

Gabriel Poveda Ramos  
Jorge Manrique H.

### **Resumen**

Este artículo se refiere a la aplicación de la curva logística, que es un instrumento matemático usado por biólogos y demógrafos para describir el crecimiento de poblaciones animales y de poblaciones humanas, así como para formular pronósticos futuros, a los censos de la ciudad de Medellín – Colombia, efectuados entre 1912 y 2005. Presentamos aquí la aplicación del método para calcular la curva logística de una población a partir de recuentos anteriores, y usando la técnica de mínimos cuadrados, desarrollada por Gabriel Poveda Ramos en la referencia [1]. La parte numérica del proceso se ilustra de una manera simple para un mejor entendimiento del manejo de esta técnica que, aunque es muy conocida y sencilla de usar, hasta hoy no se ha empleado (que sepamos) para ajustar dicha curva.

**Palabras clave:** Modelos matemáticos de población, Demografía, Medellín, Censos, Crecimiento logístico, Matemática aplicada.

### **Abstract**

*This paper refers to the logistic curve application which is a mathematical tool often used by biologists and demographers to describe the growth of animal and human populations, as well as to make forecastings on that growth to the censuses realized in Medellín- Colombia city, between 1912 and 2005. Here is presented this method in order to calculate the logistic curve for a given population on the basis of previous censuses, and using the technique of minimum squares, developed by Gabriel Poveda Ramos in reference [1]. The numerical part of the process is simply illustrated for better understanding of this technique that is well known and easy to use but, it has not been employed heretofore to adjust the logistic curve.*

**Keywords:** Mathematical Models of Population, Demography, Medellín, Censuses, Logistic growth, Applied Mathematics.

**Clasificación JEL:** J10, N36, C10.

# Aplicación de la curva logística a los censos de la ciudad de Medellín

Gabriel Poveda Ramos (\*)

Jorge Manrique H (\*\*)

## Historia

La curva logística es suficientemente conocida entre estadísticos, biólogos y demógrafos como modelo determinístico para describir y para pronosticar el crecimiento de una población humana o de poblaciones de algunas especies animales, a lo largo del tiempo, cuando están sujetas a ciertas condiciones que su régimen de crecimiento. Fue François Quetelet quien formuló las ideas básicas (en 1845, en Bélgica), que llevaron a su amigo P. F. Verhulst a darle forma matemática (entre 1838 y 1847) a las ideas del primero. En 1920, sin conocer los trabajos de aquellos científicos europeos, dos biometristas americanos, R. Pearl y L. J. Reed redescubrieron la curva de Verhulst, a la cual hoy llamamos curva logística de Verhulst – Pearl.

En las referencias bibliográficas, se puede consultar más sobre las aplicaciones biométricas y demográficas de la curva logística, y las limitaciones de los métodos tradicionales para el ajuste de la curva logística.

## Régimen de crecimiento logístico

Para un mejor entendimiento de la aplicación, se transcribe la teoría que soporta el método.

1. El régimen de crecimiento logístico resulta cuando se supone que la dinámica de crecimiento en el tiempo está determinada por tres hipótesis básicas, a saber:

---

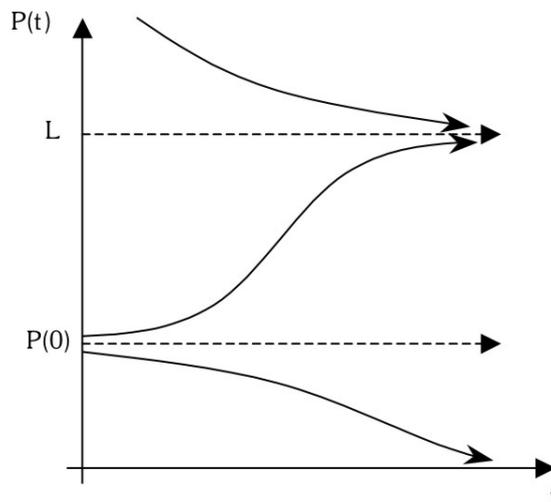
(\*) Ingeniero Químico y Electricista. Doctor Honoris Causa en Ingeniería y Magister en Matemática Aplicada. Profesor emérito de la Universidad Pontificia Bolivariana, Medellín- Colombia. Fecha de recepción 16 de julio de 2007. Fecha de aceptación 26 de septiembre de 2007

(\*\*) Ingeniero Mecánico. Magíster en Gestión Tecnológica. Profesor de la Universidad Pontificia Bolivariana, Medellín- Colombia.

- a. El crecimiento vegetativo (más las inmigraciones, si las hay), es en todo momento proporcional al número de individuos presentes, al cual llamaremos  $p$ .
- b. La población vive en un espacio limitado y/o bajo condiciones externas restrictivas que no le permitirán, ni aun a largo plazo, exceder de un tamaño límite, al cual llamaremos  $L$ .
- c. En todo momento la velocidad de crecimiento es proporcional al factor  $(L - p)$  que representa las fuerzas exógenas restrictivas.

En la Gráfica 1, se observan los comportamientos creciente y decrecientes de la población en el tiempo según las condiciones iniciales de: superpoblación desde  $P(t)$  hacia una población límite  $L$ ; de crecimiento ascendente desde una población inicial  $P(0)$  hasta la población límite  $L$ ; y de crecimiento negativo o decrecimiento, desde una población crítica hasta la extinción. Ver referencia [12].

**Gráfica 1.**  
Comportamientos de la Curva Logística



Dentro de las hipótesis anteriores, y si se sabe que el dato censal que es históricamente más reciente corresponde al valor más alto ya

censado, se determina que la ecuación diferencial para el crecimiento para una población ascendentemente creciente, es:

$$\frac{dp}{dt} = bp(L - p) \quad (01)$$

en donde

$p$ : Número de individuos de la población que viven en el instante  $t$  del tiempo.

$b$ : Coeficiente constante y positivo que es propio de las características vitales de la población y del régimen biológico y ecológico en que vive.

$L$ : Población límite posible, que está determinada por el entorno vital de la población.

Integrando la ecuación diferencial (01) ya escrita, se obtiene la función de crecimiento de la población:

$$\hat{p}(t) = \frac{L}{1 + Ke^{-bt}}, \quad (02)$$

en donde  $K$  es un constante que se introduce al efectuar la integración y que, numéricamente, depende del volumen que tenga la población en el momento que se tome como instante inicial ( $t = 0$ ) para comenzar a contar el tiempo. En efecto, de la ecuación (02) se deduce que

$$K = \frac{L}{p(0)} - 1$$

Es evidente que la forma de la ecuación (02) es invariante ante un cambio en el origen de tiempo y ante un cambio en la unidad de medida del tiempo. En efecto, haciendo un cambio en ambos sentidos, es decir, tomando una nueva variable temporal

$$T = ct + a,$$

donde  $a$  expresa un cambio en el origen del tiempo (o cronología) y  $c$  un cambio en la unidad de medida del tiempo (o cronometría), la ecuación toma la forma:

$$\hat{p}(t) = \frac{L}{1 + K_1 e^{-Bt}}$$

en donde

$$K_1 = Ke^{ab/c} \text{ y } B = b/c$$

2. En las aplicaciones demográficas de la logística, se trata usualmente de “ajustar” una función de este tipo a una sucesión consecutiva de valores numéricos de la población

$$p_1, p_2, \dots, p_n,$$

los cuales son medidos experimentalmente (por ejemplo, mediante recuentos o censos de la población, en momentos sucesivos y distintos del tiempo

$$t_1, t_2, \dots, t_n$$

los cuales se disponen en orden creciente, pero no son necesariamente equidistantes.

En tales situaciones los datos de población ( $p_1, p_2, \dots, p_n$ ) suelen ser crecientes en el tiempo. Estos datos presentan tres características prácticas, que son importantes para entender el problema teórico del ajuste de la curva logística:

a. Casi inevitablemente están afectados por errores de recuento (p.e. sobre-enumeración, sub-enumeración) o pueden ser el resultado de estimaciones y no de enumeraciones precisas.

b. Aunque tales datos fueran rigurosamente exactos, no puede esperarse que todos caigan rigurosa y exactamente sobre una función basada en hipótesis determinísticas (como lo es la logística) ya que en el crecimiento de un agregado colectivo, como en el caso de una población biológica, hay inevitablemente desviaciones aleatorias más o menos apreciables con respecto a la función determinística.

c. Sucede entonces que, aun cuando la realidad geográfica, ambiental, ecológica y social justifique la adopción de una curva logística de crecimiento, el problema del ajuste a los datos empíricamente observados en el tiempo, presenta varias dificultades

de técnica numérica. Ese problema consiste, en esencia, en calcular los tres parámetros  $L$ ,  $K$ ,  $b$  de la ecuación (02), a partir de la serie de recuentos, que suele darse en forma tabulada:

<b>Fecha</b>	<b>Población</b>
$t_1$	$p_1$
$t_2$	$p_2$
.	.
$t_n$	$p_n$

Es evidente que, para hacer el ajuste, no se puede esperar de la curva – la cual sólo tiene tres parámetros para estimar – que pase exactamente por los puntos tabulados, lo cual daría lugar a  $n$  condiciones y como, en general,  $n$  es distinto de 3, el problema no tendría solución. Aun el caso de tener tres observaciones censales ( $n = 3$ ), las observaciones teórico-prácticas mencionadas descartan este método.

Estadísticos como Lyra Madeira [1] y [4], y Toranzos [5], dedujeron métodos para calcular los parámetros  $L$ ,  $K$  y  $b$ , pero presentan errores de truncamiento y de suposición de hipótesis espurias.

### **El ajuste de la logística por mínimos cuadrados.**

3. La ecuación (02) de la curva logística puede escribirse en la forma equivalente:

$$\text{Ln}\left(\frac{L}{p} - 1\right) = \text{Ln}(k) - bt$$

o bien como

$$z = A - bt, \tag{03}$$

siendo

$$z = \text{Ln}\left(\frac{L}{p} - 1\right), \quad (04)$$

$$A = \text{Ln}(k) \quad (05)$$

y donde todos los logaritmos son neperianos (naturales).

Lo anterior significa que si, en lugar de representar la ley logística en coordenadas cartesianas e isométricas para  $t$  en las abscisas y  $p$  en las ordenadas, la representamos en coordenadas cartesianas o isométricas para  $t$  en abscisas y  $z$  en ordenas; si a  $L$  se le atribuye el valor correcto, se obtendrá una línea recta, que representa a la ecuación (03). Si los valores de  $z$  se calculan en la ecuación (04) con valores de  $L$  que no son los verdaderos valores de la ecuación (03), la gráfica entre  $t$  y  $z$  no dará una recta, sino una curva.

Esto mismo equivale a decir que, calculando  $z$  con el valor  $L$  correcto de la logística, entonces la gráfica de  $z$  versus  $t$  es una *parábola*

$$z = A + Bt + Ct^2 \quad (6)$$

en donde  $C = 0$ , y siendo  $A = \text{Ln}(K)$  y  $B = -b$

Ahora bien, es sabido en estadística elemental, que la parábola

$$z = A + Bt + Ct^2$$

puede ajustarse a una serie  $\{t_i, z_i\}$  de datos empíricos, usando el criterio de mínimos cuadrados, y su método de análisis (ver por ejemplo, la referencia [7] de la bibliografía). En ese caso, las *ecuaciones normales* que determinan los tres parámetros  $A, B, C$  son las del sistema

$$\sum z_i = An + B\sum t_i + C\sum t_i^2$$

$$\sum z_i t_i = A\sum t_i + B\sum t_i^2 + C\sum t_i^3$$

$$\sum z_i t_i^2 = A\sum t_i^2 + B\sum t_i^3 + C\sum t_i^4$$

en donde  $n$  es el número de fechas censales que se tienen.

En todos los casos de aplicación práctica se comprueba que el sistema es resoluble, o sea que el sistema no es homogéneo (es decir,

los tres términos  $\sum z_i$ ,  $\sum z_i t_i$ ,  $\sum z_i t_i^2$ , no se anulan simultáneamente), y que el determinante del sistema

$$D \equiv \begin{vmatrix} n & \sum t_i & \sum t_i^2 \\ \sum t_i & \sum t_i^2 & \sum t_i^3 \\ \sum t_i^2 & \sum t_i^3 & \sum t_i^4 \end{vmatrix}$$

es distinto de cero.

Para simplificar la escritura, pondremos:

$$T_1 = \sum t_i, \quad T_2 = \sum t_i^2, \quad T_3 = \sum t_i^3, \quad T_4 = \sum t_i^4$$

Aplicando a las ecuaciones normales la regla de Cramer, se obtiene para el coeficiente  $C$  la expresión

$$C = \frac{\begin{vmatrix} n & T_1 & \sum z_i \\ T_1 & T_2 & \sum z_i t_i \\ T_2 & T_3 & \sum z_i t_i^2 \end{vmatrix}}{D}$$

Para la curva logística, con  $L$  correcto, se debe tener  $C = 0$ , como ya se dijo. Por tanto, se debe tener la condición

$$\begin{vmatrix} n & T_1 & \sum z_i \\ T_1 & T_2 & \sum z_i t_i \\ T_2 & T_3 & \sum z_i t_i^2 \end{vmatrix} = 0 \quad (07)$$

que determina implícitamente el parámetro  $L$ , ya que éste está contenido en cada  $z_i$ .

Escribiendo por abreviar,

$$\begin{cases} P = T_1 T_3 - T_2^2 \\ Q = -n T_3 + T_1 T_2 \\ R = n T_2 - T_1^2 \end{cases}$$

entonces, la ecuación (07) se puede escribir en la forma

$$P \sum z_i + Q \sum z_i t_i + R \sum z_i t_i^2 = 0 \quad (08)$$

o bien

$$\sum \left( P + Q t_i + R t_i^2 \right) \cdot \text{Ln} \left( \frac{L}{p_i} - 1 \right) = 0 \quad (09)$$

4. La anterior es una ecuación no-algebraica, que contiene a L como incógnita, del tipo

$$f(L) = 0 \quad (10)$$

en donde  $f$  es un polinomio no algebraico, que aparece en el lado izquierdo de la ecuación (09). Este tipo de ecuaciones sólo puede resolverse por métodos numéricos, es decir, buscando numéricamente los valores de  $f(L)$  correspondientes a distintos valores numéricos, hasta localizar el valor  $L^*$  que anula a  $f(L)$ .

El hecho de que la ecuación  $f(L) = 0$  tiene efectivamente una raíz real y única, se desprende de las siguientes consideraciones:

- a. Que  $f(L)$  es continua, y está definida desde el valor  $L=M = \max(p_1, p_2, \dots, p_n)$  (con  $M>0$ ) hasta  $L = +\infty$ . En las situaciones prácticas que se manejan en demografía, la sucesión  $p_1, p_2, \dots, p_n$  suele ser de datos de población en crecimiento. En ese caso,  $M = P_n$ .
- b. Que para  $L = M$  se tiene que  $f(L) = -\infty$  (porque  $M/p_n=1$ )
- c. Que  $f(L)$  es monótonicamente creciente con  $L$ , o sea que  $f'(L) > 0$ .
- d. Que para  $L \rightarrow +\infty$ , se tiene que  $f(L) \rightarrow +\infty$ .

Hay varios métodos numéricos utilizables para calcular la raíz de la ecuación  $f(L) = 0$ : el de Newton, el de falsa posición (*Regula Falsis*), el de bisección y muchos otros. El método y la forma de aplicarlo pueden consultarse en cualquier texto de Análisis Numérico, como los de las referencias [8], [9] y [12] de la bibliografía. En este caso hemos usado el método de la bisección para resolver esta ecuación pues, es muy fácil de aplicar al problema, converge rápidamente, permite encontrar

la raíz con la exactitud que se quiera y es muy sencillo de realizar y de aprender. Lo consideramos casi el óptimo recomendable para casos como el tratado. Ver su descripción en el ANEXO 1.

5. El cómputo numérico de la raíz  $L = L^*$  de la ecuación (10) constituye la clave del método que exponemos aquí. Antes de continuar explicando el método, es importante demostrar que esa raíz no depende del origen que se haya escogido para medir el tiempo. Es decir, que si las fechas censales  $t_1, t_2, \dots, t_n$  se dan en la cronología cristiana, o judía, o mahometana, la raíz  $L^*$  es una misma. Y, en general, que si se mide el tiempo  $t$  desde cualquier momento y en cualesquiera unidades, la ecuación (10) es invariante.

En efecto, al cambiar el origen de tiempo a cualquier otro instante, la nueva fecha  $\tau_i$  y la original  $t_i$  están relacionadas por

$$\tau_i = t_i + \alpha$$

siendo  $\alpha$  el corrimiento que se haya hecho en el origen de tiempo. Entonces:

$$T'_1 = \sum(t_i + \alpha) = T_1 + n\alpha$$

$$T'_2 = \sum(t_i + \alpha)^2 = T_2 + 2\alpha T_1 + n\alpha^2$$

$$T'_3 = \sum(t_i + \alpha)^3 = T_3 + 3\alpha T_2 + 3\alpha^2 T_1 + n\alpha^3$$

$$T'_4 = \sum(t_i + \alpha)^4 = T_4 + 4\alpha T_3 + 6\alpha^2 T_2 + 4\alpha T_1 + n\alpha^4$$

Sustituyendo y haciendo algunas operaciones algebraicas que omitimos, se obtiene

$$P' = T'_1 T'_3 - (T'_2)^2 = P - \alpha^2 T_1^2 - \alpha T_1 T_2 + n\alpha^2 T_2 + n\alpha T_3$$

$$Q' = -nT'_3 + T'_1 T'_2 = Q + 2\alpha T_1 - 2n\alpha T_2$$

$$R' = nT'_2 - (T'_1)^2 = R$$

Luego:

$$P' + Q'\tau_i + R\tau_i^2 = P + Qt_i + Rt_i + \alpha \cdot [Q + nT_3 - T_1 T_2] + \alpha \cdot (\alpha + 2t_i) \cdot [R - nT_2 + T_1^2]$$

y los dos corchetes que hay a la derecha son nulos por definición, según las ecuaciones (08). Es decir, que para cualquier cambio de origen en el tiempo, la expresión

$$P + Qt_i + Rt_i^2$$

es invariante en su forma algebraica y en su valor numérico.

Mediante operaciones algebraicas sencillas se muestra que, para un cambio en la escala de tiempo

$$\tau_i = ct_i, \text{ con } c > 0,$$

se obtiene

$$P' + Q'\tau_i + R'\tau_i^2 = c^4(P + Qt_i + Rt_i^2)$$

y por lo tanto, en la ecuación (09) solamente aparece el lado izquierdo multiplicado por el factor  $c^4$  ( $c^4 \neq 0$ ), lo cual, evidentemente, no cambia la raíz  $L^*$  de esa ecuación.

6. El parámetro  $b$  de la curva logística se calcula recordando que

$$b = -B$$

y que  $B$  se obtiene del sistema de ecuaciones normales. Por lo tanto

$$b = - \frac{\begin{vmatrix} n & \sum z_i & T_2 \\ T_i & \sum z_i t_i & T_3 \\ T_2 & \sum z_i t_i^2 & T_4 \end{vmatrix}}{D}$$

y así mismo

$$Ln K = \frac{\begin{vmatrix} \sum z_i & T_1 & T_2 \\ \sum z_i t_i & T_2 & T_3 \\ \sum z_i t_i^2 & T_3 & T_4 \end{vmatrix}}{D}$$

También es importante aquí hacer notar que el determinante  $D$  es invariante ante cualquier traslación del origen de tiempo:

$$\tau_i = t_i + a,$$

traslación en la cual

$$D \equiv \begin{vmatrix} n & T'_1 & T'_2 \\ T'_1 & T'_2 & T'_3 \\ T'_2 & T'_3 & T'_4 \end{vmatrix}$$

$$D = n[T'_2 T'_4 - (T'_3)^2] - T'_1(T'_1 T'_4 - T'_2 T'_3) - T'_2 [T'_1 T'_3 - (T'_2)^2] \quad (11)$$

Sustituyendo según las ecuaciones (11), y después de un largo cálculo algebraico, se encuentra que

$$\begin{aligned} D' &\equiv n[T'_2 T'_4 - (T'_3)^2] - T'_1(T'_1 T'_4 - T'_2 T'_3) - T'_2 [T'_1 T'_3 - (T'_2)^2] \\ &= n[T_2 T_4 - T_3^2] - T_1 [T_1 T_4 - T_2 T_3] - T_2 [T_1 T_3 - T_2^2] \\ &= \begin{vmatrix} n & T_1 & T_2 \\ T_1 & T_2 & T_3 \\ T_2 & T_3 & T_4 \end{vmatrix} = D \end{aligned}$$

Es decir, que el determinante  $D$  es invariante en forma algebraica y en valor numérico ante un cambio de origen de tiempo (o cambio de cronología). Se comprueba también que ante un cambio de unidades de tiempo (o cronometría),

$$\tau_i = c t_i$$

el nuevo determinante es  $D^i = c^6 D$

Por lo tanto un cambio en cronología y en cronometría de la forma

$$\tau_i = c t_i + a$$

modifica el determinante  $D$  en la forma

$$D^i = c^6 D$$

7. Las anteriores observaciones permiten formular un algoritmo explícito, directo y general para calcular los tres parámetros  $L$ ,  $b$ ,  $K$ , de una curva logística

$$p(t) = \frac{L}{1 + K e^{-bt}}$$

a partir de los datos de población (presumiblemente crecientes),  $p_1, p_2, \dots, p_n$  (siendo  $n > 3$ ),

medidos empíricamente (por ejemplo, por recuentos censales), en las fechas

$t_1, t_2, \dots, t_n$ ,

y según el criterio de mínimos cuadrados. Para medir y expresar las fechas, se supone que hemos tomado un origen de tiempo (que puede ser cualquiera), y una unidad de medida, de la cual dependerán los parámetros  $K, b$ , pero no el parámetro  $L$ .

8. La tasa relativa de crecimiento poblacional es una variable muy importante en todos los estudios demográficos y está definida por la identidad:

$$r(t) = \frac{1}{p} \frac{dp}{dt} \tag{12}$$

que se interpreta como la velocidad del crecimiento de la población, dividida por la población. Al efectuar el proceso de derivación correspondiente, del cual se muestran algunos pasos elementales, se obtiene:

$$\frac{dp}{dt} = \frac{d}{dt} \left( \frac{L}{1 + Ke^{-bt}} \right) = \frac{Lkbe^{-bt}}{(1 + ke^{-bt})^2}$$

Ahora, al reemplazar en la ecuación (12)

$$r(t) = \frac{1}{p} \frac{dp}{dt} = \left( \frac{1 + Ke^{-bt}}{L} \right) \cdot \frac{Lkbe^{-bt}}{(1 + ke^{-bt})^2}$$

y luego de las correspondientes cancelaciones de términos, la tasa relativa de crecimiento para la curva logística resulta ser:

$$\hat{r}(t) = \frac{K \cdot b}{K + e^{bt}} \tag{13}$$

Siendo  $t =$  año que se calcula – año base, p.e. 1950

Generalmente el año de base se puede tomar como un año intermedio en la serie de fechas de los datos censales.

## El algoritmo de solución

Según lo explicado, el algoritmo para el ajuste de la curva logística es el siguiente:

1. Tómese el origen de tiempo o año base en cualquier momento (o fecha), y la unidad de tiempo que se desee. Si se trata de censos de población en Colombia durante el siglo XX, una fecha conveniente para fijar el origen del tiempo es uno de los años aproximadamente del intermedio del rango de años censados por ejemplo, el 1 de enero de 1950, y la unidad de medida obvia es el año – calendario. En esa forma, la fecha del 15 de octubre de 1993, por ejemplo, corresponde al valor  $t = 43,82$  años, y la del 17 de noviembre de 1928 es  $t = -21,14$  años.

2. Calcúlense los valores numéricos de

$$T_1, T_2, T_3, T_4: \quad T_1 = \sum t_i, \quad T_2 = \sum t_i^2, \quad T_3 = \sum t_i^3, \quad T_4 = \sum t_i^4$$

Calcúlense cada uno de los valores numéricos de

$$P = T_1 T_3 - T_2^2, \quad Q = -n T_3 + T_1 T_2, \quad R = n T_2 - T_1^2$$

3. Calcúlense cada uno de los valores numéricos de los  $n$  coeficientes

$$P + Q t_i + R t_i^2 \quad \text{para } i = 1, 2, \dots, n$$

4. Fórmese la ecuación

$$f(L) = \sum (P + Q \cdot t_i + R \cdot t_i^2) \cdot \ln \left( \frac{L}{p_i} - 1 \right) = 0$$

5. Calcúlese el valor de  $L$  en la forma siguiente:

Explorando distintos valores numéricos para  $L$ , encuéntrase aquel que hace  $f(L)=0$ . Esto puede lograrse partiendo arbitrariamente de un valor pequeño p.e.  $L_1=p_1$ , que dará un resultado negativo; elegir luego otro de valor numérico alto, p.e.  $L_2=2M$  (2 veces el valor más alto censado, o mayor si se requiere), que dará un resultado positivo. A partir de esos dos valores con signo diferente, calcular el promedio aritmético de  $L_1$  y  $L_2$ , que da como resultado  $L_3$ . Calcular ahora  $f(L_3)$ ,

compararlo con  $f(L_1)$  y  $f(L_2)$  e identificar el que tiene signo diferente. Hallar el promedio aritmético entre estos nuevos valores y así sucesivamente hasta encontrar aquel valor de  $L$  cuyo resultado para  $f(L)$  sea cero. Este procedimiento se denomina el método de bisección (ver ANEXO 1), y sirve para localizar con la exactitud que se desee y de una manera rápida, la raíz de la ecuación  $f(L)=0$ . Evidentemente, puede usarse también el método de falsa posición u otro adecuado para este fin.

6. Con el valor anterior de  $L$ , calcúlense los  $n$  valores numéricos

$$z_i = Ln\left(\frac{L}{p_i} - 1\right), \text{ para } i = 1, 2, \dots, n$$

7. Calcúlense las sumas

$$\sum z_i, \sum z_i t_i, \sum z_i t_i^2$$

8. Valórese el determinante

$$D \equiv \begin{vmatrix} n & T_1 & T_2 \\ T_1 & T_2 & T_3 \\ T_2 & T_3 & T_4 \end{vmatrix}$$

9. Calcúlese:

$$b = - \frac{\begin{vmatrix} n & \sum z_i & T_2 \\ T_1 & \sum z_i t_i & T_3 \\ T_2 & \sum z_i t_i^2 & T_4 \end{vmatrix}}{D}$$

10. Calcúlese:

$$K = \exp \frac{\begin{vmatrix} \sum z_i & T_1 & T_2 \\ \sum z_i t_i & T_2 & T_3 \\ \sum z_i t_i^2 & T_3 & T_4 \end{vmatrix}}{D}$$

11. Enumérense los tres parámetros  $L$ ,  $b$ ,  $K$ , y escríbase la función logística según

$$p(t) = \frac{L}{1 + Ke^{-bt}}$$

$t =$  año de que se trata – año base, p.e. 1950

12. Calcúlese la tasa relativa de crecimiento poblacional para la curva logística según la ecuación:

$$\hat{r}(t) = \frac{K \cdot b}{K + e^{(bt)}}$$

$t =$  año de que se trata – año base, p.e. 1950

Este algoritmo para calcular la ecuación de la curva logística y la tasa relativa de crecimiento poblacional se puede ejecutar manualmente, en una calculadora de bolsillo programable, en una hoja electrónica de cálculo o en cualquier aplicación de *software* para estadística.

### **Aplicación del método a los censos de la ciudad de Medellín**

A partir de los datos de los censos oficiales que se han hecho desde 1912 hasta 2005, suministrados por el Departamento Nacional de Estadística DANE (Tabla 1), se calculará la curva logística. Debe considerarse sin embargo, que aún existe discusión sobre la exactitud de algunas de las cifras correspondientes a los años 1928, 1951, 1973 y 1993. No se tendrá en cuenta para los cálculos el dato censal del año 1905 por tener una discrepancia muy alta con el valor censal corregido, lo cual puede incluir ruido durante el proceso de cálculo.

En la Tabla 1 aparecen dos columnas de datos de población censada. La primera corresponde a los datos censales obtenidos entre 1905 y el año 1993. La segunda corresponde a los datos censales desde 1905 hasta el año 2005 que incluye correcciones y actualizaciones realizados por el DANE, por efectos de sobre y subestimación de los datos anteriores. Con estos datos corregidos se realizarán entonces, los diferentes cálculos de la curva logística.

**Tabla 1.**  
Datos de población para la ciudad de Medellín

<b>Año</b>	<b>Día</b>	<b>Presidente de la República</b>	<b>Población censada a 1993</b> <i>(habitantes)</i>	<b>Población censada y corregida a 2005</b> <i>(habitantes)</i>	$t_i$ <i>(año)</i>
1905*	1° al 15 de junio	Rafael Reyes Prieto	30740	53936	-44,55
1912	Marzo 5	Carlos E. Restrepo	65504	71004	-37,83
1918	Octubre 14	Marco Fidel Suárez	79146	79146	-31,22
1928	Noviembre 17	Miguel Abadía Méndez	120044	120044	-21,12
1938	Julio 5	Alfonso López Pumarejo	168266	168266	-11,49
1951	Mayo 9	Laureano Gómez Castro	358189	358189	1,35
1964	Julio 15	Guillermo León Valencia	772887	791589	14,54
1973	Octubre 24	Misael Pastrana Borrero	1092191	1163865	23,81
1985	Octubre 15	Belisario Betancur Cuartas	1418554	1480382	35,79
1993	Octubre 15	César Gaviria Trujillo	1551160	1834881	43,79
2005	Mes de mayo	Álvaro Uribe Vélez	-	2223078	55,41

Fuente: Datos suministrados por el DANE

$t_i$ : es el año censal, contado a partir de 1950: año que se calcula - año base (1950)

\*: no será considerado e los cálculos por las razones mencionadas.

Para un mejor entendimiento del procedimiento, se realizarán los cálculos paso a paso, empleando una hoja electrónica de cálculo (Excel®).

1. Para los diez datos censales de la columna de Población censada y corregida 2005 (Tabla 1) -omitiendo el dato correspondiente al año 1905- tómesese el origen de tiempo en cualquier momento (o fecha), y la unidad de tiempo que se desee: 1° de enero de 1950 (valor intermedio entre los años censados); y unidad de tiempo: año calendario.
2. Calcúlense los valores numéricos de  $T_1$ ,  $T_2$ ,  $T_3$ ,  $T_4$ .

En la Tabla 2 se obtienen los valores de los tiempos elevados a las diferentes potencias y sus respectivas sumatorias.

**Tabla 2**

$t_i$ (año)	$t_i^2$ (año <sup>2</sup> )	$t_i^3$ (año <sup>3</sup> )	$t_i^4$ (año <sup>4</sup> )
-37,83	1430,817	-54122,3069	2047238,31
-31,22	974,4995	-30420,9265	949649,2908
-21,12	446,1608	-9424,03993	199059,4607
-11,49	132,1005	-1518,29663	17450,53871
1,35	1,821853	2,459065427	3,319149229
14,54	211,2735	3070,912924	44636,48656
23,81	566,9679	13500,12334	321452,6289
35,79	1280,67	45830,62398	1640115,225
43,79	1917,253	83949,69795	3675859,054
55,41	3070,419	170136,1016	9427473,128
<b>73,02</b>	<b>10031,98</b>	<b>221004,35</b>	<b>18322937,44</b>

Los resultados son:

**Tabla 3**

<b>T1</b> (año)	73,02
<b>T2</b> (año <sup>2</sup> )	10031,98
<b>T3</b> (año <sup>3</sup> )	221004,35
<b>T4</b> (año <sup>4</sup> )	18322937,44

3. Con estos resultados se obtiene:

**Tabla 4**

<b>P</b> (año <sup>4</sup> )	$P = T_1 T_3 - T_2^2$	-84502682,1
<b>Q</b> (año <sup>3</sup> )	$Q = -nT_3 + T_1 T_2$	-1477495,85
<b>R</b> (año <sup>2</sup> )	$R = nT_2 - T_1^2$	94987,73466

4. Calcúlense cada uno de los valores numéricos de los  $n$  coeficientes

$$P + Qt_i + Rt_i^2 \text{ para } i = 1, 2, \dots, n.$$

**Tabla 5**

$t_i$ (año)	$P + Q \cdot t_i + R \cdot t_i^2$ (año <sup>4</sup> )
-37,83	107295391,7
-31,22	54185768,99
-21,12	-10914444,5
-11,49	-54973161,1
1,35	-86323893,8
14,54	-85910063,4
23,81	-65828467,2
35,79	-15729083,4
43,79	32918542,34
55,41	125279410,4

5. Fórmese la ecuación

$$f(L) = \sum (P + Q \cdot t_i + R \cdot t_i^2) \cdot \text{Ln} \left( \frac{L}{p_i} - 1 \right) = 0$$

$$\text{y sea } z_i = \text{Ln} \left( \frac{L}{p_i} - 1 \right)$$

6. Utilizando el método de la bisección, se exploran distintos valores numéricos para  $L$  hasta encontrar  $f(L)=0$ . Se establece arbitrariamente uno de ellos ( $L_0$ ) y otro ( $L_1$ ), entre los cuales  $f(L)$  cambia de signo:

Para seleccionar el primer valor de ( $L_0$ ) y encontrar más rápidamente el valor adecuado evitando un gran número de iteraciones innecesarias, podemos considerar que, como estamos buscando la población límite de la ciudad y el último valor de población censado según la Tabla 1 es  $p= 2'223,078$  habitantes entonces, este valor u otro muy cercano podría servir como punto de partida, p. e.  $L_0=2'223,079$  habitantes.

Con este valor se realiza una primera iteración para encontrar el resultado de  $f(L_0)$ , obteniendo un valor negativo, que se muestra en la Tabla 6.

**Tabla 6**

$t_i$ (años)	$P + Q \cdot t_i + R \cdot t_i^2$ (año <sup>4</sup> )	$z_i = \text{Ln} \left( \frac{L}{p_i} - 1 \right)$	$(P + Q \cdot t_i + R \cdot t_i^2) \cdot z_i$ (año <sup>4</sup> )
-37,83	107295391,7	3,41145156	366033031,1
-31,22	54185768,99	3,29910302	178764434,3
-21,12	-10914444,5	2,86327848	-31251093,89
-11,49	-54973161,1	2,50239406	-137564512,1
1,35	-86323893,8	1,64989657	-142425496,3
14,54	-85910063,4	0,59242882	-50895597,56
23,81	-65828467,2	-0,09421924	6202308,126
35,79	-15729083,4	-0,68976729	10849407,17
43,79	32918542,34	-1,55321939	-51129718,24
55,41	125279410,4	-14,6144033	-1830883826
		$f(L_0) =$	-1682301064

En una segunda iteración, seleccionando arbitrariamente un valor mayor que  $L_0$ , p.e., dos veces el valor de la última población censada se tiene que  $L_1 = 2M = 2 \cdot 2,223,078 = 4,446,156$  habitantes. Con este valor, el resultado para  $f(L_1)$  es positivo:

**Tabla 7**

$t_i$ (años)	$P + Q \cdot t_i + R \cdot t_i^2$ (año <sup>4</sup> )	$z_i = \ln\left(\frac{L}{p_i} - 1\right)$	$(P + Q \cdot t_i + R \cdot t_i^2) \cdot z_i$ (año <sup>4</sup> )
-37,83	107295391,7	4,12096033	442160052,4
-31,22	54185768,99	4,0105396	217314172
-21,12	-10914444,46	3,58456615	-39123548,19
-11,49	-54973161,11	3,23566912	-177874959,6
1,35	-86323893,83	2,43474228	-210176434
14,54	-85910063,37	1,52969058	-131415814,6
23,81	-65828467,25	1,03679529	-68250644,86
35,79	-15729083,39	0,69483788	-10929162,98
43,79	32918542,34	0,35285898	11615603,23
55,41	125279410,4	0	0
		<b><math>f(L_1)</math></b>	<b>33319263,39</b>

Iterando varias veces según el método de la bisección (ver ANEXO 1) entre estos dos valores de  $L_0 = 2'223,079$  habitantes y  $L_1 = 4'446,156$  habitantes, se continúa con el valor intermedio  $L_2 = 3334617,5$  habitantes, y así sucesivamente hasta llegar al valor de  $L = 3'442,234.382903$  habitantes con el cual se obtienen los resultados de la Tabla 8.

Puesto que  $f(L) = 2,26498E-06 \approx 0$ , puede considerarse que la población límite  $L$  para la ciudad de Medellín, según los datos censales al año 2005, es  $L = 3'442,234.382903$  habitantes.

**Tabla 8**

$t_i$ (años)	$P + Q \cdot t_i + R \cdot t_i^2$ (año <sup>4</sup> )	$z_i = \text{Ln} \left( \frac{L}{p_i} - 1 \right)$	$(P + Q \cdot t_i + R \cdot t_i^2) \cdot z_i$ (año <sup>4</sup> )
-37,83	107295391,7	3,86029684	414192061,6
-31,22	54185768,99	3,74932074	203159827,7
-21,12	-10914444,46	3,32052126	-36241644,82
-11,49	-54973161,11	2,96821204	-163171998,7
1,35	-86323893,83	2,15293667	-185849876,5
14,54	-85910063,37	1,20851611	-103823695,8
23,81	-65828467,25	0,67171364	-44217879,41
35,79	-15729083,39	0,28158896	-4429136,198
43,79	32918542,34	-0,13239066	-4358107,693
55,41	125279410,4	-0,60073359	-75259550,21
		$f(L) =$	<b>2,26498E-06</b>

7. Con estos datos se calculan ahora los valores de:

**Tabla 9**

$\sum z_i$	17,47998201
$\sum z_i \cdot t_i$ (año)	-360,1030421
$\sum z_i \cdot t_i^2$ (año <sup>2</sup> )	9966,700688

8. Luego se forma el determinante:

$$D = \begin{vmatrix} n & T_1 & T_2 \\ T_1 & T_2 & T_3 \\ T_2 & T_3 & T_4 \end{vmatrix} = \begin{vmatrix} 10 & 73.021 & 10031.98 \\ 73.021 & 10031.98 & 221004.3 \\ 10031.98 & 221004.3 & 18322937 \end{vmatrix} \text{ (año}^6\text{)}$$

Al calcular dicho determinante por la regla de Sarrus, con la ayuda de una hoja electrónica de cálculo u otra aplicación de *software*, se obtiene el resultado:

<b>D =</b>	<b>5,66192E+11</b> (año <sup>6</sup> )
------------	---

9. De aquí deducimos

$$b = - \frac{\begin{vmatrix} n & \sum z_i & T_2 \\ T_1 & \sum z_i t_i & T_3 \\ T_2 & \sum z_i t_i^2 & T_4 \end{vmatrix}}{D} = - \frac{\begin{vmatrix} 10 & 17.47998 & 10031.98 \\ 73.021 & -359.8565 & 221004.348 \\ 10031.98 & 9953.061 & 18322937.44 \end{vmatrix}}{5,66192E + 11 (\text{año}^6)} \quad (\text{año}^5)$$

$b =$	$0,051322155$ $(\text{año}^{-1})$
-------	--------------------------------------

10. De igual forma se procede con el cálculo de  $k$ :

$$k = \exp \frac{\begin{vmatrix} \sum z_i & T_1 & T_2 \\ \sum z_i t_i & T_2 & T_3 \\ \sum z_i t_i^2 & T_3 & T_4 \end{vmatrix}}{D} = \exp \frac{\begin{vmatrix} 17.47998 & 73.0212 & 10031.9833 \\ -359.8565 & 10031.9833 & 221458.7785 \\ 9953,0616 & 221004.3489 & 18322937.44 \end{vmatrix}}{5,66192E + 11 (\text{año}^6)} \quad (\text{año}^6)$$

$k =$	$8,354153403$
-------	---------------

11. Concluimos que la curva logística que mejor ajusta estos datos censales de Medellín está dada por la expresión:

$$\hat{p}(t) = \frac{3'442,234.38290}{1 + 8,354153403 \cdot e^{(-0,051322155 \cdot t / \text{año})}} \text{ habtes} \quad (14)$$

en donde

$$t = (\text{Año censado}) - 1950.$$

12. La tasa relativa de crecimiento poblacional para esta curva logística es:

$$\hat{r}(t) = \frac{Kb}{K + e^{bt}}$$

que, para el caso de Medellín vale:

$$\hat{r}(t) = \frac{0.428753155}{8.35415340 + e^{(0.05132215 \cdot t / \text{año})}} [\text{año}]^{-1} \quad (15)$$

$$t = \text{año} - 1950$$

De esta manera puede formarse una tabla que compara los valores históricos de los censos con la estimación retrospectiva que da la curva logística para los diferentes años desde 1905 inclusive, con las correspondientes tasas relativas instantáneas de crecimiento de la población; la tabla es la siguiente:

**Tabla 10.**

Tasa relativa de crecimiento y error relativo con datos censales (1905- 2005)

Fecha censal (DD/MM/AA)	ti (año)	Población censada (miles de hab.)	Población estimada (miles de hab.)	Tasa relativa de crecimiento por año según logística	Error relativo $\epsilon(t)$ de la logística	Cuadrado del error relativo $\epsilon(t)^2$
15/06/1905	-44,55	53,936	41,378	5,07%	-23,28%	0,0542
05/03/1912	-37,83	71,004	58,135	5,05%	-18,12%	0,0329
14/10/1918	-31,22	79,146	81,058	5,01%	2,42%	0,0006
17/11/1928	-21,12	120,044	133,938	4,93%	11,57%	0,0134
05/07/1938	-11,49	168,266	214,218	4,81%	27,31%	0,0746
09/05/1951	1,35	358,189	391,384	4,55%	9,27%	0,0086
15/07/1964	14,54	791,589	693,703	4,10%	-12,37%	0,0153
24/10/1973	23,81	1163,865	994,469	3,65%	-14,55%	0,0212
15/10/1985	35,79	1480,382	1476,559	2,93%	-0,26%	0,0000
15/10/1993	43,79	1834,881	1828,081	2,41%	-0,37%	0,0000
31/05/2005	55,41	2223,078	2316,083	1,68%	4,18%	0,0018

Fuente: Datos censales suministrados por el Dane.

De igual manera puede construirse una tabla con los valores de población proyectada y su respectiva tasa relativa de crecimiento, año por año desde el año 1900 hasta el año 2020 o más, si se desea:

**Tabla 11.**  
Población estimada para la ciudad de Medellín año por año (1900 – 2020) según la Curva Logística

Fecha	$t_i$ (año)	Población estimada (miles de hab.)	Tasa relativa de crecimiento según logística	Fecha	$t_i$ (año)	Población estimada (miles de hab.)	Tasa relativa de crecimiento según logística
01/01/1900	-50,00	31,370	5,09%	01/01/1942	-8,00	253,190	4,75%
01/01/1901	-49,00	33,006	5,08%	01/01/1943	-7,00	265,487	4,74%
01/01/1902	-48,00	34,727	5,08%	01/01/1944	-6,00	278,328	4,72%
01/01/1903	-47,00	36,536	5,08%	01/01/1945	-5,00	291,771	4,70%
01/01/1904	-46,00	38,439	5,07%	01/01/1946	-4,00	305,762	4,68%
01/01/1905	-45,00	40,441	5,07%	01/01/1947	-3,00	320,355	4,65%
01/01/1906	-44,00	42,543	5,07%	01/01/1948	-2,00	335,570	4,63%
01/01/1907	-43,00	44,752	5,07%	01/01/1949	-1,00	351,471	4,61%
01/01/1908	-42,00	47,075	5,06%	01/01/1950	0,00	367,990	4,58%
01/01/1909	-41,00	49,524	5,06%	01/01/1951	1,00	385,188	4,56%
01/01/1910	-40,00	52,091	5,05%	01/01/1952	2,00	403,085	4,53%
01/01/1911	-39,00	54,788	5,05%	01/01/1953	3,00	421,751	4,50%
01/01/1912	-38,00	57,623	5,05%	01/01/1954	4,00	441,101	4,47%
01/01/1913	-37,00	60,610	5,04%	01/01/1955	5,00	461,204	4,44%
01/01/1914	-36,00	63,741	5,04%	01/01/1956	6,00	482,076	4,41%
01/01/1915	-35,00	67,030	5,03%	01/01/1957	7,00	503,793	4,38%
01/01/1916	-34,00	70,485	5,03%	01/01/1958	8,00	526,252	4,35%
01/01/1917	-33,00	74,125	5,02%	01/01/1959	9,00	549,525	4,31%
01/01/1918	-32,00	77,937	5,02%	01/01/1960	10,00	573,625	4,28%
01/01/1919	-31,00	81,941	5,01%	01/01/1961	11,00	598,632	4,24%
01/01/1920	-30,00	86,146	5,00%	01/01/1962	12,00	624,420	4,20%
01/01/1921	-29,00	90,572	5,00%	01/01/1963	13,00	651,065	4,16%
01/01/1922	-28,00	95,207	4,99%	01/01/1964	14,00	678,573	4,12%
01/01/1923	-27,00	100,071	4,98%	01/01/1965	15,00	707,028	4,08%
01/01/1924	-26,00	105,176	4,98%	01/01/1966	16,00	736,277	4,03%
01/01/1925	-25,00	110,549	4,97%	01/01/1967	17,00	766,397	3,99%
01/01/1926	-24,00	116,170	4,96%	01/01/1968	18,00	797,386	3,94%
01/01/1927	-23,00	122,067	4,95%	01/01/1969	19,00	829,328	3,90%
01/01/1928	-22,00	128,251	4,94%	01/01/1970	20,00	862,042	3,85%
01/01/1929	-21,00	134,754	4,93%	01/01/1971	21,00	895,604	3,80%
01/01/1930	-20,00	141,554	4,92%	01/01/1972	22,00	930,001	3,75%
01/01/1931	-19,00	148,682	4,91%	01/01/1973	23,00	965,317	3,69%
01/01/1932	-18,00	156,151	4,90%	01/01/1974	24,00	1001,339	3,64%
01/01/1933	-17,00	163,999	4,89%	01/01/1975	25,00	1038,142	3,58%
01/01/1934	-16,00	172,198	4,88%	01/01/1976	26,00	1075,701	3,53%
01/01/1935	-15,00	180,784	4,86%	01/01/1977	27,00	1114,096	3,47%
01/01/1936	-14,00	189,773	4,85%	01/01/1978	28,00	1153,086	3,41%
01/01/1937	-13,00	199,208	4,84%	01/01/1979	29,00	1192,741	3,35%
01/01/1938	-12,00	209,055	4,82%	01/01/1980	30,00	1233,025	3,29%
01/01/1939	-11,00	219,355	4,81%	01/01/1981	31,00	1274,012	3,23%
01/01/1940	-10,00	230,127	4,79%	01/01/1982	32,00	1315,436	3,17%
01/01/1941	-9,00	241,420	4,77%	01/01/1983	33,00	1357,363	3,11%

**Tabla 11.a**  
Población estimada para la ciudad de Medellín año por año  
(1900 – 2020) según la Curva Logística

<b>Fecha</b>	<b><math>t_i</math></b> (año)	<b>Población esti- mada</b> (miles de hab.)	<b>Tasa relativa de crecimiento</b> según logística
01/01/1984	34,00	1399,748	3,05%
15/10/1985	35,79	1476,559	2,93%
01/01/1986	36,00	1485,808	2,92%
01/01/1987	37,00	1529,261	2,85%
01/01/1988	38,00	1572,963	2,79%
01/01/1989	39,00	1616,980	2,72%
01/01/1990	40,00	1661,012	2,66%
01/01/1991	41,00	1705,124	2,59%
01/01/1992	42,00	1749,256	2,52%
15/10/1993	43,79	1828,081	2,41%
01/01/1994	44,00	1837,473	2,39%
01/01/1995	45,00	1881,321	2,33%
01/01/1996	46,00	1924,960	2,26%
01/01/1997	47,00	1968,454	2,20%
01/01/1998	48,00	2011,510	2,13%
01/01/1999	49,00	2054,195	2,07%
01/01/2000	50,00	2096,458	2,01%
01/01/2001	51,00	2138,366	1,94%
01/01/2002	52,00	2179,642	1,88%
01/01/2003	53,00	2220,358	1,82%
01/01/2004	54,00	2260,472	1,76%
31/05/2005	55,41	2316,083	1,68%
01/01/2006	56,00	2338,853	1,65%
<b>01/01/2007*</b>	<b>57,00</b>	<b>2376,944</b>	<b>1,59%</b>
01/01/2008	58,00	2414,298	1,53%
01/01/2009	59,00	2450,988	1,48%
01/01/2010	60,00	2486,789	1,42%
01/01/2011	61,00	2521,783	1,37%
01/01/2012	62,00	2555,952	1,32%
01/01/2013	63,00	2589,372	1,27%
01/01/2014	64,00	2621,848	1,22%
01/01/2015	65,00	2653,464	1,18%
01/01/2016	66,00	2684,214	1,13%
01/01/2017	67,00	2714,175	1,09%
01/01/2018	68,00	2743,182	1,04%
01/01/2019	69,00	2771,319	1,00%
01/01/2020	70,00	2798,589	0,96%

\*: Fecha actual

## Gráficas

**La Gráfica 2** muestra la curva logística calculada con la ecuación (14), desde el año 1900 hasta el año 2070 según los datos censales tomados entre los años 1912 y 2005, estos representados por puntos negros. Se muestra también el dato censal correspondiente al año 1905, el cual no se tomó en cuenta para calcular la curva por razones expuestas anteriormente. El año tomado como base para los cálculos fue 1950.

Se observa como la curva toma una forma similar a la mencionada al principio de este documento, sobre comportamiento logístico (Gráfica 1, curva intermedia), con crecimiento ascendente desde una población inicial  $P(0)$  hasta la población límite  $L$ .

**En la Gráfica 3** se observa el comportamiento de la tasa relativa de crecimiento de la población calculada también, a partir de los datos censales tomados entre 1912 y 2005. Se observa como la tasa de crecimiento disminuye desde un valor de 5.09% en el año 1900, pasando por 1.70% en el año 2005 y tendiendo a un valor cercano a cero (0.09%) en la medida en que la población se acerca a su valor límite hacia el año 2070.

**La Gráfica 5** muestra en una cuadrícula semilogarítmica el comportamiento de la curva logística según la ecuación (4) versus el tiempo  $t$  en años, desde 1900 hasta 2070.

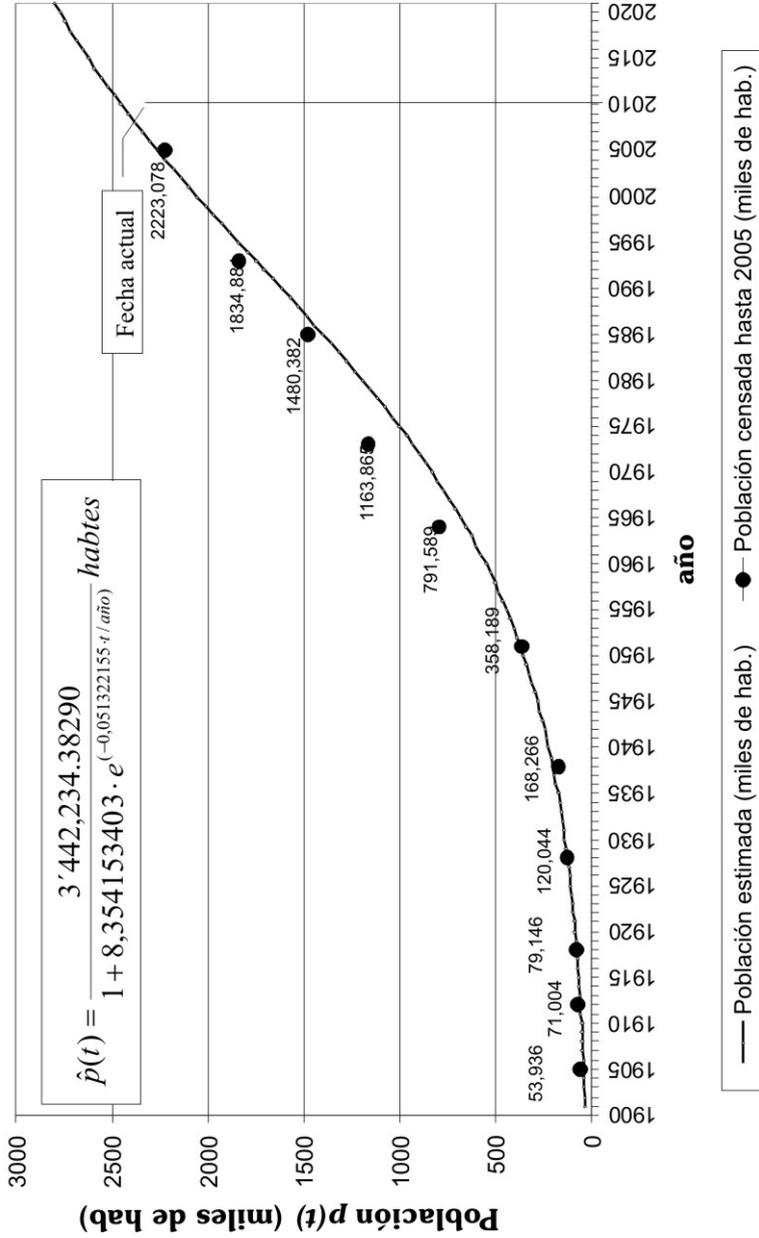
$$z = Ln\left(\frac{L}{p} - 1\right), \quad (04)$$

La curva tiende a ser lineal con pendiente negativa, iniciando en el valor 11.606 y terminado en un valor de 6.924 hacia el año 2070. Esto confirma el carácter exponencial creciente de la función logística. Los valores respectivos se muestran en la Tabla 12.

Como un aporte adicional al método descrito, los Anexos 4 y 5 presentan la aplicación del algoritmo de cálculo en el software estadístico de uso libre “R”, disponible en la web, y el correspondiente código de programación.

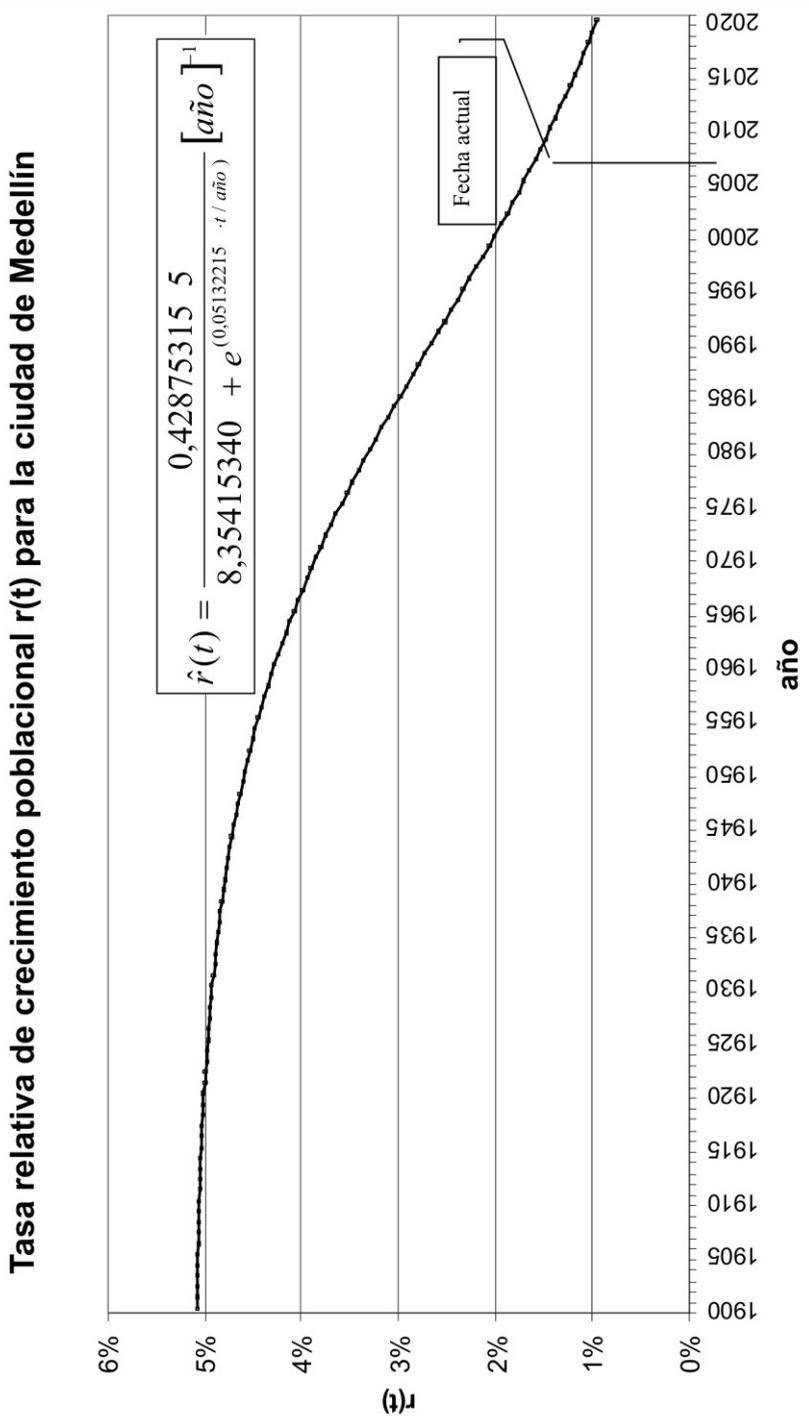
**Gráfica 2.**  
 Curva Logística de población calculada y proyectada para la ciudad de Medellín con censos entre 1912 y 2005

**Curva Logística calculada para la ciudad de Medellín**



Fuente: Datos censales suministrados por el Dane

**Gráfica 3.** Tasa relativa de crecimiento poblacional para la ciudad de Medellín, calculada con los datos censales tomados entre los años 1912 y 2005, proyectados hasta 2020.

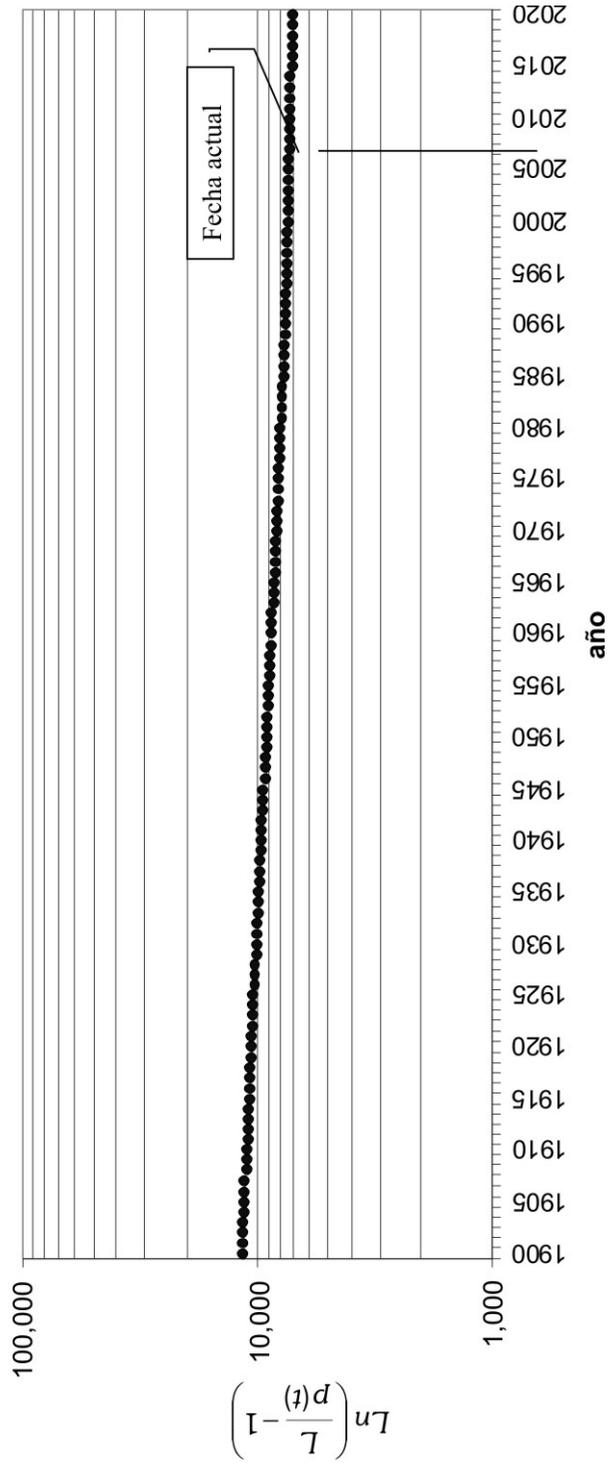


Fuente: Datos censales suministrados por el Dane

**Gráfica 4.**

Comportamiento de  $Ln\left(\frac{L}{p(t)} - 1\right)$ , versus  $t$ , trazado en cuadrícula semilogarítmica para la población de Medellín entre los años 1905 y 2020.

**Gráfica de  $Ln\left(\frac{L}{p(t)} - 1\right)$  vs  $t$  para la población de Medellín**



**Tabla 12.**  
Comportamiento del crecimiento relativo de la población

Año	$\text{Ln}\left(\frac{L}{p(t)} - 1\right)$	Año	$\text{Ln}\left(\frac{L}{p(t)} - 1\right)$	Año	$\text{Ln}\left(\frac{L}{p(t)} - 1\right)$
1900	11,60576781	1940	9,612932656	1980	7,934047449
1901	11,55492499	1941	9,565023706	1981	7,901335052
1902	11,50410716	1942	9,517417299	1982	7,869325866
1903	11,45331562	1943	9,469990005	1983	7,837937531
1904	11,40255171	1944	9,422749799	1984	7,80717722
1905	11,35178216	1945	9,375576325	1985	7,776969732
1906	11,30111259	1946	9,328735922	1986	7,747485859
1907	11,2504751	1947	9,282108261	1987	7,718647128
1908	11,19987134	1948	9,235702457	1988	7,690457675
1909	11,14916454	1949	9,189401738	1989	7,662846268
1910	11,0986336	1950	9,143468828	1990	7,635966269
1911	11,04814183	1951	9,097786879	1991	7,609742963
1912	10,99769123	1952	9,052366153	1992	7,584177112
1913	10,94714586	1953	9,007093875	1993	7,559201273
1914	10,89678411	1954	8,962228269	1994	7,534950958
1915	10,84647012	1955	8,917656338	1995	7,5113552
1916	10,79620633	1956	8,87338941	1996	7,488411326
1917	10,74585776	1957	8,829319092	1997	7,46605563
1918	10,69570221	1958	8,785698036	1998	7,444405936
1919	10,64560482	1959	8,742417414	1999	7,423395156
1920	10,5955685	1960	8,699489446	2000	7,403017398
1921	10,5454595	1961	8,656810437	2001	7,383212775
1922	10,49555484	1962	8,614626195	2002	7,364082153
1923	10,44572086	1963	8,572832306	2003	7,345562559
1924	10,39596108	1964	8,531441581	2004	7,327645351
1925	10,34614314	1965	8,490355242	2005	7,310274637
1926	10,29654312	1966	8,449810786	2006	7,293535532
1927	10,24702882	1967	8,409708354	2007	7,277369472
1928	10,19760441	1968	8,370060914	2008	7,261765701
1929	10,14813927	1969	8,330774699	2009	7,246672551
1930	10,09890828	1970	8,292077254	2010	7,232160924
1931	10,04978093	1971	8,253873307	2011	7,218176954
1932	10,0007622	1972	8,21617541	2012	7,204708484
1933	9,95172347	1973	8,178894768	2013	7,191708273
1934	9,902938145	1974	8,142247226	2014	7,179234739
1935	9,854277728	1975	8,106142008	2015	7,16723893
1936	9,805748103	1976	8,070590581	2016	7,155707988
1937	9,757223007	1977	8,035508954	2017	7,144599232
1938	9,708973992	1978	8,001099507	2018	7,133960365
1939	9,660874913	1979	7,967275786	2019	7,123747448
				2020	7,113947541

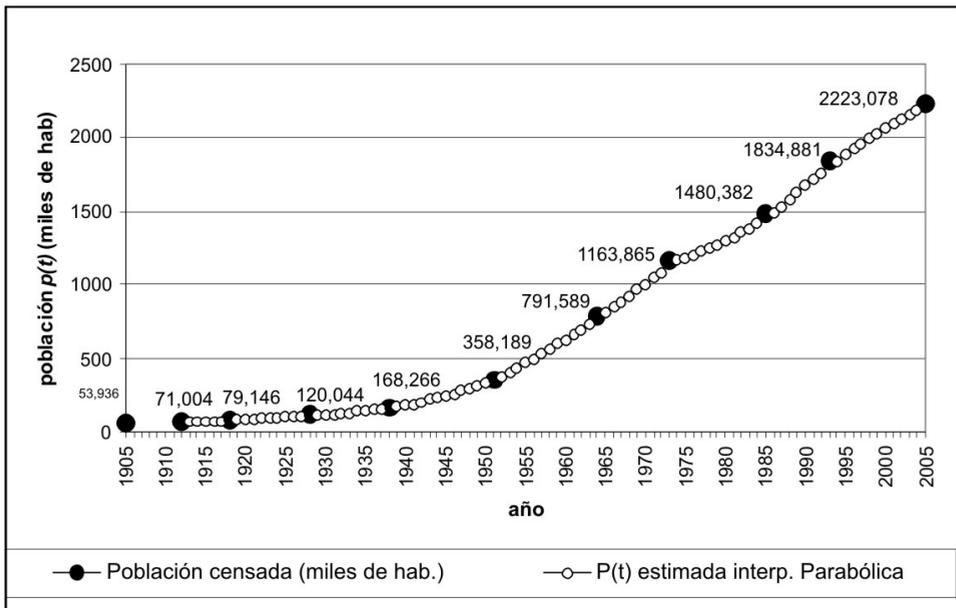
## Interpolación entre censos por el método de la parábola de segundo grado, año-a-año, de 1912 a 2005

Con el fin de determinar los valores de los datos de población intermedios año por año entre los datos censados, se ha realizado con estos, una interpolación por el método de la parábola o polinomial de segundo grado. En el ANEXO 4 se describe el método y sus respectivas ecuaciones. En la Gráfica 6 se observa su comportamiento y en la Tabla 13 se encuentran los resultados obtenidos.

Para este caso, se ha tomado cada punto final del rango a interpolar, como punto inicial del rango siguiente. Solo en el último rango de años (1985- 2005), se ha tomado el punto intermedio del penúltimo rango como punto inicial del rango final de interpolación.

### Gráfica 5.

Población estimada de la ciudad de Medellín, por el método de interpolación parabólica



Fuente: Datos censales suministrados por el Dane

**Tabla 13.**  
Población estimada de la ciudad de Medellín, por el método de interpolación parabólica

Año	Población censada (miles de hab.)	P(t) estimada interp. Parabólica	Año	Población censada (miles de hab.)	P(t) estimada interp. Parabólica	Año	Población censada (miles de hab.)	P(t) estimada interp. Parabólica
<b>1912</b>	<b>71,004</b>	<b>71,004</b>	1943		218,378	1974		1166,865
1913		71,216	1944		231,912	1975		1183,797
1914		71,779	1945		246,357	1976		1202,516
1915		72,680	1946		261,632	1977		1223,082
1916		73,918	1947		277,776	1978		1245,383
1917		75,498	1948		294,790	1979		1269,471
1918	79,146	79,146	1949		312,723	1980		1295,347
1919		79,662	1950		331,478	1981		1323,090
1920		82,249	<b>1951</b>	<b>358,189</b>	<b>358,189</b>	1982		1352,547
1921		85,182	1952		376,886	1983		1383,792
1922		88,445	1953		406,295	1984		1416,824
1923		92,045	1954		436,271	<u>1985</u>	<u>1480,382</u>	1480,382
1924		95,982	1955		466,892	1986		1490,770
1925		100,268	1956		498,160	1987		1538,707
1926		104,880	1957		530,161	1988		1585,533
1927		109,830	1958		562,723	1989		1631,371
<b>1928</b>	<b>120,044</b>	<b>120,044</b>	1959		595,930	1990		1675,971
1929		120,151	1960		629,783	1991		1719,460
1930		121,510	1961		664,377	1992		1761,838
1931		123,738	1962		699,524	<b>1993</b>	<b>1834,881</b>	<b>1834,881</b>
1932		126,835	1963		735,317	1994		1843,368
1933		130,813	1964	791,589	791,589	1995		1882,409
1934		135,651	1965		808,942	1996		1920,339
1935		141,358	1966		846,675	1997		1957,257
1936		147,935	1967		885,053	1998		1992,961
1937		155,402	1968		924,077	1999		2027,554
1938	168,266	168,266	1969		963,857	2000		2061,036
1939		172,905	1970		1004,175	2001		2093,494
1940		182,960	1971		1045,139	2002		2124,750
1941		193,916	1972		1086,749	2003		2154,895
1942		205,712	<b>1973</b>	<b>1163,865</b>	<b>1163,865</b>	2004		2183,929
						<u>2005</u>	<u>2223,078</u>	2223,078

Fuente: Datos censales suministrados por el Dane.

Los valores en negrilla de la Tabla 13 representan los valores iniciales y finales de los rangos de datos tomados para la interpolación por el método de la parábola. Los valores subrayados al final de la tabla, corresponden a los valores inicial y final tomados para interpolar el último rango.

La población promedio  $\bar{p}$  de la ciudad durante el período de 93 años comprendido entre 1912 y 2005, calculado con los valores obtenidos por interpolación parabólica de la Tabla 13 y según la ecuación (14) fue:

$$\bar{p} = \frac{1}{93 \text{ año}} \sum_{t=1}^{93} p(t) \quad (14)$$

$$\bar{p} = 811,864 \text{ habitantes}$$

Con este valor de población promedio se realizan los cálculos siguientes.

La varianza poblacional  $s_p^2$  es: 492,745.143hab., resultado de:

$$s_p^2 = \frac{1}{93 \text{ año}} \sum_{t=1}^{93} [p(t) - \bar{p}]^2 \quad (15)$$

El error cuadrático promedio para el rango de 93 años entre 1912 y 2005, es  $S_p^2 = \mathbf{0.013833}$  obtenido de la ecuación:

$$S_p^2 = \frac{1}{93a} \sum_1^{93} [\varepsilon(t)]^2 \quad (16)$$

siendo  $\varepsilon(t)$  el error relativo entre los datos censados (y los valores interpolados) de la población, y los valores estimados con la función logística.

El error relativo promedio se calcula con la siguiente expresión:

$$\varepsilon(t)_{prom} = \frac{\sum |\varepsilon(t)|}{\sum p(t)} = \frac{\sum |\varepsilon(t)|}{93a \cdot \bar{p}(t)} \quad (17)$$

que arroja como resultado un valor de 0,0001169.

En la Tabla 14 se muestran los valores correspondientes a los errores relativos y absolutos de las estimaciones de la población para el rango de 93 años comprendido entre 1912 y 2005.

El coeficiente de correlación (o de conformidad)  $r^2$  de la curva logística con los censos realizados, es 0,99999997 que se obtiene a partir de la siguiente expresión:

$$r^2 = 1 - \left( \frac{S_p^2}{s_p^2} \right) \quad (18)$$

donde

$s_p^2$ : varianza poblacional

$S_p^2$ : error cuadrático promedio

**Tabla 14.**

Población censada e interpolada vs población estimada

Año	P(t) interp parabólica (miles de hab)	Población estimada (miles de hab.)	Error relativo e(t) P(t)parab vs P(t)logist	Cuadrado del error relativo $\varepsilon(t)^2$	Valor absoluto del Error relativo   $\varepsilon(t)$
1912	71,004	58,135	-22%	0,04900	0,2214
1913	71,216	60,610	-17%	0,03061	0,1750
1914	71,779	63,741	-13%	0,01590	0,1261
1915	72,680	67,030	-8%	0,00710	0,0843
1916	73,918	70,485	-5%	0,00237	0,0487
1917	75,498	74,125	-2%	0,00034	0,0185
1918	79,146	81,058	2%	0,00056	0,0236
1919	79,662	81,941	3%	0,00077	0,0278
1920	82,249	86,146	5%	0,00205	0,0452
1921	85,182	90,572	6%	0,00354	0,0595
1922	88,445	95,207	7%	0,00504	0,0710
1923	92,045	100,071	8%	0,00643	0,0802
1924	95,982	105,176	9%	0,00764	0,0874
1925	100,268	110,549	9%	0,00865	0,0930
1926	104,880	116,170	10%	0,00944	0,0972
1927	109,830	122,067	10%	0,01005	0,1002
1928	120,044	133,938	10%	0,01076	0,1037

<b>Año</b>	<b>P(t) interp parabólica (miles de hab)</b>	<b>Población estimada (miles de hab.)</b>	<b>Error relativo e(t) P(t) parab vs P(t) logist</b>	<b>Cuadrado del error relativo <math>\varepsilon(t)^2</math></b>	<b>Valor absoluto del Error relativo <math> \varepsilon(t) </math></b>
1929	120,151	134,754	11%	0,01174	0,1084
1930	121,510	141,554	14%	0,02005	0,1416
1931	123,738	148,682	17%	0,02815	0,1678
1932	126,835	156,151	19%	0,03525	0,1877
1933	130,813	163,999	20%	0,04095	0,2024
1934	135,651	172,198	21%	0,04504	0,2122
1935	141,358	180,784	22%	0,04756	0,2181
1936	147,935	189,773	22%	0,04861	0,2205
1937	155,402	199,208	22%	0,04836	0,2199
1938	168,266	214,218	21%	0,04601	0,2145
1939	172,905	219,355	21%	0,04484	0,2118
1940	182,960	230,127	20%	0,04201	0,2050
1941	193,916	241,420	20%	0,03872	0,1968
1942	205,712	253,190	19%	0,03516	0,1875
1943	218,378	265,487	18%	0,03149	0,1774
1944	231,912	278,328	17%	0,02781	0,1668
1945	246,357	291,771	16%	0,02423	0,1557
1946	261,632	305,762	14%	0,02083	0,1443
1947	277,776	320,355	13%	0,01767	0,1329
1948	294,790	335,570	12%	0,01477	0,1215
1949	312,723	351,471	11%	0,01215	0,1102
1950	331,478	367,990	10%	0,00984	0,0992
1951	358,189	391,384	8%	0,00719	0,0848
1952	376,886	403,085	6%	0,00422	0,0650
1953	406,295	421,751	4%	0,00134	0,0366
1954	436,271	441,101	1%	0,00012	0,0110
1955	466,892	461,204	-1%	0,00015	0,0123
1956	498,160	482,076	-3%	0,00111	0,0334
1957	530,161	503,793	-5%	0,00274	0,0523
1958	562,723	526,252	-7%	0,00480	0,0693
1959	595,930	549,525	-8%	0,00713	0,0844
1960	629,783	573,625	-10%	0,00958	0,0979
1961	664,377	598,632	-11%	0,01206	0,1098
1962	699,524	624,420	-12%	0,01447	0,1203
1963	735,317	651,065	-13%	0,01675	0,1294
1964	791,589	693,703	-14%	0,01991	0,1411
1965	808,942	707,028	-14%	0,02078	0,1441
1966	846,675	736,277	-15%	0,02248	0,1499

<b>Año</b>	<b>P(t) interp parabólica</b> (miles de hab)	<b>Población estimada</b> (miles de hab.)	<b>Error relativo e(t) P(t)parab vs P(t)logist</b>	<b>Cuadrado del error relativo <math>\varepsilon(t)^2</math></b>	<b>Valor absoluto del Error relativo <math> \varepsilon(t) </math></b>
1967	885,053	766,397	-15%	0,02397	0,1548
1968	924,077	797,386	-16%	0,02524	0,1589
1969	963,857	829,328	-16%	0,02631	0,1622
1970	1004,175	862,042	-16%	0,02719	0,1649
1971	1045,139	895,604	-17%	0,02788	0,1670
1972	1086,749	930,001	-17%	0,02841	0,1685
1973	1163,865	994,469	-17%	0,02902	0,1703
1974	1166,865	1001,339	-17%	0,02733	0,1653
1975	1183,797	1038,142	-14%	0,01969	0,1403
1976	1202,516	1075,701	-12%	0,01390	0,1179
1977	1223,082	1114,096	-10%	0,00957	0,0978
1978	1245,383	1153,086	-8%	0,00641	0,0800
1979	1269,471	1192,741	-6%	0,00414	0,0643
1980	1295,347	1233,025	-5%	0,00255	0,0505
1981	1323,090	1274,012	-4%	0,00148	0,0385
1982	1352,547	1315,436	-3%	0,00080	0,0282
1983	1383,792	1357,363	-2%	0,00038	0,0195
1984	1416,824	1399,748	-1%	0,00015	0,0122
1985	1480,382	1476,559	0%	0,00001	0,0026
1986	1490,770	1485,808	0%	0,00001	0,0033
1987	1538,707	1529,261	-1%	0,00004	0,0062
1988	1585,533	1572,963	-1%	0,00006	0,0080
1989	1631,371	1616,980	-1%	0,00008	0,0089
1990	1675,971	1661,012	-1%	0,00008	0,0090
1991	1719,460	1705,124	-1%	0,00007	0,0084
1992	1761,838	1749,256	-1%	0,00005	0,0072
1993	1834,881	1828,081	0%	0,00001	0,0037
1994	1843,368	1837,473	0%	0,00001	0,0032
1995	1882,409	1881,321	0%	0,00000	0,0006
1996	1920,339	1924,960	0%	0,00001	0,0024
1997	1957,257	1968,454	1%	0,00003	0,0057
1998	1992,961	2011,510	1%	0,00009	0,0092
1999	2027,554	2054,195	1%	0,00017	0,0130
2000	2061,036	2096,458	2%	0,00029	0,0169
2001	2093,494	2138,366	2%	0,00044	0,0210
2002	2124,750	2179,642	3%	0,00063	0,0252
2003	2154,895	2220,358	3%	0,00087	0,0295
2004	2183,929	2260,472	3%	0,00115	0,0339
2005	2223,078	2316,083	4%	0,00161	0,0402

## Conclusiones

Con base en los datos censales procesados y los cálculos realizados pueden realizarse las siguientes observaciones sobre la demografía de la ciudad de Medellín:

- Luego de realizar el cálculo para deducir la ecuación de la curva con la cual se comparan los datos censales disponibles a la fecha, de la ciudad de Medellín, se observa como la curva de la Gráfica 2 toma una forma sigmoïdal similar a la curva intermedia de la Gráfica 1, sobre comportamiento logístico con crecimiento ascendente desde una población inicial  $P(0)$  hasta la población límite  $L$ .
- El método de cálculo utilizado es sencillo y los valores calculados presentan buen ajuste con los datos reales por lo tanto, con este método original, la curva logística se ajusta muy bien a los datos de los censos realizados.
- Debido a que los valores censados se encuentran distribuidos a lo largo de una curva, la interpolación parabólica parece ser la más adecuada, además de ser un método fácil de aplicar. El método usado aquí para interpolar puede afinarse o modificarse en varios sentidos, con cambios solo insignificantes en los resultados de este documento. Es decir, este es un método estadísticamente muy robusto.
- Entre los años de 1960 y mediados de los 80's, la población estuvo por encima de los valores estimados por la curva logística. Esto se explica por el fenómeno de explosión demográfica que se presentó en el país en esa época por varias razones, como la migración de campesinos a las ciudades en busca de mejores oportunidades a raíz de la violencia política de los años 50, entre otras, lo que generó una gran explosión demográfica en los 60's y que se redujo a finales de los 80's por la propagación en todos los estratos sociales de los métodos de planificación familiar (control de natalidad) [13].
- Hacia el año 2030, la población de la ciudad de Medellín (sin sus municipios vecinos de hoy 2007) habrá alcanzado los tres millones de personas, según los datos existentes a la fecha.

- La población seguirá aumentando durante los próximos años, aun cuando con tasas porcentuales declinantes.
- La tasa de crecimiento relativo poblacional ha venido reduciéndose desde 5,07% en el año 1912 a 1,68% en el año 2005. Para el año 2020 se espera un valor de 0,96%; lo cual indica que, de continuar la tendencia, el crecimiento se estabilizará al llegar al límite de población hacia el último cuarto del presente siglo.
- A mediados del siglo XXI la población estará en un valor cercano a 3'280,448 habitantes y estará acercándose asintóticamente al límite máximo de saturación de habitantes hacia el 2070 (valor calculado por fuera de las gráficas mostradas), es decir cerca de los 3'400,000 habitantes. A partir del valor de la población límite (3'442234 habitantes), cifra que la historia censal permite pronosticar, a la fecha, una de las interpretaciones que puede tener este valor es la indicación de que, a partir de este número de habitantes, se iniciará un proceso de hacinamiento, lo cual repercutirá en el incremento progresivo de problemas sociales por cada habitante nuevo en la ciudad.
- El hecho de tomar cualquiera de los años de la serie de años censados como origen del tiempo o base de cálculo, no influye en el cálculo de la curva logística. Tampoco influye, por este método, medir el tiempo en meses, lustros, quincenas u otra unidad de tiempo histórico.
- La curva logística permite pronosticar valores de población hacia el futuro y también, hacia el pasado, pues, con la ecuación logística solo deben reemplazarse los valores numéricos del año en el que se desea hacer la estimación demográfica.
- El coeficiente de correlación o de conformidad  $r^2=0,99999997$  resultante de la comparación entre los valores censados (incluidos los interpolados) durante todo el rango de tiempo entre 1912 y 2005, con la curva logística, es altamente satisfactorio, lo cual confirma la precisión del método.
- El comportamiento de la curva según la ecuación (4) en escala semilogarítmica, tiende a ser lineal decreciente, evidenciado las

características de la función exponencial creciente de la curva logística

- El método de cálculo que hemos ideado y presentado aquí además de ser sencillo de aplicar, es mucho más satisfactorio que los otros dos o tres métodos que se encuentran en la literatura sobre este tema (ver bibliografía al final).
- Tanto en la hoja de cálculo de Excel ® , como en el programa para estadística “R”, los resultados fueron consistentes.
- El método de bisección presenta varios puntos a favor, siendo el más importante de estos, su sencillez conceptual, su facilidad de interpretación e implementación; no se requiere mucho poder de cómputo para su ejecución pues, puede calcularse en herramientas tan sencillas como una hoja de Excel® y además, proporciona una estimación predeterminada de la precisión de la solución calculada. Es de anotar que con el conocimiento y aplicación de herramientas computacionales actuales, es muy fácil usar los métodos de interpolación y convergencia existentes.
- Los resultados de este ejercicio, pueden servir de insumo para procesos de planeación estratégica de los recursos y políticas de la ciudad y el área metropolitana.

### **Agradecimientos:**

Los autores expresan su agradecimiento los profesores Freddy Hernández B. y Jair Montoya, de la Escuela de Ingenierías de la Universidad Pontificia Bolivariana de Medellín, por su colaboración con los códigos de programación para el cálculo de la curva logística en el programa R y del pseudocódigo para el método de bisección, respectivamente.

## **Bibliografía**

- Poveda R.G., Curva Logística y Mínimos Cuadrados. Sociedad Colombiana de Matemáticas . Lecturas Matemáticas, Vol 3 [2] (1982). [Artículos matemáticos]. [www.scm.org.co](http://www.scm.org.co)
- Lotka, A.J., Elements of Mathematical Biology, New York, Dover Publications, 1956, 465 p.
- Gini, C., Teorías de la Población, Editorial Aguilar, Madrid, 1955, 420 p.
- De Lyra Madeira, J., “A Curva Logística”, Revista Brasileira de Estadística, Año XVII, N° 68, Río de Janeiro, 1957.
- Toranzos, F.I., Estadística, Buenos Aires, Editorial Kapelusz, 1962, 370 p.
- Davis, H.T., Introduction to Nonlinear Differential and Integral Equations, New York, Dover Publications, 1956, 465 p.
- Arkin, H. y Colton, R., An Outline of Statistical Methods, New York, Barnes and Noble Inc., 4ª. ed., 1950, 224 p.
- Ralston, A., A First Course in Numerical Analysis, Tokio, Kogakusha Company, 1965, 578 p.
- Scheid, F., Numerical Analysis, New York, Mc. Graw Hill Company, 1968, 422 p.
- Montgomery, D. y Rounge, G. Probabilidad y estadística aplicada a la ingeniería. México. Mc Graw Hill, 1997.
- Engel, A. Elementos de biomatemática. Washington, D.C. Secretaría General de la Organización de los Estados Americanos. Programa Regional de Desarrollo científico y Tecnológico. 76 p. 1978.
- Burden, R. y Faires, J. Análisis numérico. Sexta ed. México, Internacional Thomson Editores. 811 p.
- Poveda R. G., Historia económica de Colombia en el Siglo XX. Medellín, Editorial Universidad Pontificia Bolivariana. 2005. 790p.
- Páginas web visitadas:
- Mora, W. y Espinoza, J., El método de la bisección. [online]. [citado en oct. 29 de 2006]. Disponible en :
- <http://www.cidse.itcr.ac.cr/cursos-linea/NUMERICO/SitioWebEcuaciones/node3.html>

## ANEXO 1. Método de la Bisección

Este es uno de los métodos más sencillos para resolver ecuaciones en una variable. Se basa en el Teorema de los Valores Intermedios, el cual establece que toda función continua  $f$  en un intervalo cerrado  $[a, b]$  toma todos los valores que se hallan entre  $f(a)$  y  $f(b)$ . Esto es, que todo valor entre  $f(a)$  y  $f(b)$  es la imagen de al menos un valor en el intervalo  $[a, b]$ . En caso de que  $f(a)$  y  $f(b)$  tengan signos opuestos  $f(a) \cdot f(b) < 0$ , el valor cero sería un valor intermedio entre  $f(a)$  y  $f(b)$ , por lo que con certeza existe un punto  $p$  en  $[a, b]$  que cumple  $f(p)=0$ . De esta forma, se asegura la existencia de al menos una solución de la ecuación  $f(x)=0$  [12] y [14].

El método consiste en lo siguiente:

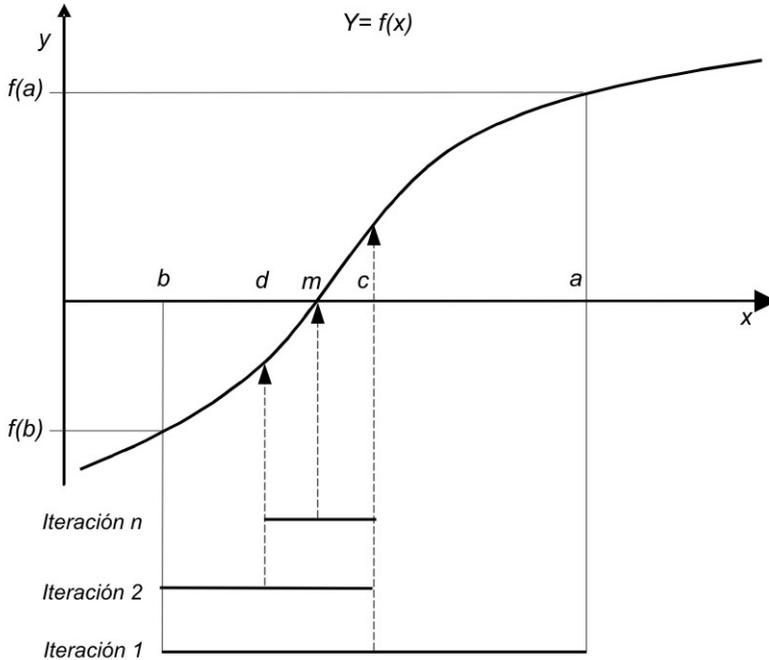
Verificar la continuidad de la función  $f(x)$  en el intervalo  $[a, b]$ - ver numeral 4 del Ajuste de la logística por mínimos cuadrados.

Luego se verifica que  $f(a) \cdot f(b) < 0$

$$L = \lim_{n \rightarrow \infty} L_n \quad L_{n+1} = \frac{1}{2}(L_n + L_{n+1}) \quad L_n L_{n-1} < 0$$

Se calcula el punto medio  $c$  del intervalo  $[a, b]$ . A continuación se calcula  $f(c)$ . En caso de que  $f(c)$  sea igual a cero, la raíz buscada será  $c$ . En caso de que no lo sea, se verifica si  $f(c)$  tiene signo opuesto con  $f(a)$  o con  $f(b)$ . Se redefine el intervalo  $[a, b]$  como  $[a, c]$  o  $[c, b]$  según se haya determinado en cuál de estos intervalos ocurre un cambio de signo. Con este nuevo intervalo se continúa sucesivamente encontrando el centro y encerrando la solución en un intervalo cada vez más pequeño, hasta alcanzar la precisión deseada. En la Gráfica 6 del presente apéndice, se ilustra el procedimiento descrito en el cual, el cambio de signo ocurre inicialmente entre  $[a, b]$ ; se busca el punto intermedio  $c$  y se redefine el rango  $[a, c]$ . Se verifica el cambio de signo y finalmente, entre  $[d, c]$  se encuentra el punto intermedio  $m$  como raíz de  $f(x)$ .

**Gráfica 6.**  
Ilustración del método de la bisección



Desde la primera iteración del algoritmo de bisección, se observa que la raíz  $m$  se halla a una distancia menor o igual que  $\frac{b-a}{2}$ . En la segunda iteración, el nuevo intervalo  $[a, c]$  mide  $\frac{b-a}{2}$  y de nuevo la distancia entre el nuevo punto medio y  $m$  es menor o igual que  $[d, c] \left(\frac{b-a}{2}\right) \div 2 = \frac{b-a}{4}$ . Si se continúa así sucesivamente, en la  $n$ -ésima iteración, al aproximar la raíz  $p$  de la función  $f(x)$ , con el punto de acercamiento  $m$  (renombrado como el punto medio durante  $n$  iteraciones  $m_n$ ) se tiene que:

$$|m_n - p| \leq \frac{b-a}{2^n} \quad (19)$$

y según esta desigualdad,

$$n = \frac{\ln\left(\frac{b-a}{\varepsilon}\right)}{\ln 2} \quad (20)$$

$n$  en su valor entero, es el número de iteraciones necesarias para aproximar la solución de la ecuación mediante el punto medio con un parámetro máximo  $\varepsilon$  de error.

En vista de que el uno de los objetivos del presente trabajo es su fácil interpretación por cualquier persona interesada en las solución de este tipo de problemas, hemos decidido utilizar el método de bisección el cual, presenta entonces, varios puntos a favor: el más importante, su sencillez conceptual, su facilidad de interpretación e implementación; no se requiere mucho poder de cómputo para su ejecución y además, este método proporciona una estimación predeterminada de la precisión de la solución calculada, según la ecuación (20).

En algunas ocasiones, otros métodos pueden converger más rápidamente como el de Newton, sin embargo, requiere más operaciones de procesamiento y conceptualmente son un poco más complejos. Actualmente existen muchos métodos de interpolación o convergencia y con el conocimiento y aplicación de herramientas computacionales, es muy fácil ahora usarlos todos.

A continuación se presenta el pseudo-código del método de bisección para que pueda ser implementado en cualquier lenguaje de programación:

### **Algoritmo:**

Entrada: extremos  $a$ ,  $b$ ; tolerancia  $TOL$ ; número máximo de iteraciones  $N0$

Salida: solución aproximada  $p$  o mensaje de error

Paso 1 tome  $i=1$

$FA=f(a)$

Un ejemplo muy elemental de código para aplicar el método de bisección podría ser el siguiente:

```
clc
clear all
close all
fprintf('Bienvenido al programa de bisección:\n\n');
```

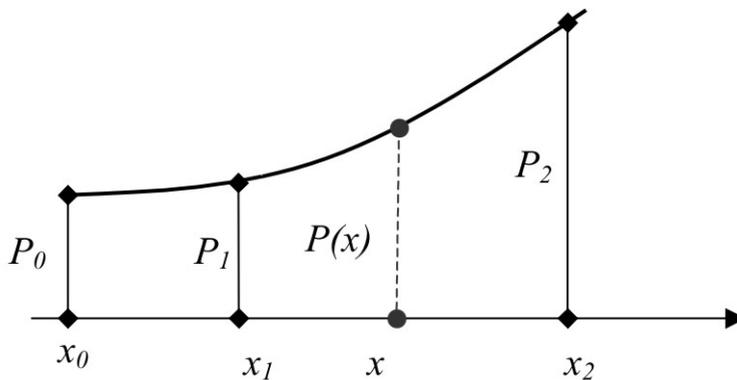
```
c=input('ingrese f(x)= ','s');
f = inline(c);
a=input('ingrese a=');
b=input('ingrese b=');
tol = input('ingrese tol=');
e=input('lmax=');
i=1;
while i<=e
    p=(a+b)/2;
    if (f(p)<=tol) & (b-a<=tol)
        fprintf('La solucion es:%f\n',p);
        break
    end
    i=i+1;
    if f(a)*f(b)<0
        b=p;
    else
        a=p;
    end
end
end
```



### ANEXO 3. Interpolación por el método de la parábola

Cuando se tienen tres puntos en un plano, puede determinarse la ecuación de un polinomio de segundo grado (una parábola) que contenga dichos puntos. De esta forma, pueden encontrarse los puntos intermedios que pertenezcan a la curva, entre los puntos extremos conocidos. Si se tienen cuatro puntos conocidos, el polinomio será de tercer grado y así sucesivamente. Este procedimiento se denomina interpolación polinomial de Lagrange, y específicamente para tres puntos conocidos, interpolación cuadrática o parabólica.

**Gráfica 7.**  
Interpolación parabólica



El teorema que define el  $n$ -ésimo polinomio interpolante de Lagrange es el siguiente [12]:

Si  $x_0, x_1, \dots, x_n$  son  $n+1$  números distintos y  $f$  es una función cuyos valores están dados en esos números, entonces existe un único polinomio  $P$  de grado a lo más  $n$ , con la propiedad de que

$$f(x_k) = p(x_k) \quad \text{para cada } k=0, 1, \dots, n$$

Este polinomio está dado por

$$P(x) = f(x_0)L_{n,0}(x) + \dots + f(x_n)L_{n,n}(x) = \sum_{k=0}^n f(x_k)L_{n,k}(x) \quad (18)$$

Donde

$$L_{n,k}(x) = \frac{(x-x_0)(x-x_1)\dots(x-x_{k-1})(x-x_{k+1})\dots(x-x_n)}{(x_k-x_0)(x_k-x_1)\dots(x_k-x_{k-1})(x_k-x_{k+1})\dots(x_k-x_n)} = \prod_{\substack{i=0 \\ i \neq k}}^n \frac{(x-x_i)}{(x_k-x_i)} \quad (19)$$

Para cada  $k=0,1,\dots,n$

$L_{n,k}(x)$  suele escribirse como  $L_k(x)$ .

Al aplicar esta ecuación a los puntos de la gráfica se obtiene:

$$p(x) = \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} p_0 + \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} p_1 + \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} p_2$$

Al evaluar el polinomio en los puntos conocidos:

$$P(x_0) = p_0$$

$$P(x_1) = p_1$$

$$P(x_2) = p_2$$

Otra forma de escribir el polinomio de interpolación es, mediante los determinantes de las ecuaciones simultáneas de segundo grado que resultan al evaluar cada uno de los tres puntos conocidos de la parábola de la Gráfica 8. Para el caso del punto  $p_0$ , la ecuación es:

$$p(x_0) = a_0 + a_1 x_0 + a_2 x_0^2$$

y la ecuación para cualquier punto  $p(x)$ :

$$p(x) = \frac{D_0}{D} - \frac{D_1}{D} x + \frac{D_2}{D} x^2 \quad (20)$$

Donde

$$D = \begin{vmatrix} 1 & x_0 & x_0^2 \\ 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \end{vmatrix}, \quad D_0 = \begin{vmatrix} p_0 & x_0 & x_0^2 \\ p_1 & x_1 & x_1^2 \\ p_2 & x_2 & x_2^2 \end{vmatrix}, \quad D_1 = \begin{vmatrix} p_0 & 1 & x_0^2 \\ p_1 & 1 & x_1^2 \\ p_2 & 1 & x_2^2 \end{vmatrix}, \quad D_2 = \begin{vmatrix} p_0 & 1 & x_0 \\ p_1 & 1 & x_1 \\ p_2 & 1 & x_2 \end{vmatrix}$$

Los cuales, luego de ser solucionados por la regla de Sarrus, se traducen en las siguientes expresiones, que al reemplazarse en la ecuación (20), permiten realizar la interpolación correspondiente. Este procedimiento puede hacerse con una calculadora programable o una hoja de cálculo:

$$D = x_0x_1(x_1 - x_0) + x_1x_2(x_2 - x_1) - x_2x_0(x_2 - x_0) \quad (21)$$

$$D_0 = p_0x_1x_2(x_2 - x_1) - p_0x_2x_0(x_2 - x_0) + p_2x_0x_1(x_1 - x_0) \quad (22)$$

$$D_1 = p_0(x_2^2 - x_1^2) - p_1(x_2^2 - x_0^2) + p_2(x_1^2 - x_0^2) \quad (23)$$

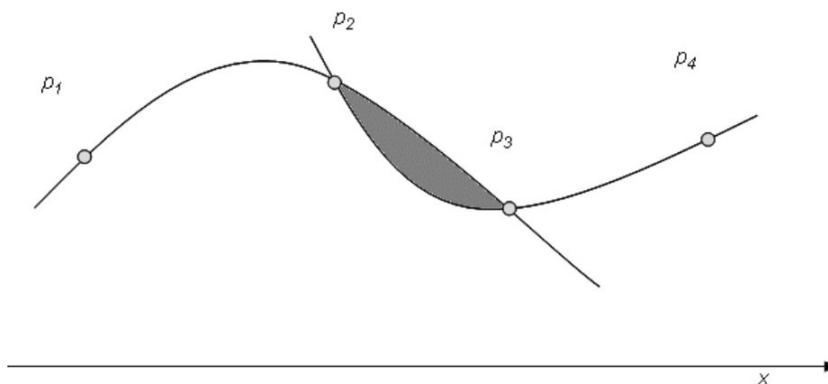
$$D_2 = p_0(x_2 - x_1) - p_1(x_2 - x_0) + p_2(x_1 - x_0) \quad (24)$$

Existen dos formas de realizar la interpolación parabólica con tres puntos conocidos: una de ellas es, tomando el tercer punto como origen del rango siguiente y así sucesivamente (ver Gráfica 10). La otra consiste en tomar el punto intermedio como origen del rango siguiente, lo cual se muestra en la Gráfica 8, donde  $p_1$ ,  $p_2$  y  $p_3$  son los puntos conocidos del primer rango de interpolación. El segundo rango lo conforman los puntos  $p_2$ ,  $p_3$  y  $p_4$  y así sucesivamente.

Puede observarse como los valores de la función entre  $p_2$  y  $p_3$  pueden ser diferentes según la parábola seleccionada y se encuentran en el área sombreada. Por lo tanto, los valores obtenidos por este método, presentan alguna discrepancia en este trayecto.

### Gráfica 8.

Interpolación parabólica con un punto traslapado

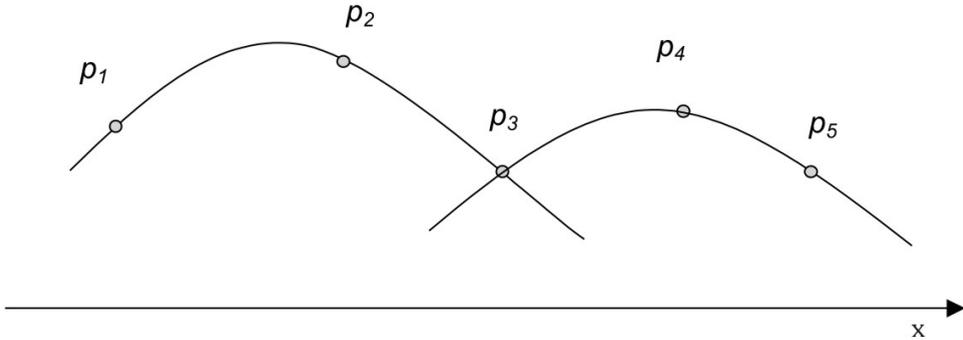


En ambos casos, los valores resultantes de la interpolación, son aproximados y los métodos se consideran válidos.

El método a utilizar en este ejercicio será el de interpolación mediante parábolas consecutivas representado en la Gráfica 10:

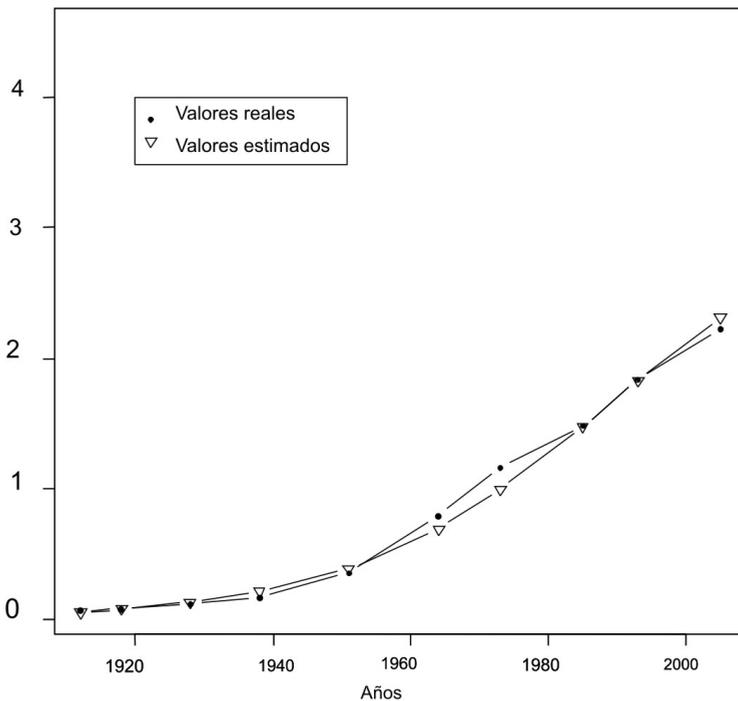
### Gráfica 9.

Interpolación parabólica mediante parábolas consecutivas



ANEXO 4. Curva logística calculada a partir de los datos censales de la ciudad de Medellín de 1912 a 2005- Resultados obtenidos con el programa estadístico “R”.

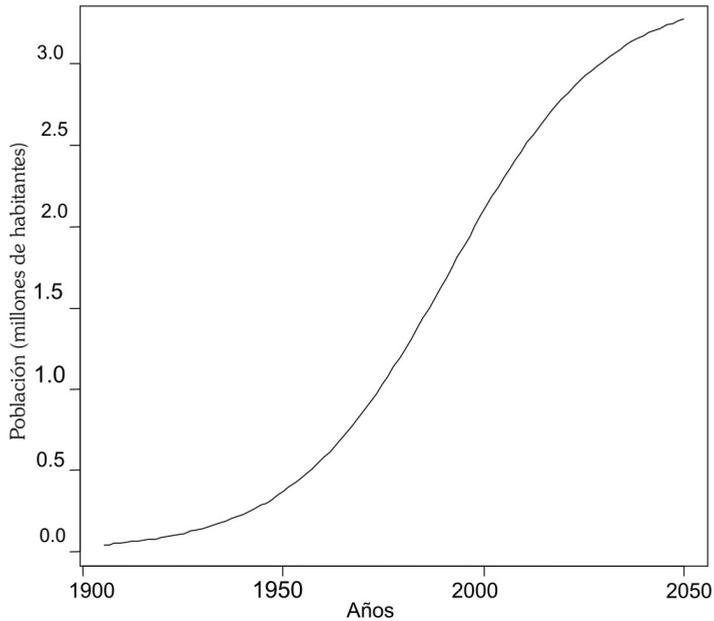
### Curva Logística censos Medellín siglo XX



La ecuación de la curva logística que se encuentra es:

$$p(t) = \frac{3,442,234.382903}{1 + 8.355437e^{-0.05131809(t-50)}} \text{ [habtes]}$$

**Población estimada para Medellín desde 1905 hasta 2050**



<b>Año</b>	<b>Población (habitantes)</b>	<b>t<sub>i</sub> (años)</b>	<b>Estimado (habitantes)</b>	<b>Error de estimación (%)</b>
1912	71004	-37.83	58123.54	-0.18
1918	79146	-31.22	81043.61	0.02
1928	120044	-21.12	133945.87	0.12
1938	168266	-11.49	214232.39	0.27
1951	358189	1.35	391333.17	0.09
1964	791589	14.54	693719.98	-0.12
1973	1163865	23.81	994252.01	-0.15
1985	1480382	35.79	1476460.75	0.00
1993	1834881	43.79	1827952.83	0.00
2005	2223078	55.41	2315743.08	0.04

Resultados parciales del algoritmo de solución:

<b>Indicador</b>	<b>Valor</b>
T1	73.03 años
T2	10032.78 años <sup>2</sup>
T3	221006.4 años <sup>3</sup>
T4	18324941 años <sup>4</sup>
Pop. Límite	3442234.382903 habitantes
P	-84516542 años <sup>4</sup>
Q	-1477370 años <sup>3</sup>
R	94994.4 años <sup>2</sup>
D	566322744125 años <sup>6</sup>
b	0.05131809 años <sup>-1</sup>
K	8.355437

ANEXO 5. Código para calcular la curva logística en el programa “R”.

```

require(xtable)
datos=read.table(file="C:/Documents and Settings/Administrador/
Mis documentos/Trabajos en R/Censos.csv",header=T,sep=";")
attach(datos)
datos
T1=sum(ti[-1])
T2=sum(ti[-1]^2)
T3=sum(ti[-1]^3)
T4=sum(ti[-1]^4)
T1 ; T2 ; T3 ; T4
P=T1*T3-T2^2
Q=-length(ti[-1])*T3+T1*T2
R=length(ti[-1])*T2-T1^2
P ; Q ; R
L=seq(from=3443685.01,to=3443687.01,length=10)
FL=matrix(0,ncol=1,nrow=length(L))
for (i in 1:length(L)) {
FL[i,1]= sum( (P+Q*ti[-1]+R*ti[-1]^2) * log(L[i]/Poblacion[-1]-1) )
}

```

```

cbind(L,FL)
plot(x=L,y=FL)
abline(a=0,b=0)
Lopt=3442234.382903
Z=log(Lopt/Poblacion[-1]-1)
A=sum(Z)
B=sum(Z*ti[-1])
C=sum(Z*ti[-1]^2)
A;B;C
D=det(matrix(c(length(ti[-1]),T1,T2,T1,T2,T3,T2,T3,T4),ncol=3,byr
ow=T))
b=-det(matrix(c(length(ti[-1]),A,T2,T1,B,T3,T2,C,T4),ncol=3,byrow
=T))/D
K=exp((det(matrix(c(A,T1,T2,B,T2,T3,C,T3,T4),ncol=3,byrow=T))
/D))
D;b;K
Estimados=Lopt/(1+K*exp(-b*ti[-1]))
#####
plot(x=Ano[-1],y=Poblacion[-1]/1000000,type='b',ylim=c(70000,
4500000)/1000000,
main='Curva Logística censos Medellín siglo XX',xlab='Años',
ylab='Población (millones de habitantes)',pch=20)
par(new=T)
plot(x=Ano[-1],y=Estimados/1000000,type='b',ylim=c(70000,45
00000)/1000000,
xlab='',ylab='',pch=25)
legend(x=1920,y=4,legend=c('Valores reales','Valores
estimados'),pch=c(20,25))
curve((Lopt/(1+K*exp(-b*(x-1950))))/10^6,from=1905,to=2050,
main='Población estimada para Medellín \n desde 1905 hasta
2050',
xlab='Años',ylab='Población (millones de habitantes)')
#####
Error=round((Estimados-Poblacion[-1])/Poblacion[-1],2)
cbind(datos[-1,],Estimados,Error)

```