Ajuste de una función expologística a la evolución de la población total de México, 1930-1985

Manuel Ordorica Mellado*

En la investigación demográfica se ha desarrollado un gran número de funciones matemáticas con el fin de analizar la dinámica de la población total, entre las que destacan la función exponencial, la curva de Gompertz y la logística por sólo mencionar unas cuantas. Sin embargo, ninguna de estas funciones se ajusta en forma adecuada ai caso de México, debido a que los supuestos que subyacen a las representaciones matemáticas antes mencionadas no corresponden a la dinámica observada de los componentes del crecimiento natural. El propósito del presente trabajo es ajustar una función matemática a la evolución de la población total de México entre 1930-1985, que reproduzca en forma adecuada la evolución de la natalidad y la mortalidad observadas en el periodo mencionado.

Introducción

El propósito del presente trabajo es ajustar una función matemática a la evolución de la población total de México entre 1930 y 1985.

En la investigación demográfica se ha desarrollado un gran número de funciones matemáticas con el fin de analizar la dinámica de la población total, entre las que destacan: la función exponencial, la curva de Gompertz y la logística, sólo por mencionar unas cuantas. Estos modelos matemáticos han permitido estudiar y describir la evolución de la población de algunos países en determinados momentos de su historia. Sin embargo, ninguna de estas funciones se ajusta en forma adecuada al caso de México, debido a que los supuestos que subyacen a las representaciones matemáticas antes mencionadas no corresponden a la dinámica observada de los componentes del crecimiento natural.

La función exponencial ($P(t) = P(o)e^{rt}$, donde $P(t)$ es la población en el instante t, $P(o)$ es la población en el momento o, r es la

* Coordinador de los programas de maestría y doctorado en población del Centro de Estudios Demográficos y de Desarrollo Urbano de El Colegio de México.

tasa anual de crecimiento demográfico y t es el tiempo) supone que las tasas de natalidad, de mortalidad y, en consecuencia, la de crecimiento demográfico permanecen invariables en el tiempo. A partir de estos supuestos se genera una ecuación diferencial como la siguiente:

$$\frac{1}{P(t)} \frac{dP(t)}{dt} = r = \text{constante}$$

La función de Gompertz ($P(t) = e^{b/m} e^{-ce^{-mt}}$ en donde c es una constante) se obtiene a partir de la siguiente ecuación diferencial:

$$\frac{dP(t)}{dt} = P(t)(b - m \ln P(t))$$

siendo b y m >0. La constante b es la tasa de natalidad y la constante m representa la tasa de mortalidad. Esta última es proporcional al $\ln(P(t))$.

$$\text{La logística } (P(t) = \frac{bP(o)}{mP(o) + (b - mP(o))e^{-bt}})$$

construida por el matemático y biólogo belga P. F. Verhulst, se obtiene a partir de la siguiente ecuación diferencial:

$$\frac{dP(t)}{dt} = P(t)(b - mP(t))$$

en donde b y m > 0. La constante b es la tasa de natalidad y la constante m es la tasa de mortalidad, la cual es proporcional a la población $P(t)$.

Uno de los primeros en aplicar una función matemática a la evolución de la población fue Quetelet en el año de 1835. Él señaló que la población se incrementa a ritmo acelerado hasta llegar a un punto en que empieza a crecer más lentamente.

Verhulst analizó este planteamiento y sugirió una curva teórica a la que llamó logística. Esta obra fue olvidada por muchos años hasta que en 1920 Pearl y Reed redescubrieron la función logística.

En la logística se supone que la tasa de natalidad o la tasa de mortalidad permanecen constantes y el otro componente del crecimiento demográfico desciende en forma lineal.

La función logística y la exponencial se pueden generar de otra forma (Lotka, 1969).

$$\text{Sea } \Omega\,(P(t)) = \frac{dP(t)}{dt}$$

Esta función se puede desarrollar como una serie de potencias a través de la serie de Taylor.

$$\Omega(P(t)) = \frac{dP(t)}{dt} = \sum_{n=1}^{\infty} r_n P^n(t)$$

donde $\frac{dP(t)}{dt}$ es la densidad anual de aumento o disminución en t.

La función exponencial y la logística pueden ser consideradas como un caso particular del desarrollo en serie del incremento anual instantáneo de la población. La función

$$\frac{dP(t)}{dt}$$

es una función de la población, digamos $\Omega(P(t))$; por tanto, se puede desarrollar según la fórmula de Taylor:

$$\frac{dP(t)}{dt} = \Omega(P(t)) = r_1P(t) + r_2P^2(t) + r_3P^3(t) + \ldots$$

Cuando la serie queda reducida a un solo término $(\frac{dP(t)}{dt} = r_1P(t))$ se genera la exponencial. En el caso de que se conserven los dos primeros sumandos $\frac{dP(t)}{dt} = r_1P(t) + r_2P^2(t))$ se obtiene la logística.

En este caso se supone que la tasa de natalidad b(t) es una función lineal decreciente de la población P, de tal manera que

$$b(t) = b_0 + b_1P(t).$$

donde b_0 y $b_1 > 0$ y la tasa de mortalidad $m(t) = m_0$ permanece constante en el tiempo. Podría suponerse que la tasa de natalidad permanece invariable y que la tasa de mortalidad desciende en forma lineal. En este caso, también se genera una función logística.

En todas estas formulaciones, los supuestos respecto a la evolución real de los componentes del crecimiento natural se alejan de la dinámica demográfica observada en México durante el periodo 1930-1985.

¿Cómo han descendido los componentes del aumento vegetativo de la población en México durante el periodo mencionado?

En el cuadro 1 se presentan las series de las tasas de natalidad y de mortalidad:

CUADRO 1
Tasas de natalidad y de mortalidad

Años	*Tasa de natalidad (por mil)*	*Tasa de mortalidad (por mil)*
1930-1934	49.4[1]	25.6[4]
1940-1944	44.3[1]	22.0[4]
1950-1954	46.7[2]	16.2[2]
1960-1964	44.9[2]	11.3[2]
1965-1969	44.2[2]	10.2[2]
1970-1974	42.7[2]	9.2[2]
1975-1979	37.6[2]	7.9[2]
1980-1984	33.9[3]	7.1[3]
1985-1989	31.2[3]	6.5[3]

[1] *Dinámica de la población de México*, El Colegio de México, CEED, 1970, p. 47. Datos referidos a 1930 y 1940.
[2] *México: Estimaciones y proyecciones de población, 1950-2000*, SPP, Conapo, Celade, Naciones Unidas, 1983, p. 11.
[3] *Perspectivas de la población mundial. Estimaciones y proyecciones en 1982*. Nueva York, Naciones Unidas, 1985, p. 353.
[4] *Dinámica de la población de México*, El Colegio de México, CEED, 1970, p. 14.

Como se puede observar en el cuadro anterior, la tasa de natalidad permanece en niveles altos hasta el año de 1965 y empieza a descender en el segundo quinquenio del decenio de los sesenta. En una primera etapa, la declinación es acelerada y posteriormente se observa un descenso menos rápido.

Por lo que se refiere a los niveles de mortalidad, éstos disminuyen rápidamente en un principio y luego se empieza a observar una desaceleración en la caída a partir de 1960.

¿Qué función matemática podría describir la evolución de las tasas de natalidad y de las tasas de mortalidad?

Al observar la evolución de los componentes del crecimiento natural de la población, podríamos intentar el ajuste de una función logística a cada una de las variables.

Ajuste

$$\text{Sea } b(t) = k'_1 + \frac{k'_2}{1 + e^{a+bt}}$$

donde $b(t)$ representa la tasa de natalidad en función del tiempo, $k'_1 + k'_2$ es la asíntota superior, k'_1 es la asíntota inferior, a y b son los parámetros a estimar y e es igual a 2.7182. . .

Desarrollando la función anterior tenemos que

$$b(t) - k'_1 = \frac{k'_2}{1 + e^{a+bt}}$$

Reagrupando

$$1 + e^{a+bt} = \frac{k'_2}{b(t) - k'_1}$$

y

$$e^{a+bt} = \frac{k'_2}{b(t) - k'_1} - 1$$

Por lo que

$$a + bt = L\left[\frac{k'_2}{b(t) - k'_1} - 1\right]$$

Haciendo

$$L\left[\frac{k'_2}{b(t) - k'_1} - 1\right] = Y(t)$$

Tenemos que

$$Y(t) = a + b \cdot t$$

A partir de esta ecuación es posible estimar los parámetros a y b, utilizando el método de regresión lineal simple:

$$Y(t) = a + b \cdot t + e_t$$

El procedimiento matemático descrito para la fecundidad se realiza de la misma manera para la mortalidad, a la cual denotaremos con el símbolo m(t).

Resultados del ajuste de los componentes del crecimiento natural

Los modelos obtenidos son los siguientes:

Para la mortalidad (m(t)):

$$m(t) = .004 + \frac{.026}{1 + e^{-1.41777 + .06998t}}$$

donde k'_1 es igual a 4 por mil y k'_2 es 26 por mil. El coeficiente de determinación (R^2) calculado es igual a .986.

Para la natalidad (b(t)):

$$b(t) = .010 + \frac{.040}{1 + e^{-3.53927 + .059483t}}$$

donde k'_1 es igual a 10 por mil y k'_2 es 40 por mil. El coeficiente de determinación (R^2) calculado es igual a .827.

En ambos ajustes es posible destacar el elevado coeficiente de correlación calculado.

¿Cómo determinar una función de población que cumpla con las representaciones estimadas para la evolución de la tasa de natalidad y de la tasa de mortalidad?

Ajuste de los datos de población a una función matemática

Definamos los siguientes términos:

P (t) es la población en el momento t; t es la variable tiempo.

$\frac{dP\ (t)}{dt}$ es el incremento anual instantáneo de población.

$k_{1,\beta}$ es la asíntota inferior de la tasa de natalidad.

$k_{1,\beta} + k_{2,\beta}$ es la asíntota superior de la tasa de natalidad.

$k_{1,\mu}$ es la asíntota inferior de la tasa de mortalidad.

$k_{1,\mu} + k_{2,\mu}$ es la asíntota superior de la tasa de mortalidad.

β_1, β_2, μ_1 y μ_2 son parámetros: β_1 y μ_1 representan el nivel de la natalidad y de la mortalidad, respectivamente; β_2 y μ_2 representan la velocidad de la caída de la natalidad y de la mortalidad, respectivamente.

La tasa anual instantánea de crecimiento de la población puede ser representada de la siguiente manera:

$$\frac{1}{P\,(t)}\;\frac{dP\,(t)}{dt} = \mathrm{b}(t) - \mathrm{m}(t) \tag{1}$$

En el caso de que b(t) y m(t) sean funciones logísticas, es posible expresar la tasa de crecimiento de la población de la siguiente manera:

$$\frac{1}{P\,(t)}\;\frac{dP\,(t)}{dT} = \underbrace{k_{1,\beta} + \frac{k_{2,\beta}}{1+e^{\beta_1+\beta_2 t}}}_{\mathrm{b}(t)} - \underbrace{(k_{1,\mu} + \frac{k_{2,\mu}}{1+e^{\mu_1+\mu_2 t}})}_{m(t)} \tag{2}$$

Se sabe que

$$\frac{1}{P\,(t)}\;\frac{dP\,(t)}{dt} = \frac{dLP\,(t)}{dt} \text{ entonces}$$

$$\frac{dLP(t)}{dt} = k_{1,\beta} + \frac{k_{2,\beta}}{1+e^{\beta_1+\beta_2 t}} - (k_{1,\mu} + \frac{k_{2,\mu}}{1+e^{\mu_1+\mu_2 t}}) \tag{3}$$

Integrando $\frac{dLP(t)}{dt}$ se tiene que $LP\,(t)$ es igual a

$$k_{1,\beta}t + \int \frac{k_{2,\beta}}{1+e^{\beta_1+\beta_2 t}}\,dt - (k_{1,\mu}\,t + \int \frac{k_{2,\mu}}{1+e^{\mu_1+\mu_2 t}}\,dt) + C \tag{4}$$

Reagrupando términos se tiene que $LP\,(t)$

$$(k_{1,\beta} - k_{1,\mu})\,t + \int \frac{k_{2,\beta}}{1+e^{\beta_1+\beta_2 t}}\,dt - \int \frac{k_{2,\mu}}{1+e^{\mu_1+\mu_2 t}}\,dt) + C \tag{5}$$

se tiene que

$$* \int \frac{k_{2,\beta}}{1+e^{\beta_1+\beta_2 t}}\,dt = k_{2,\beta}(1/\beta_2(\beta_2 t - L\;(1+e^{\beta_1+\beta_2 t}))) + C$$

y

$$* \int \frac{k_{2,\mu}}{1+e^{\mu_1+\mu_2 t}}\,dt = k_{2,\mu}(1/\mu_2(\mu_2 t - L\;(1+e^{\mu_1+\mu_2 t}))) + C$$

* Véase el apéndice.

Por tanto

$$LP\ (t) = (k_{1,\beta}-k_{1,\mu})t + k_{2,\beta}(1/\beta_2(\beta_2 t - L(1 + e^{\beta_1 + \beta_2 t}))) - k_{2,\mu}(1/\mu_2(\mu_2 t - L(1 + e^{\mu_1 + \mu_2 t}))) + C \quad (6)$$

y

$$P\ (t) = e^{(k_{1,\beta}-k_{1,\mu})t} + k_{2,\beta}(1/\beta_2(\beta_2 t - L(1 + e^{\beta_1 + \beta_2 t}))) - k_{2,\mu}(1/\mu_2(\mu_2 t - L(1 + e^{\mu_1 + \mu_2 t}))) + C \quad (7)$$

O sea

$$P\ (t) = e^{(k_{1,\beta}-k_{1,\mu})t} \cdot e^{(k_{2,\beta}-k_{2,\mu})t} \frac{(1 + e^{\mu_1 + \mu_2 t})^{k_{2,\mu}/\mu_2}}{(1 + e^{\beta_1 + \beta_2 t})^{k_{2,\beta}/\beta_2}}\ e^{c}$$

Reagrupando se tiene que

$$P\ (t) = e^{(k_{1,\beta}+k_{2,\beta}-(k_{1,\mu}+k_{2,\mu}))t} \frac{(1 + e^{\mu_1 + \mu_2 t})^{k_{2,\mu}/\mu_2}}{(1 + e^{\beta_1 + \beta_2 t})^{k_{2,\beta}/\beta_2}}\ e^{c} \quad (9)$$

Haciendo $t = 0$ se obtiene lo siguiente:

$$P\ (0) = \frac{(1 + e^{\mu_1})^{k_{2,\mu}/\mu_2)}}{(1 + e^{\beta_1})^{k_{2,\beta}/\beta_2}}\ e^{c} \quad (10)$$

Por tanto

$$e^{c} = P\ (0)\frac{(1 + e^{\beta_1})^{k_{2,\beta}/\beta_2}}{(1 + e^{\mu_1})^{k_{2,\mu}/\mu_2}} \quad (11)$$

Sustituyendo (11) en (9) se obtiene la ecuación final siguiente:

$$P\ (t) = P(0)\ e^{(k_{1,\beta}+k_{2,\beta}-(k_{1,\mu}+k_{2,\mu}))t} \frac{(1 + e^{\mu_1 + \mu_2 t})^{k_{2,\mu}/\mu_2}}{(1 + e^{\beta_1 + \beta_2 t})^{k_{2,\beta}/\beta_2}} \cdot \frac{(1 + e^{\beta_1})^{k_{2,\beta}/\beta_2}}{(1 + e^{\mu_1})^{k_{2,\mu}/\mu_2}} \quad (12)$$

En este modelo se supone que *la* población es cerrada

¿Qué características tiene la función construida **bajo** las **hipótesis** señaladas?

La función, a la que he llamado *expologística* está conformada de la siguiente manera:

— Dos constantes que son: la población inicial (P(0)) y un cociente de dos logísticas evaluadas en el momento 0. Esta última constante representa la relación entre la tasa de mortalidad y la de natalidad en el instante 0.

$$\frac{(1+e^{\beta_1})^{k_{2,\beta}/\beta_2}}{(1+e^{\mu_1})^{k_{2,\mu}/\mu_2}} =$$

$$= \frac{\dfrac{1}{(1+e^{\mu_1})^{k_{2,\mu}/\mu_2}}}{\dfrac{1}{(1+e^{\beta_1})^{k_{2,\beta}/\beta_2}}} \quad \begin{array}{l}\text{(logística de la mortalidad)}\\ \\ \text{(logística de la natalidad)}\end{array}$$

—Una función exponencial $[e^{(k_{1,\beta}+k_{2,\beta}-(k_{1,\mu}+k_{2,\mu}))t}]$ que representa la diferencia de la asíntota superior de la tasa de natalidad menos la asíntota superior de la tasa de mortalidad.

—Un cociente de 2 logísticas elevadas a una constante. Dicha constante es diferente en cada caso. Ambas logísticas son funciones del tiempo.

$$\frac{(1+e^{\mu_1+\mu_2 t})^{k_{2,\mu}/\mu_2}}{(1+e^{\beta_1+\beta_2 t})^{k_{2,\beta}/\beta_2}} =$$

$$= \frac{\dfrac{1}{(1+e^{\beta_1+\beta_2 t})^{k_{2,\beta}/\beta_2}}}{\dfrac{1}{(1+e^{\mu_1+\mu_2 t})^{k_{2,\mu}/\mu_2}}} \quad \begin{array}{l}\text{(logística de la natalidad)}\\ \\ \text{(logística de la mortalidad)}\end{array}$$

En el cuadro 2 se presentan los resultados de las diferentes funciones:

Cada una de las funciones que aparecen en la primera, tercera y cuarta columnas tienden a infinito, cuando t tiende a infinito. La quinta columna es la inversa de la función de la cuarta columna. Esta función actúa como freno a la nueva ecuación de $P(t)$. Si bien la función de la primera y tercera columnas crecen rápidamente, la función

$$\frac{1}{(1+e^{-3.539275+.059483t})^{.040/.059483}}$$

detiene el crecimiento de la función $P(t)$.

CUADRO 2
Resultados de las funciones

Años	(1) $e^{(.010+.040-(.004+.026))t}$	(2) $\frac{(1+e^{-3.53927})^{.040/.059483}}{(1+e^{-1.41777})^{.026/.069980}}$
1930	1.000000000	.939875534
1940	1.221402758	.939875534
1950	1.491824698	.939875534
1960	1.822118800	.939875534
1970	2.225540928	.939875534
1980	2.718281828	.939875534
1990	3.320116923	.939875534
2000	4.055199967	.939875534

Años	(3) $(1+e^{-1.41777+.06998t})^{.026/.06998}$	(4) $(1+e^{-3.539275+.059483t})^{.040/.059483}$
1930	1.083932874	1.019432914
1940	1.159044030	1.035094480
1950	1.289388643	1.063194742
1960	1.499790465	1.113234330
1970	1.815783312	1.201268324
1980	2.263553514	1.353371369
1990	2.873669008	1.609909676
2000	3.685894558	2.030680958

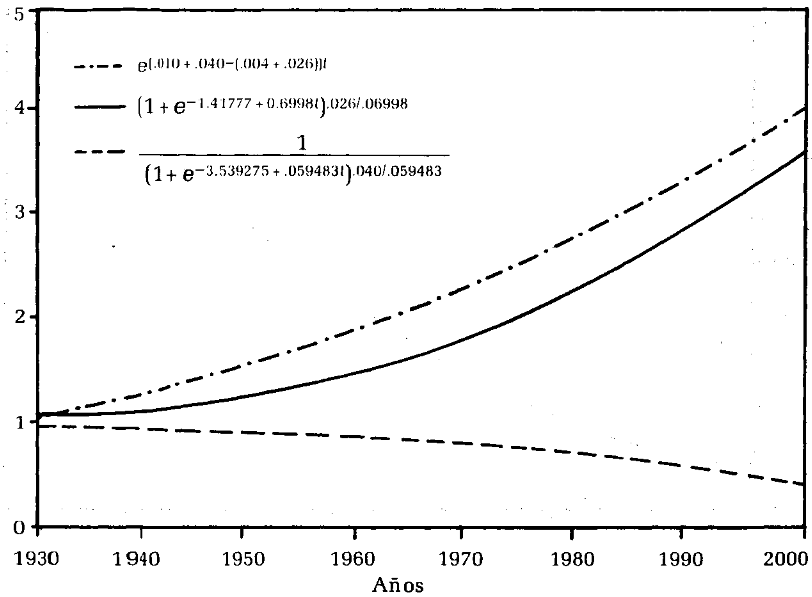
Años	(5) $\frac{1}{(1+e^{-3.539275+0.594831t})^{.040/.059483}}$
1930	.980937525
1940	.966095384
1950	.940561461
1960	.898283473
1970	.832453483
1980	.738895489
1990	.621152860
2000	.492445647

Resultados y restricciones

En el cuadro 3 y la gráfica 2 se presenta el ajuste obtenido, partiendo de una población inicial para 1930 de 17 063 300 habitantes (Alba, 1984:17).

Las estimaciones obtenidas son superiores a los resultados de las diferentes proyecciones de población realizadas en épocas recientes por los demógrafos mexicanos. Este hecho es normal, en

GRÁFICA 1
EUM 1930-2000: diferentes funciones matemáticas

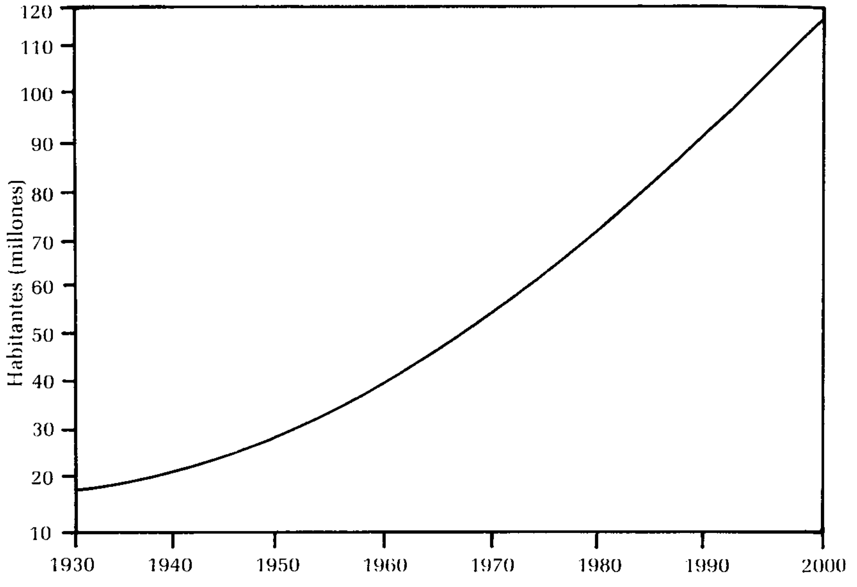


CUADRO 3
Ajuste

Año	*Población (P(t))* *(en millones)*
1930	17.1
1940	21.9
1950	29.0
1960	39.4
1970	53.9
1980	72.9
1990	95.0
2000	118.0

la medida en que el modelo aquí trabajado supone una población cerrada. No obstante, sería fácil incorporar en el modelo la migración internacional. Sin embargo, el objetivo del trabajo ha sido construir uno nueva función matemática para describir la dinámica de la población.

GRÁFICA 2
EUM 1930-2000: población total estimada



Es importante señalar que en el ajuste del modelo se presenta un problema estadístico. El modelo se ha ajustado con pocos datos, por lo que se cuenta con sólo 7 grados de libertad en las regresiones calculadas.

Una desventaja del modelo es que no incorpora la estructura por edad de la población. Al incorporar este elemento básico en el modelo, la matemática se complica demasiado.

Una bondad de este modelo es que es posible trabajar matemáticamente la Teoría de la Transición Demográfica entre otros enfoques teóricos, es decir, es posible describir matemáticamente los cambios ocurridos en los componentes del crecimiento de la población, los que a su vez se traducirán en efectos sobre la función expologística.

Asimismo, puede representar un instrumento pertinente para realizar lo que se ha dado en llamar "conciliación intercensal", y de esta manera ser un elemento indispensable en la elaboración de proyecciones de población, en donde queden claramente establecidas las hipótesis a partir de los componentes demográficos. Este modele permite ir más allá de las proyecciones mecánicas que muchas veces no sabemos qué supuestos tienen implícitos y tampoco disponemos de elementos objetivos para realizar la evaluación de resultados.

Apéndice (desarrollo matemático)

$$k_{2,\beta}\int \frac{1}{1+e^{\beta_1+\beta_2 t}}\,dt = k_{2,\beta}\int \frac{e^{-\beta_1-\beta_2 t}}{1+e^{-\beta_1-\beta_2 t}}\,dt =$$

Sea $u = 1 + e^{-\beta_1-\beta_2 t}$

$du = -\beta_2 e^{-\beta_1-\beta_2 t}\,dt$

entonces

$$k_{2,\beta}\int \frac{e^{-\beta_1-\beta_2 t}(-\beta_2)}{(1+e^{-\beta_1-\beta_2 t})(-\beta_2)}\,dt =$$

$$\frac{k_{2,\beta}}{-\beta_2}\cdot\int\frac{du}{u} =$$

$$\frac{k_{2,\beta}}{-\beta_2}\,Lu + C =$$

$$\frac{k_{2,\beta}}{-\beta_2}\,L(1+e^{-\beta_2-\beta_2 t}) + C =$$

$$\frac{k_{2,\beta}}{-\beta_2}\,L(1+e^{\beta_1+\beta_2 t}) + C =$$

$$\frac{k_{2,\beta}}{-\beta_2}\,L\left(\frac{1+e^{\beta_1+\beta_2 t}}{e^{\beta_1+\beta_2 t}}\right) + C =$$

$$\frac{k_{2,\beta}}{\beta_2}\,L\left(\frac{e^{\beta_1+\beta_2 t}}{1+e^{\beta_1+\beta_2 t}}\right) + C =$$

$$\frac{k_{2,\beta}}{\beta_2}\,(L(e^{\beta_1+\beta_2 t}) - L(1+e^{\beta_1+\beta_2 t})) + C =$$

$$\frac{k_{2,\beta}}{\beta_2}\,((\beta_1+\beta_2 t) - L(1+e^{\beta_1+\beta_2 t})) + C =$$

$$\frac{k_{2,\beta}}{\beta_2}\cdot\beta_1 = \text{constante} \therefore$$

$$\frac{k_{2,\beta}}{\beta_2}\cdot(\beta_2 t - L(1+e^{\beta_1+\beta_2 t})) + \mathbf{c}$$

Bibliografía

Alba, Francisco (1984). *La población de México: evolución y dilemas*, México, El Colegio de México, CEDDU.
G. Zill, Dennis (1988). *Ecuaciones diferenciales con aplicaciones*, Grupo Editorial Iberoamérica, segunda edición.
Keyfitz, Nathan (1979). *Introducción a las matemáticas de población*, Santiago de Chile, Centro Latinoamericano de Demografía.
Lotka, Alfred (1969). *Teoría analítica de las asociaciones biológicas*, Santiago de Chile, Centro Latinoamericano de Demografía.
Naciones Unidas (1970). "El concepto de población estable. Aplicación al estudio de la población de países que no tienen buenas estadísticas demográficas", Nueva York (ST/SOA/SER.A/39).