que $Y = \frac{k}{2}$ y para $Y > \frac{k}{2}$ la tasa de crecimiento disminuye, por ser $\frac{d^2Y}{dt^2} < 0$, o sea, por ser $Y = f\,(t)$ cóncava hacia abajo.

Podemos, pues, decir, con Morice y Chartier (2), que el crecimiento de la población es proporcional al camino ya recorrido Y por el que queda por recorrer k — Y. Esta propiedad la consideramos como la fundamental del problema logístico, ya que a partir de ella podemos obtener la ecuación de la denominada *curva logística*.

En efecto, como $b)$ es una ecuación diferencial de variables separables, tenemos:

$$\frac{dY}{Y\,(k - Y)} = C\,dt.$$

(2) MORICE, E. y F. CHARTIER, *Méthode Statistique* (I. N. S. E. E.), segunda parte, pág. 436. París, 1954.

e integrando los dos miembros,

$$\frac{1}{k} \cdot \int \left(\frac{1}{Y} + \frac{1}{k - Y} \right) dY =$$

$$= \frac{1}{k} [\log Y - \log (k - Y)] = Ct + D$$

de donde,

$$\frac{Y}{k - Y} = e^{kCt+kD} = m\, e^{at}$$

que nos permite hallar la ecuación

$$Y = \frac{m\, k\, e^{at}}{1 + m\, e^{at}}$$

que podemos transformar en la

$$Y = \frac{k}{1 + \frac{1}{m} e^{-at}} = \frac{k}{1 + e^{b-at}}, \qquad [2]$$

cuya última expresión será para nosotros la de la curva logística.

Se comprueba fácilmente que tiene como asíntotas las rectas de ecuaciones $Y = 0$ e $Y = k$, y el punto de inflexión en el de coordenadas $\left(\frac{b}{a}, \frac{k}{2} \right)$ que es también un centro de simetría de la curva.

La curva logística de la población de España.

Con este título se publicó en un *Boletín de Estadística* (3) del año 1941, un trabajo de la entonces Dirección General de Estadís-

(3) DIRECCIÓN GENERAL DE ESTADÍSTICA, *La curva logística de la población de España.* "Boletín de Estadística", núm. 12, pág. 121, octubre-diciembre 1941.

tica, que es hasta ahora el último en el que se aplica la teoría expuesta a la población de España. El primero publicado sobre la materia lo fué en la "Revista de Estudios Políticos" de este Instituto, por José Vergara Doncel (4).

Estos trabajos difieren esencialmente en la ecuación de la curva logística empleada: mientras en el segundo se utiliza la función clásica [2], en el de la Dirección de Estadística se ajusta una curva de ecuación.

$$Y = h + \frac{K}{1 + e^{b-at}} \qquad [3]$$

que consideramos más adecuada al caso de un país viejo, como España, en donde, por razones históricas, han debido existir varios ciclos de crecimiento de población y, sobre todo, por estar de acuerdo con la propiedad IV, antes considerada.

El Sr. Vergara obtiene los parámetros de la logística apoyado en los Censos de población de los años 1797 y 1860 y en la población estimada del año 1923, con cuyos datos obtiene un límite de 46,970 millones de habitantes, cifra muy parecida a la deducida por la Dirección de Estadística, que es de 45,458 millones de habitantes. No sabemos cómo se ha determinado esta última cifra, pues para calcular el parámetro h se dice, textualmente, en el trabajo original: "y, una vez fijados los valores de la asíntota inferior compatibles con el principio esencial de la curva logística..." (5), pero lo que sí deducen, en ambos casos, es que dentro de los supuestos fundamentales de la tendencia logística, España podría alcanzar una población superior a los 45 millones de habitantes.

Don José Vergara reconoce él mismo la poca fiabilidad de sus estimaciones al tratar de justificar la hipótesis de que "nuestras poblaciones censales no siguen la ley logística" (6), de acuerdo con la teoría de Knibbs (éste intenta demostrar que el punto de vista de

---

(4) José Vergara Doncel, *El movimiento de la población de España*. "Revista de Estudios Políticos". Madrid, julio 1941, pág. 453.

(5) Trab. cit. (3), pág. 123.

(6) Trab. cit. (4), pág. 458.

Verhulst Pearl y Yule (7), acerca del crecimiento de la población, no está justificado) (8).

La Dirección General de Estadística da un ajuste mejor. El parámetro h, que define el límite inferior de la población de España, tiene un valor de 14,443, muy parecido al $h = 14{,}756$, que a continuación obtendremos por un procedimiento no inspirado en el anterior, ya que, como hemos dicho, desconocemos el que allí se siguió. No obstante, y al apoyarnos en informaciones censales más recientes, nuestra asíntota superior es de 36 millones de habitantes, solamente.

A continuación exponemos minuciosamente el proceso de cálculo que nosotros utilizamos.

Determinación de la asíntota superior.

TABLA I

t	Y	Δt	ΔY	y	$\frac{\Delta Y}{\Delta t}$
t_1	Y_1	$\Delta t_1 = t_2 - t_1$	$\Delta Y_1 = Y_2 - Y_1$	$y_1 = \frac{1}{2}(Y_1 + Y_2)$	$\frac{\Delta Y_1}{\Delta t_1}$
t_2	Y_2	$\Delta t_2 = t_3 - t_2$	$\Delta Y_2 = Y_3 - Y_2$	$y_2 = \frac{1}{2}(Y_2 + Y_3)$	$\frac{\Delta Y_2}{\Delta t_2}$
t_3	Y_3				
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
t_{i-1}	Y_{i-1}	$\Delta t_{i-1} = t_i - t_{i-1}$	$\Delta Y_{i-1} = Y_i - Y_{i-1}$	$y_{i-1} = \frac{1}{2}(Y_{i-1} + Y_i)$	$\frac{\Delta Y_{i-1}}{\Delta t_{i-1}}$
t_i	Y_i	$\Delta t_i = t_{i+1} - t_i$	$\Delta Y_i = Y_{i+1} - Y_i$	$y_i = \frac{1}{2}(Y_i + Y_{i+1})$	$\frac{\Delta Y_i}{\Delta t_i}$
t_{i+1}	Y_{i+1}				
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
t_{n-1}	Y_{n-1}	$\Delta t_{n-1} = t_n - t_{n-1}$	$\Delta Y_{n-1} = Y_n - Y_{n-1}$	$y_{n-1} = \frac{1}{2}(Y_{n-1} + Y_n)$	$\frac{\Delta Y_{n-1}}{\Delta t_{n-1}}$
t_n	Y_n				

(7) Yule, "Journal of the Royal Statistical Society", vol. LXXXVIII, part. I, pág. 22, January 1925.

(8) Knibbs, G. H., *The Growth of human populations, and the laws of their increase.* "Metron", vol. V, núm. 3, pág. 147, diciembre 1925.

La asíntota $Y = k$ de la curva logística se caracteriza por la propiedad fundamental de la curva, esto es, la velocidad de crecimiento de la población, en un instante t, es proporcional al producto del número de habitantes alcanzado hasta ese instante por la población que falta hasta llegar al máximo k, o sea,

$$\frac{dY}{dt} = C\,Y\,(k - Y)\,. \qquad [4]$$

Como podemos considerar que los valores de $\frac{dY}{dt}$ corresponden a la variación absoluta de la población por unidad de tiempo, ajustando una parábola de segundo orden, del tipo [4], a la distribución estadística bidimensional definida por las dos últimas columnas de la tabla I, podemos obtener una estimación de k, a partir de una serie cronológica formada por las cifras de los censos de población de que se disponga.

Consideremos en nuestro caso las cifras correspondientes a los censos oficiales de la población de nuestra Península e islas adyacentes publicadas por el Instituto Nacional de Estadística en su "Anuario Estadístico de España" (edición manual) del año 1955 (pág. 37). En la tabla II publicamos estos datos en unión de los resultados correspondientes a las variables estadísticas que figuran en la tabla I.

TABLA II

Años	t	Y (en miles de habitantes)	Δt	ΔY	y	$\frac{\Delta Y}{\Delta t}$
1857	0	15.455	3	190	15.550	63,3
1860	3	15.645	17	977	16.133	57,5
1877	20	16.622	10	928	17.086	92,8
1887	30	17.550	10	559	17.829	55,9
1897	40	18.109	3	485	18.351	161,7
1900	43	18.594	10	1.333	19.260	133,3
1910	53	19.927	10	1.376	20.615	137,6
1920	63	21.303	10	2.261	22.433	226,1
1930	73	23.564	10	2.314	24.721	231,4
1940	83	25.878	10	2.099	26.927	209,9
1950	93	27.977				

Si representamos en un sistema cartesiano la variable bidimensional $\left(y, \dfrac{\Delta Y}{\Delta t} \right)$ nos encontramos con que la variación de $\dfrac{\Delta Y}{\Delta t}$ no es muy regular, debido, posiblemente, a que los resultados de algunos censos anteriores a 1900 no son muy fiables. Para obviar esta dificultad hemos prescindido de las informaciones censales que nos ofrecían menos garantía, desde el punto de vista del decrecimiento de $\dfrac{\Delta Y}{\Delta t}$: las correspondientes a los años 1857, 1887, 1897 y 1910, con lo que manejaremos los resultados de la tabla III a los que corresponde una marcha normal según se observa en el diagrama dado por la línea poligonal B L M H I J.

TABLA III

Años	t	Y (en miles de habitantes)	Δt	ΔY	y	$\dfrac{\Delta Y}{\Delta t}$
1860	0	15.645	17	977	16.133	57,5
1877	17	16.622	23	1.972	17.608	85,7
1900	40	18.594	20	2.709	19.948	135,5
1920	60	21.303	10	2.261	22.433	226,1
1930	70	23.564	10	2.314	24.721	231,4
1940	80	25.878	10	2.099	26.927	209,9
1950	90	27.977				

Para ajustar la parábola de segundo orden [4], consideremos las nuevas variables

$$1.000\, u = y - 16.133, \quad v = \frac{\Delta Y}{\Delta t} - 57,5 .$$

Entonces, la ecuación de la línea que hemos de ajustar será del tipo

$$v = a u + b u^2$$

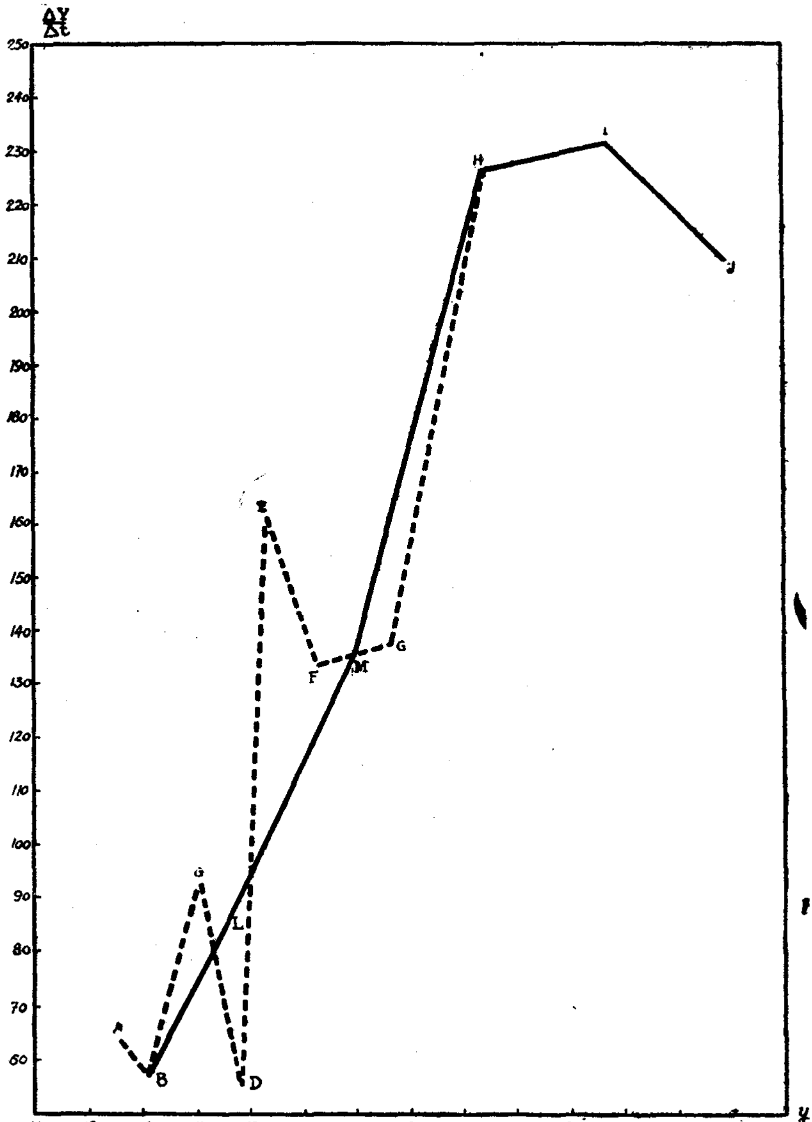

Gráfico núm. 1.

Si analizamos un ajuste mínimo-cuadrático, los valores de a y b se determinarán a partir del sistema de ecuaciones normales

$$a \sum u_i^2 + b \sum u_i^3 = \sum u_i v_i$$

$$a \sum u_i^3 + b \sum u_i^4 = \sum u_i^2 v_i$$

en el que, sustituyendo los valores obtenidos en la tabla IV, se obtiene la solución,

$$a = 33{,}455 \quad \text{y} \quad b = -\ 1{,}688 \,,$$

por lo que la parábola buscada es la

$$v = 33{,}455\, u - 1{,}688\, u^2 = 1{,}688\, u\, (19{,}819 - u) \,,$$

TABLA IV

u	v	
0	0	$\Sigma\, u\, v = 4.542{,}34$
1,5	28,2	$\Sigma\, u^2\, v = 38.519{,}084$
3,8	78,0	$\Sigma\, u^3 = 2.204{,}062$
6,3	168,6	$\Sigma\, u^4 = 20.863{,}834$
8,6	173,9	
10,8	152,4	

que es del tipo [4] con $k = 19{,}819$. Es decir, el tope máximo que puede alcanzar la población española en el actual ciclo evolutivo de la población es k más la población de partida utilizada en nuestro artificio, que era de 16,133 millones de habitantes.

En resumen, la asíntota superior Y = L de la curva logística de la población española, estimamos que vale:

$$Y = L = 16{,}133 + 19{,}819 = 35{,}952$$

millones de habitantes, por lo que tomaremos como valor aproximado el de

L = 36 millones de habitantes.

Ajuste de la curva logística.

La curva logística que vamos a ajustar es del tipo

$$Y = h + \frac{K}{1 + e^{b-at}}$$

en la que el límite de Y cuando t tiende a infinito es

$$h + K = L = 36.$$

De aquí se deduce que

$$Y - h = \frac{K}{1 + e^{b-at}}$$

o sea,

$$e^{b-at} = \frac{K - (Y - h)}{Y - h} = \frac{(K + h) - Y}{Y - h} = \frac{36 - Y}{Y - h}.$$

Eligiendo tres censos tales que el del centro sea equidistante en el tiempo de los censos de los extremos, es decir, en los momentos

$$-d,\ 0,\ d,$$

a los que corresponden las poblaciones censadas

$$Y_{-1},\ Y_0,\ Y_1$$

respectivamente, se tendrá

$$e^{b+ad} = \frac{36 - Y_{-1}}{Y_{-1} - h}, \qquad e^{b} = \frac{36 - Y_0}{Y_0 - h}, \qquad e^{b-ad} = \frac{36 - Y_1}{Y_1 - h}$$

y multiplicando miembro a miembro las ecuaciones de los extremos e igualando al cuadrado de la del centro,

$$\frac{36 - Y_{-1}}{Y_{-1} - h} \cdot \frac{36 - Y_1}{Y_1 - h} = \left(\frac{36 - Y}{Y_0 - h} \right)^2$$

tenemos una ecuación de segundo grado en *h*, que nos permite, en nuestro caso, calcular el valor de este parámetro.

Para resolver nuestro problema hemos considerado los censos de los dos años extremos que nos ofrecen garantía, o sea, los de los años 1860 y 1950; el año equidistante de estos dos es el 1905 en donde no se ha realizado censo, por cuya razón hemos de estimar su cifra y esto lo hemos hecho calculando la media geométrica de las poblaciones correspondientes a 1900 y 1910, es decir,

$$Y_{1.905} = \sqrt{Y_{1.900} \cdot Y_{1.910}} = \sqrt{18.594 \times 19.927} = 19.249 \, .$$

Por tanto, los valores que vamos a manejar son:

$$d = 45, \; Y_{-1} = 15{,}645, \; Y_0 = 19{,}249, \; Y_1 = 27{,}977$$

y la ecuación que permite determinar *h* toma la forma,

$$117{,}288 \, h^2 - 5.953{,}127 \, h + 62.307{,}330 = 0 \, .$$

Como la ecuación original en *h* se satisface para $h = 36$, ésta es divisible por $h - 36$, y efectuada la división queda

$$117{,}288 \, h - 1.730{,}759 = 0$$

de donde

$$h = 14{,}756$$

por lo que

$$K = 36 - 14{,}756 = 21{,}244 \, .$$

De la expresión,

$$e^{b} = \frac{36 - Y_0}{Y_0 - h} = \frac{16,751}{4,493} = 3,72824\,,$$

se deduce que

$$b = \log_e 3,72824 = 1,316\,,$$

y de la

$$e^{b-ad} = \frac{36 - Y_1}{Y_1 - h} = \frac{8,023}{13,221} = 0,60684\,,$$

de donde

$$b - ad = \log_e 0,60684 = -\,0,4995\,,$$

y de aquí

$$a = \frac{b + 0,4995}{d} = \frac{1,8155}{45} = 0,040\,344$$

por lo que, la ecuación de la curva logística de España, será

$$Y = 14,756 + \frac{21,244}{1 + e^{1,316 - 0,040\,344\,t}} \qquad [5]$$

con origen del tiempo en el año 1905.

Nos parece interesante destacar que la abscisa del punto de inflexión

$$t = \frac{b}{a} = \frac{1,316}{0,040\,344} = 32,6$$

corresponde al año

$$1905 + 32,6 = 1937,6 \text{ (7 de agosto de 1937)}\,,$$

es decir, en el centro de nuestro período bélico, lo que es posible que justifique el cambio de signo en la aceleración del crecimiento de la población de España. De todas formas, nosotros no tenemos autoridad científica para discutir este sugestivo problema que entra de lleno en el campo de los sociólogos y demógrafos.

Poblaciones calculadas.

Presentamos en la tabla V los valores teóricos de la población de España, en los años censales, en el supuesto de que ésta siguiese la ley logística que hemos deducido, y previsiones para algunos años posteriores al 1950.

TABLA V

Años	Población censada o estimada (*) Y	Población calculada $\overline{Y}$
1800	—	14,838
1857	15,455	15,547
1860	15,645	15,645
1877	16,622	16,450
1887	17,550	17,196
1897	18,109	18,210
1900	18,594	18,575
1905	19,249 (*)	19,249
1910	19,927	20,005
1920	21,303	21,753
1930	23,564	23,762
1940	25,878	25,890
1950	27,977	27,977
1955	—	28,957
1960	—	29,876
1965	—	30,717
1970	—	31,468
1980	—	32,742
1990	—	33,615
2000	—	34,408
2100	—	35,970

Puede observarse, tanto por la inspección de esta tabla como por el gráfico número 2, que sólo hay dos desviaciones de cierta cuantía, entre los valores teóricos y observados, los correspondientes a los años 1887 y 1920. Estas desviaciones son de 0,354 y — 0,450 millones de habitantes, o sea, de un 2 por 100 de la población censada, en ambos casos.

La desviación obtenida en el año 1920 sospechamos que se debe a la influencia de la "gripe" del año 1918 que, además del impacto directo sobre la población, produjo una disminución de la nup-

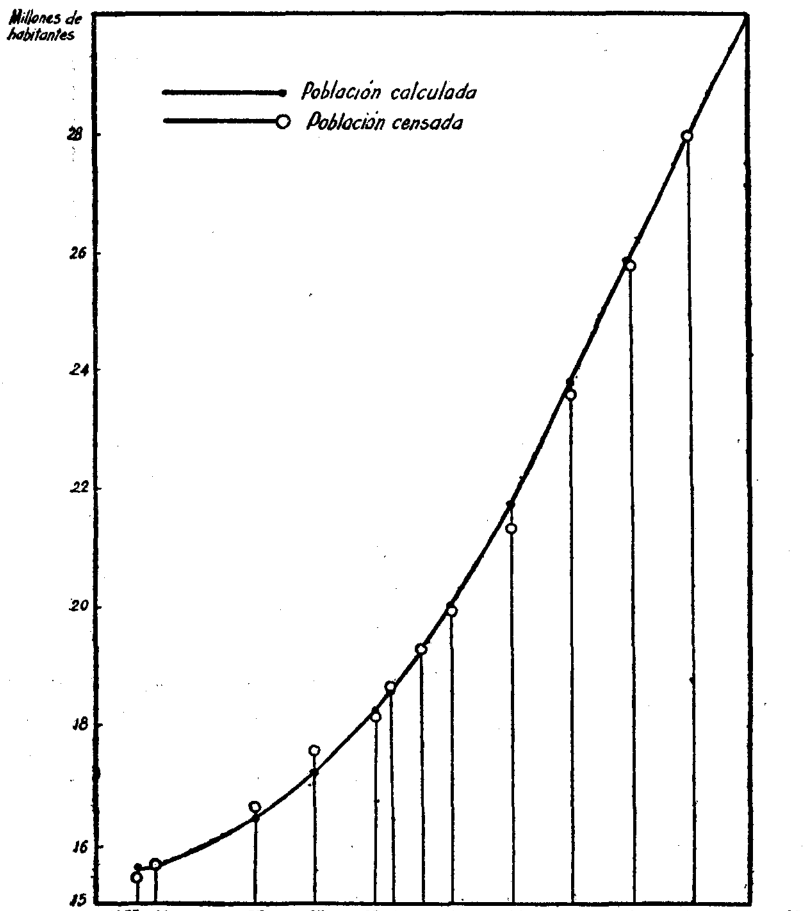

Gráfico núm. 2.

cialidad y la natalidad. La desviación correspondiente al año 1887 parece que se debe a falta de veracidad de las cifras del censo, como puede comprobarse observando los valores de la última columna de la tabla II.

Estas previsiones obtenidas nos han permitido estimar poblaciones teóricas, en el intervalo de tiempo 1857-1950, con un error relativo menor o igual al 2 por 100 de la población verdadera. Ahora bien, esto no quiere decir que las poblaciones calculadas en los años siguientes al 1950 lo sean también con un error menor o

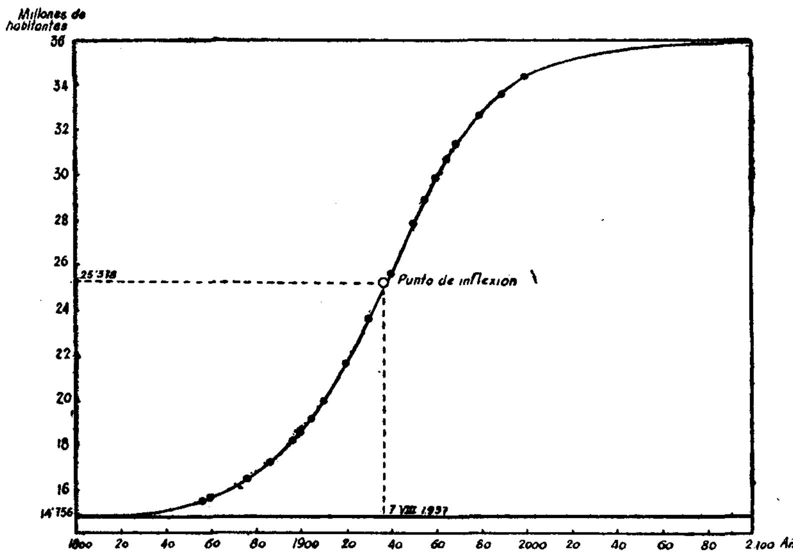

Gráfico núm. 3.

igual del 2 por 100, y ni siquiera estamos en condiciones de poder hacer proposiciones en términos de probabilidad del error que pueda haberse cometido en el cálculo de estas poblaciones futuras. Lo único que podemos afirmar es que si aceptamos la tendencia logística en la población de España, y, a partir de la información de que disponemos, nuestras previsiones de poblaciones futuras son las ordenadas de la curva dibujada en los gráficos números 2 y 3 o las obtenidas dando valores a t en la ecuación [5], pero sin poder decir nada más respecto a la fiabilidad de estas estimaciones.

Las conclusiones de este trabajo que, a nuestro parecer, tienen más interés, son las siguientes:

a) La población máxima de España es de 36 millones de habitantes, dentro del ciclo actual de crecimiento de la población.

b) La velocidad de crecimiento de la población de España ha ido en aumento hasta el 7 de agosto del año 1937, y a partir de ese día ha empezado a disminuir.

ANGEL ALCAIDE INCHAUSTI