MODELO LOGÍSTICO ADAPTADO A LA POBLACIÓN MEXICANA EN EL PERIODO 1900-1990

M. en C. Jaime Salvador Medina González*
José de Jesús Suárez Hernández**

INTRODUCCIÓN

Una de las aplicaciones de las ecuaciones diferenciales es el análisis del crecimiento de poblaciones de seres vivos; en particular, de poblaciones humanas. En el presente trabajo se pretende mostrar que el cambio en el tamaño de la población mexicana, en el periodo 1900-1990, tiene un comportamiento como el que establece la ley logística de crecimiento poblacional.

Por otro lado, actualmente existe la necesidad de obtener parámetros de referencia que permitan evaluar la cobertura de eventos censales futuros; de ahí la importancia de encontrar un modelo que proporcione algún criterio de evaluación, un modelo que describa el cambio en el tamaño de la población a través del tiempo y que sea lo más confiable posible.

JUSTIFICACIÓN DEL MODELO

Antecedentes

Varios autores se han dedicado al análisis del crecimiento de poblaciones de seres vivos y en particular de poblaciones humanas. Hagamos una revisión histórica.

En propias palabras de Malthus en su ensayo sobre el principio de la población 1798: La población, si no encuentra obstáculos, aumenta en progresión geométrica, mientras que los alimentos lo hacen en progresión aritmética.

Quetelet fue uno de los primeros en utilizar un modelo matemático (1835) para explicar la evolución de la población, señalando que la población se incrementa a ritmo acelerado hasta llegar a un punto en que empieza a crecer más lentamente.

Como influencia de las teorías de Malthus en lo referente al desequilibrio entre el incremento de la población y el de los alimentos, Darwin supuso (1838) la necesidad de los seres vivos de competir para obtener suficiente alimento que les garantice la mejor situación posible en el entorno.

Verhulst analizó este planteamiento y sugirió una curva teórica a la que llamó logística, la cual dio a conocer en 1837, y actualmente es conocida como Ley logística del crecimiento de la población. Usando esta ley encontró dos modelos, con los cuales pronosticó para Bélgica una población máxima de 6 millones y para Francia de 40 millones. En 1930, la población de Francia concordaba sorprendentemente con lo obtenido por Verhulst (1).

En 1920 Pearl y Reed la utilizaron para explicar el comportamiento del crecimiento de la población en Estados Unidos (1)

MARCO CONCEPTUAL

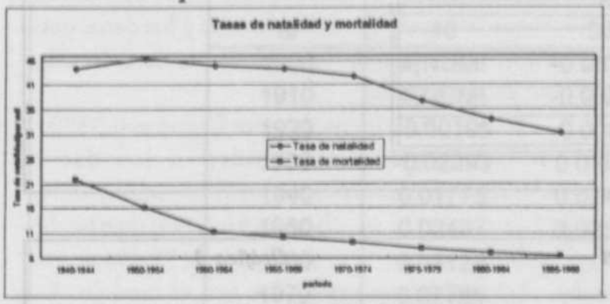
Población cerrada. El modelo logístico supone una población cerrada, es decir, sin migración. En este sentido, una población cerrada es aquella cuyo crecimiento es resultado de la fecundidad, mortalidad y otros factores como la competencia por los alimentos, pero no de la entrada o salida de sus integrantes del espacio que la delimita.

Teoría de transición demográfica. La transición demográfica es el cambio que sufre una población al pasar de niveles altos en fecundidad y mortalidad a niveles bajos, pero con un cierto desfasamiento. Algunos autores cuestionan que se trate de una teoría, argumentando que es simplemente una descripción de los cambios que sucedieron en los países en vías de desarrollo en los siglos XVIII y XIX. Veamos cómo se ha presentado esta característica en el caso de México. En la época de la Revolución Mexicana la población permaneció, en volumen, casi constante durante más de veinte años, debido a una mortalidad casi igual a la natalidad. Al término de dicho período la población

*Profesor-Investigador del Departamento de Matemáticas y Física. Jsmedina@correo.uaa.mx
**Estudiante de 8° Semestre de Matemáticas Aplicadas ma960021@correo.uaa.mx

empezó a aumentar su velocidad de crecimiento, esta velocidad permaneció en la década de los sesentas y principios de los setentas cuando aun políticamente se impulsó el crecimiento del tamaño de las familias, con la finalidad de aprovechar aún más los recursos naturales del territorio mexicano (ver gráfica 1).

Evolución de las tasas de natalidad y mortalidad en México en el periodo 1940-1990 *Gráfica 1*



FUENTE: El Colegio de México, SPP-CONAPO-CELADE, Organización de las Naciones Unidas.

En este mismo periodo se empezó a mejorar la calidad de la vida de la población debido a la incorporación de los avances tecnológicos en el área biomédica; por ende, la mortalidad empezó a disminuir mientras que la natalidad siguió manteniendo su mismo nivel elevado. Una vez logrado el objetivo de aumentar la población en el país y con la creciente demanda de servicios sociales (salud, educación, etc.) a mediados de los setentas se cambió la anterior política por una de "planificación familiar". Esto se pretendió lograr dando auge a todo tipo de anticonceptivos que evitaba que las mujeres pudieran procrear. Además, a ambos sexos se les alienta a practicarse operaciones que esterilizan su reproducción total o parcialmente. Así la última etapa se caracteriza por un freno en la tasa de natalidad mientras la tasa de mortalidad disminuía imperceptiblemente.

MÉTODOS

Para el desarrollo de este trabajo utilizamos el método de "variables separables" de las ecuaciones diferenciales ordinarias de primer orden, así como una hoja de cálculo para obtener los valores de los parámetros del "modelo logístico" que mejor describe el cambio en el tamaño de la población conforme el tiempo transcurre. Para ello se recurrió al método de mínimos cuadrados.

La formulación matemática de ley logística de crecimiento de la población es la siguiente:

$$\frac{dP(t)}{dt} = aP(t) - bP^2(t) \qquad (1)$$

Sujeta a $P(t_0) = P_0$

Donde $P(t_0)$ denota el tamaño de la población en el tiempo t_0, y se le conoce como población inicial o población de referencia.

Los valores de a y b son los parámetros que se deben calcular, ya que éstos caracterizan el comportamiento del tamaño de la población.

El primero especifica la diferencia entre sus tasas de natalidad y mortalidad.

El segundo especifica una tasa de competencia intraespecies, en el caso de poblaciones humanas: muertes prematuras por desnutrición, servicios médicos insuficientes o inadecuados, enfermedades contagiosas o crímenes violentos, etcétera (4).

Los valores de los parámetros se calculan a partir de la información de los censos oficiales de población efectuados en México en el periodo 1900-1990, cuyas cifras se presentan en el cuadro 1.

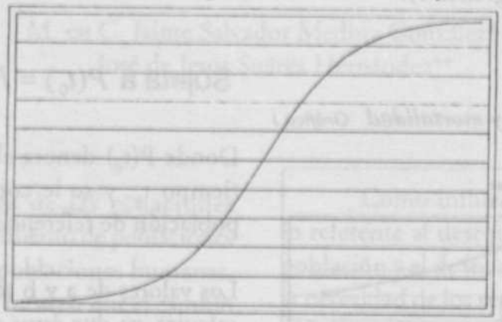
Población en México en el periodo 1900-1990 *Cuadro 1*

AÑO	POBLACIÓN (millones)
1900	13.607
1910	15.160
1920	14.335
1930	16.553
1940	19.654
1950	25.791
1960	34.923
1970	48.225
1980	66.847
1990	81.250

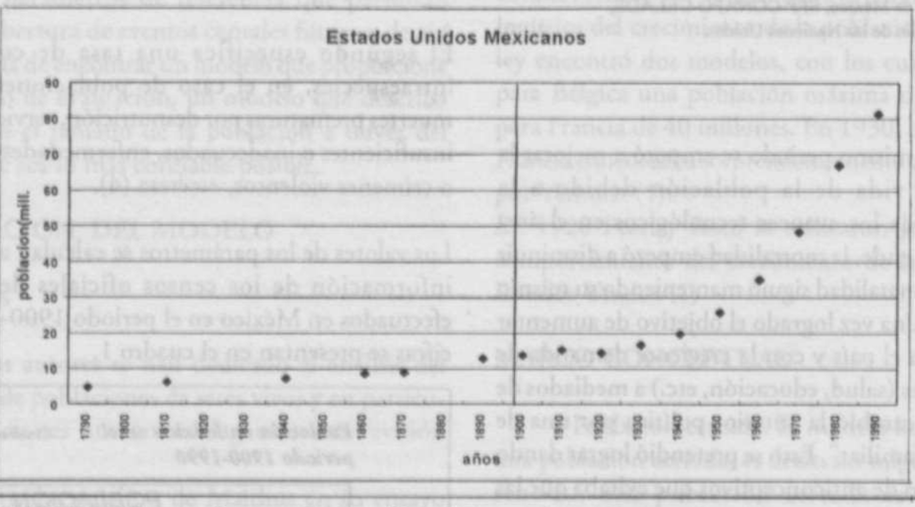
FUENTE: INEGI. II a XI Censo de Pobación y Vivienda

Se seleccionó este rango por dos razones. Una, por la regularidad temporal de los eventos censales, en este caso cada diez años, con la limitante de no haberse efectuado en la misma fecha (día y mes) del año. Aunque se dispone de información de años anteriores éstos no son uniformes en su periodicidad. Y segunda, por la enorme similitud de los datos a partir de 1900 en adelante con el comportamiento logístico (ver gráficas 2 y 3).

Curva logística típica *Gráfica 2*



Evolución de la población en México en el periodo 1900-1990 *Gráfica 3*



FUENTE: Diferentes fuentes oficiales del gobierno mexicano.

Para calcular los valores de los parámetros se requieren tres datos observados, cuyo tiempo transcurrido entre la observación de uno y otro sea el mismo (que estén igualmente espaciados). Por ejemplo si tomamos como dato de referencia t_0=1900 y el incremento en el tiempo es de 40 años se tendría la siguiente terna:

Ejemplo de terna de datos igualmente espaciados *Cuadro 2*

AÑO	*POBLACIÓN (millones)*
1900	13.607
1940	19.654
1980	66.847

Puesto que tenemos diez datos de los censos en diez tiempos distintos (de 1900 a 1990), existen varias formas de seleccionar ternas de datos igualmente espaciados. Para cada terna de datos se hace el cálculo de los parámetros que determinan la ley logística.

Valores del parámetro "a" según combinación de ternas posibles *Cuadro 3*

	Valores de a			
	Espaciamiento			
to	10	20	30	40
1900	#¡NUM!	-0.08109	-0.02960	-0.01158
1910	#¡NUM!	-0.06805	-0.00220	0.00067
1920	-0.00195	-0.00821	0.00878	
1930	-0.02393	0.00911	0.00759	
1940	0.01775	0.02433		
1950	0.02497	0.03804		
1960	0.03129			
1970	0.07785			

Valores del parámetro "b" según combinación de ternas posibles *Cuadro 4*

	Valores de b			
	Espaciamiento			
to	10	20	30	40
1900	#¡NUM!	-0.00603	-0.00245	-0.00130
1910	#¡NUM!	-0.00462	-0.00018	-0.00064
1920	-0.00106	-0.00144	-0.00056	
1930	-0.00229	-0.00063	-0.00004	
1940	-0.00042	-0.00017		
1950	-0.00018	0.00019		
1960	-0.00002			
1970	0.00078			

Ha sido demostrado (1) que las fórmulas para calcular a y b utilizando las ternas son:

$$a = \frac{1}{t_1 - t_0} Ln\left[\frac{P_2(P_1 - P_0)}{P_0(P_2 - P_1)}\right] \qquad (2)$$

$$b = \frac{a}{P_1}\left[\frac{P_1^2 - P_0 P_2}{P_0 P_1 - 2P_0 P_2 + P_1 P_2}\right] \qquad (3)$$

Donde P_{01} P_1 y P_2 son las poblaciones registradas en los tres tiempos igualmente espaciados t_0, t_1 y t_2 respectivamente.

Una vez que se tienen los valores de a y b, calculados con las fórmulas (2) y (3), éstos se sustituyen en la fórmula:

$$P(t) = \frac{aP_0}{bP_0 + (a - bP_0)e^{-a(t-t_0)}} \qquad (4)$$

Esta fórmula se encuentra resolviendo la ecuación diferencial (1) y nos permite obtener una aproximación al tamaño de la población en cualquier tiempo comprendido en el periodo 1900-1990. Una forma simplificada de esta expresión es la siguiente:

$$P(t) = \frac{k}{1 + \left(\frac{k}{P_0} - 1\right)e^{-a(t-t_0)}} \qquad (5)$$

Donde $k=a/b$

Este valor de k es importante, ya que representa la población límite, es decir la población máxima esperada a largo plazo.

RESULTADOS

Con los diez datos observados se obtuvieron veinte ternas de datos distintas y por ende veinte posibilidades para los valores de los parámetros a y b, de las cuales sólo dieciocho arrojaron valores válidos, como se muestra en el cuadro 5; los valores no válidos están indicados con #¡NUM! De las ternas válidas se seleccionó aquella cuyos parámetros correspondientes a y b, que determinan al modelo logístico, fueran tales que el coeficiente de correlación entre los datos observados en los censos poblacionales y los datos que arroja el modelo teórico sea máximo, es decir lo más cercano a uno posible.

Los valores de los parámetros fueron a=0.03804 y b=0.00019 que corresponden a los datos poblacionales registrados en los años 1950, 1970 y 1990 para estos valores se obtuvo un coeficiente de correlación de 0.99278, es decir, más de 99.2% de los datos observados son explicados por el modelo. Es necesario aclarar que estos valores obtenidos por la hoja de cálculo se han redondeado a cinco decimales.

Un criterio adicional para evaluar dichos parámetros es que la suma de los errores cuadrados sea mínima, que para este caso fue 220.07292.

Porción de una hoja de cálculo *Cuadro 5*

Valores de a				
	Espaciamiento			
to	10	20	30	40
1900	#¡NUM!	-0.08109	-0.02960	-0.01158
1910	#¡NUM!	-0.06805	-0.00220	0.00067
1920	-0.00195	-0.00821	0.00878	
1930	-0.02393	0.00911	0.00759	
1940	0.01775	0.02433		
1950	0.02497	$\boxed{0.03804}$		
1960	0.03129			
1970	0.07785			

Valores de b				
	Espaciamiento			
to	10	20	30	40
1900	#¡NUM!	-0.00603	-0.00245	-0.00130
1910	#¡NUM!	-0.00462	-0.00018	-0.00064
1920	-0.00106	-0.00144	-0.00056	
1930	-0.00229	-0.00063	-0.00004	
1940	-0.00042	-0.00017		
1950	-0.00018	$\boxed{0.00019}$		
1960	-0.00002			
1970	0.00078			

POB. REAL	AÑO	POB. TEORICA	ERROR
13.607	1900	4.315	-9.2925
	1905	5.196	
15.160	1910	6.251	-8.9097
	1915	7.512	
14.335	1920	9.016	-5.3189
	1925	10.804	
16.553	1930	12.924	-3.6285
	1935	15.427	
19.654	1940	18.367	-1.2868
	1945	21.802	
25.791	1950	25.791	0.0000
	1955	30.387	
34.923	1960	36.638	0.7153
	1965	41.579	
48.225	1970	48.225	0.0000
	1975	55.569	
66.847	1980	63.573	-3.2734
	1985	72.168	
81.250	1990	81.250	0.0000
91.158	1995	90.685	
	2000	100.316	
	2005	109.973	
	2010	119.482	
	2015	128.682	
	2020	137.431	
	2025	145.617	
	2030	153.159	
	2035	160.011	

COEF. DE CORR.
0.99278297

SUMA DE ERRORES CUADRADOS
220.0729184

El cuadro 5 es una porción de una hoja de cálculo; en ella aparecen, en primera instancia, las dos tablas para los posibles valores de los parámetros a y b. Aparece también una tabla comparativa de los datos reales y los teóricos, junto con el error correspondiente.

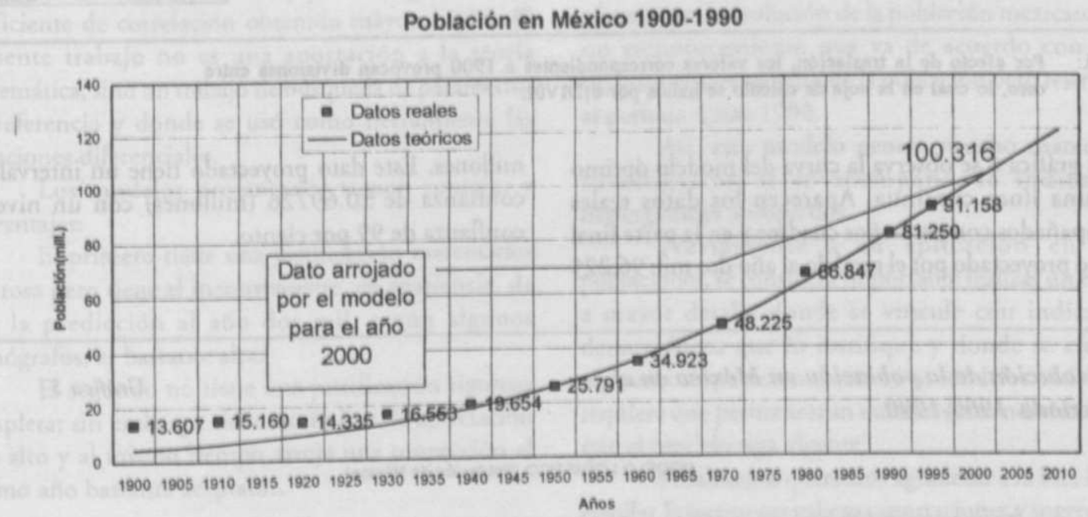
Al sustituir los datos en la fórmula simplificada (5) obtenemos el modelo:

$$P(t) = \frac{203.45881}{1+\left(\frac{203.45881}{25.791}-1\right)e^{-0.03804(t-1950)}} \qquad (6)$$

El lector interesado en trabajar con este modelo debe tener en cuenta que todos los valores sustituidos en la fórmula están redondeados a cinco decimales, por lo que obtendrá resultados aproximados.

En la gráfica 4 se observa la curva del modelo óptimo con una línea continua. Los datos reales son señalados con pequeños cuadros y en la parte derecha está el dato proyectado por el modelo al año dos mil, es decir, 100 millones 316 mil. Este dato proyectado tiene un intervalo de confianza de ±2.88286 (millones) con un nivel de confianza de 99 por ciento.

Evolución de la población en México en el periodo 1900-1990 *Gráfica 4*



FUENTE: Diferentes fuentes oficiales del gobierno mexicano y Datos obtenidos por el modelo.

Al observar este ejercicio vemos que los datos que corresponden a los años 1900, 1910, 1920 y 1930, tienen un error significativo, por lo que se decidió realizar un cambio que consistió en ajustar el modelo a una traslación vertical hacia abajo del volumen de la población y al final una traslación vertical hacia arriba del ajuste obtenido de igual magnitud obteniendo los siguientes resultados.

Los valores de los parámetros óptimos fueron a*=0.07189 y b*=0.00072 que corresponden a los datos poblacionales registrados en los años 1940, 1960 y 1980, para estos valores el coeficiente de correlación correspondiente es 0.99960. La suma de los errores cuadrados para este caso fue 8.71100.

En la figura se muestra una porción de esta hoja de cálculo, en ella aparece en primera instancia las dos tablas para los posibles valores de los parámetros a* y b*. Viene también la gráfica resultado de los cálculos con el modelo teórico. Aparece una tabla comparativa de los datos reales y los teóricos. Además se incluye una columna de errores (ver cuadro 5).

Porción de una hoja de cálculo para el sgundo modelo

Cuadro 6

VALORES DE a*				
	Espaciamiento			
to	10	20	30	40
1900	#¡DIV/0!	#¡DIV/0!	#¡DIV/0!	#¡DIV/0!
1910	#¡NUM!	0.00837	0.04181	0.05305
1920	0.17825	0.11615	0.10055	
1930	0.07371	0.07884	0.07363	
1940	0.08626	$\boxed{0.07189}$		
1950	0.06681	0.06637		
1960	0.05790			
1970	0.09268			

VALORES DE b*				
	Espaciamiento			
to	10	20	30	40
1900	#¡DIV/0!	#¡DIV/0!	#¡DIV/0!	#¡DIV/0!
1910	#¡NUM!	-0.01114	-0.00107	-0.00465
1920	0.02320	0.00395	0.00156	
1930	0.00042	0.00119	0.00080	
1940	0.00183	$\boxed{0.00072}$		
1950	0.00066	0.00064		
1960	0.00034			
1970	0.00113			

	POBLACION		
AÑO	OBSERVADA	TEORICA	ERROR
1900	13.607	13.969	0.362
1910	15.160	14.346	-0.814
1920	14.335	15.112	0.778
1930	16.553	16.648	0.095
1940	19.654	19.654	0.000
1950	25.791	25.272	-0.519
1960	34.923	34.923	0.000
1970	48.225	49.322	1.097
1980	66.847	66.847	0.000
1990	81.250	83.576	2.326
2000		$\boxed{96.225}$	

0.99960447	8.71099695
COEF. DE CORR.	SUMA DE ERRORES CUADRADOS

ERROR TIPICO	0.910133357
DESVIACION ESTANDAR	0.897786142

INTERVALO DE CONF. 95%	0.53054717
LIMITE INFERIOR	95.69442337
LIMITE SUPERIOR	96.75551771

INTERVALO DE CONF. 99%	0.697259617
LIMITE INFERIOR	95.52771092
LIMITE SUPERIOR	96.92223016

NOTA: Por efecto de la traslación, los valores correspondientes a 1900 provocan divisiones entre cero, lo cual en la hoja de cálculo se indica por #¡DIV0!.

En la gráfica 6 se observa la curva del modelo óptimo con una línea continua. Aparecen los datos reales acompañados con pequeños cuadros y en la parte final el dato proyectado por el modelo al año dos mil: 96.225 millones. Este dato proyectado tiene un intervalo de confianza de ±0.69726 (millones) con un nivel de confianza de 99 por ciento.

Evolución de la población en México en el periodo 1900-1990

Gráfica 5

FUENTE: Diferentes fuentes oficiales del gobierno mexicano y datos obtenidos por el segundo modelo.

DISCUSIÓN

En general el modelo logístico no es recomendado para utilizarse en poblaciones humanas, debido a la enorme variabilidad a la que se encuentra sometido su crecimiento como producto de los cambios en la situación socioeconómica o en los patrones culturales y avances tecnológicos.

Se decide emplearlo porque se presentaron muchas condiciones que concordaban con las características intrínsecas en el modelo. Lo que se pretende aquí es sólo dar un ejemplo de la aplicación de las ecuaciones diferenciales y dar a conocer al lector interesado la hoja de cálculo construida para modelar fenómenos de ese tipo.

Se tomó el riesgo de efectuar una predición al año dos mil por la cercanía con los datos observados (se tiene el dato de 1995 con el cual se puede medir la tendencia al final del periodo) y por la magnitud en el coeficiente de correlación obtenido mayor a 0.99. El presente trabajo no es una aportación a la teoría matemática, sino un trabajo de búsqueda de parámetros de referencia y donde se usó como herramienta las ecuaciones diferenciales.

Los modelos presentados tienen ventajas y desventajas:

El primero tiene una justificación matemática rigurosa pero tiene el inconveniente, en apariencia, de que la predicción al año dos mil, según algunos demógrafos, es bastante alta.

El segundo no tiene una justificación rigurosa completa; sin embargo tiene un índice de correlación más alto y al mismo tiempo arroja una proyección al mismo año bastante aceptable.

CONCLUSIONES

Este método ofrece la ventaja de que con tan sólo una terna de datos se pueden obtener los parámetros del modelo. Además de que entre más datos observados se incorporen al análisis mayor precisión se obtiene en la aproximación.

El método y la hoja de cálculo desarrollada es fácilmente aplicable a cualquier fenómeno que tenga el comportamiento como la ley logística, con la única condición de que los datos observados sean más de tres y estén igualmente espaciados.

El modelo se ha usado con mayor precisión en estudios de poblaciones controladas ya sea en lugares abiertos (regiones cercadas) o cerrados (laboratorios).

El análisis de poblaciones humanas se puede llevar a cabo en regiones más pequeñas como podría ser alguna entidad federativa.

De acuerdo a los datos obtenidos podemos afirmar que la evolución de la población mexicana tiene un comportamiento que va de acuerdo con la ley logística de crecimiento de la población pero restringida al periodo 1900-1990.

Así, este modelo genera mucho material de investigación si se acompaña de indicadores demográficos apropiados.

Previamente a su aplicación en otras poblaciones, se considera importante realizar un estudio a mayor detalle donde se vincule con indicadores demográficos que lo justifique y donde se efectúen algunas comparaciones con otros modelos. Además se requiere que permanezcan estables las condiciones para que el modelo siga vigente.

Finalmente queremos agradecer a la Lic. Lupita Aguilar Frías por sus valiosas aportaciones y sugerencias. Se agradece también la participación del M. en P. Fernando Padilla Lozano.

BIBLIOGRAFÍA

(1) Brawn, Martín
Ecuaciones diferenciales y sus aplicaciones.
Edit. Iberoamericana, 1990.

(2) Keyfitz, Nathan
Introducción a las Matemáticas de Población
Centro Latinoamericano de Demografía
Santiago de Chile, 1979.

(3) Lotka, Alfred J.
Teoría analítica de las asociaciones biológicas.
Centro Latinoamericano de Demografía
Santiago de Chile, 1976.

(4) Nagle, R. Kent
Fundamentos de ecuaciones diferenciales
Addison-Wesley Iberoamericana, 1992.