KENYA'S POPULATION CALCULATIONS AND PROJECTIONS USING A NUMERICAL APPROACH

BY

NYAMBANE NYABUTO HARON

A Dissertation Submitted to Graduate School, In Partial Fulfillment for the Requirements of the Degree of Master of Science in Applied Mathematics in the Department of Mathematics
Egerton University

EGERTON UNIVERSITY, NJORO (KENYA)

AUGUST 2002

DECLARATION

This dissertation is my original work and has not been submitted to any university for award of a degree or diploma.

NAME: NYAMBANE NYABUTO HARON

Reg. No. S45/0265/2000

Signature ____________

Date: 6/10/02

RECOMMENDATION

This dissertation was submitted for examination with my approval as university supervisor.

NAME: DR. JOHN KROP LONYANGAPUO
SENIOR LECTURER, DEPARTMENT OF MATHEMATICS
EGERTON UNIVERSITY

Signature ____________

Date: 6/10/02

COPYRIGHT

DEDICATION

This dissertation is dedicated to my old parents, Mr. Livingstone Nyambane Moenga and Mrs. Miriam Mokobi Nyambane. It is also dedicated to my wife Esther Mwango Nyabuto and my daughters Maureen, Emily and Naomi for their continual encouragement and financial support throughout my study and research period.

ACKNOWLEDGEMENT

I would like to thank Dr. John K. Lonyangapuo for his tireless help and guidance throughout my work in this dissertation. I also thank Egerton University for offering me admission to pursue master of science in Applied Mathematics. Much thanks goes to my lecturers in Applied Mathematics, Prof. S. Singh, Dr. S. Thorof and Dr. J. K. Lonyangapuo.

I also thank my wife Esther Mwango Nyabuto, my daughters Maureen, Emily and Naormi for their encouragement and patience during my long stay away from home.

Thanks to Dr. Ali Islam, COD Mathematics Department for allowing me unlimited use of Departmental computers in printing my dissertation and encouraging me to work hard. Much thanks to my colleagues George Nzenge and John Wahome for their support and encouragement as we studied together during the entire study period.

Lastly, I thank Macrice Makoye for her effort in typing my dissertation. May God bless abundantly all the persons mentioned above.

ABSTRACT

The aim of this dissertation is to employ numerical methods in calculating and projecting a given population. In particular, Kenya's population has been used as illustrative example and results compared with those obtained when statistical methods are employed. The numerical results obtained are tabulated and put on bar graphs and line graphs for easy comparison.

Finally, other factors that affect population growth rate such as mortality from AIDS, fertility rates and mortality from other causes are also discussed.

Summary

The aim of the project is to employ numerical methods in calculating and projecting population and compare the results with that obtained when statistical methods are used. A forward, central and backward differences tables are generated so that the numerical methods can be used.

After establishing the statistical equations (methods) used and the values obtained as calculated population and projected population, we employ our numerical methods to establish the accuracy of the method as compared to the enumerated value during the census and the values obtained by the use of statistical methods. The various numerical methods are used and the resulting solution is evaluated numerically.

There are many numerical methods that can be used but for the case of population calculation and projection we employ three formulae namely;

(i) Gregory – Newton forward and backward formula

(ii) Striling formula

(iii) Bessel's formula

Here we just apply the formula directly using the enumerated values during census as the base population. The basis of the numerical method in this case was the specification of the statistical value in projecting Kenya's population. The next approach was to specify the values in graphs i.e. bar graphs and line graphs and the make comparison in the trend from the graphs rather than the new compar values as raw data, thus allowing an explicit selection of values by both statistical method and numerical method.

Also the differences between the projected values by statistical method and numerical method are obtained and their absolute values taken and then express them as the percentage of the statistical projected values.

Projection per province is made and then the sum of the projected population is obtained for all the provinces and comparison is made when projection is done on Kenya's population (entire population) of all the provinces.

Finally a mention of AIDS/HIV and mortality from AIDS, fertility rate and mortality from other causes is made.

Table of Contents

List of Tables

List of Figures

NOMENCLATURE

$y = f(x)$ a continuous function

$y_i = f(x_i)$ a continuous function known at discrete points.

$P= (x)$ a polynomial

$P(\bar{x})$ an estimate of $f(\bar{x})$

$x_o\, x_1$ ……….… x_n ordered discrete points.

$x_o < x_1 <$ ……….… $< x_n$

x_o is less than x_1 and x_1 is less than x_2 and so on up to x_{n-1} is less than x_n

$\bar{x}$ a point in the internal $[x_o, x_n]$ or outside the interval.

p_n the year on which projection is done.

P_o the base year's population.

r the intercensal growth rate.

t the time interval between the base year and the year on which projection is done.

$a \leq x \leq b$, a is less than or equal to x and x is less than or equal to b.

$x < a$ x is less than a

$x > b$ x is greater than b

$lnf(x)$ natural logarithm of the function $f(x)$.

$ln(x)$ natural logarithm of the discrete point x

Δ a forward difference operator

∇ a backward difference operator

δ a central difference operator

$p(s)$ an approximate of the function $f(x)$.

$$s = \frac{x - x_o}{h}$$

x a discrete point on which approximation is done

x_o the initial discrete point

h intercensal period.

f_o initial function value.

f' first derivative of the function $f(x)$

f'' second derivative of the function $f(x)$

f''' first derivative of the function $f(x)$

y_o initial value of the function.

y_n final value of the function.

CHAPTER 1

INTRODUCTION

1.0 Mathematical Model

Often scientific experimentation or numerical computation results in values for a function only at discrete points along the independent variable for that function. In other words in many problems in engineering and science, the data being considered are known only at a set of discrete points not as a continuous function. For example the continuous function

$$y= f(x) \tag{1.1}$$

may be known only at values of x i.e.

$$y_i=y(x_i) \qquad (i=1,2\ldots\ldots n) \tag{1.2}$$

if the function is also known.

Discrete data or tabular data may consist of small sets of smooth data, large sets of smooth data, small sets of rough data or large sets of rough data.

In many engineering and mathematical applications, the values of the given discrete data are not directly required. The values of the function at points other than the known discrete points may be needed (i.e. interpolation points). This process is performed by fitting an approximating function to the set of discrete data and performing the desired process on the approximating function. This approximating polynomial is called an interpolating polynomial. That is, polynomial interpolation is used when the data is exact, the principles here is that, we determine a polynomial $p(x)$, which passes through all the data points and then we take $p(\bar{x})$ as an estimate of $f(\bar{x})$ It is assumed that the discrete points x_o, x_1 ....x_n are ordered so that $x_o<x_1,<$ ....$<x_n$, strictly, the process is known as interpolation only when the point $\bar{x}$ lies in

the interval $[x_o, x_n]$. Interpolation is important in its application and use for the following reasons.

(a) Interpolation is important for functions that are available only in tabular forms (perhaps from the result of an experiment)

(b) Interpolation serves to introduce the wider application of finite differences, see Hosking, et al (1996).

Since we are to interpolate the polynomial then we are going to use numerical methods to interpolate and since some values lie outside the given range, we shall apply the extrapolation approach (method).

1.1 Numerical Methods

Numerical methods are a class of methods for solving a wide variety of mathematics problems. These problems can of course, have their origins as mathematical modes of physical situations. This class of methods is unusual in that, only arithmetic and logic employed, thus the methods can be employed directly on digital computers.

Numerical methods are capable of handling nonlinearities, complex geometries , and large systems of coupled equations which are necessary for the accurate simulation of many real physical situations.

The knowledge of numerical methods enables the computer user to select, modify and program a method suitable for any specific task, aids in the selection and use of prepackaged programs and library subprograms, and makes it possible for the user to communicate with a specialist in an efficient and intelligent way when seeking help for a particularly difficult problem, see Horbeck (1975) and Jacques and Judd (1987).

In this dissertation, we shall be concerned with the problem of constructing a polynomial approximation to a set of discrete data points, these points are Kenya's population for given years, see Kenya's Population Census (1989 and 1999).

1.2 Literature Review

In the present study, the application and use of numerical methods in population calculation and projection has never been used before and therefore no known literature material is available. However, in the ever increasing field of application to engineering and complex organization, mathematics often assumes a character less exact, possibly less aesthetically satisfying, then when it is pursued for its own sake.

A problem may arise in which many variables or 'parameters' are involved; only when attention is confined to the more significant of these in the problem susceptible to known mathematical techniques, and even then the functions concerned are often only manageable when reduced to an approximate form. Thus in view of the foregoing literature, we employ the numerical method in this dissertation to project and calculate population.

On the other hand however, a lot of written and published literature, is available for the statistical method. For example Kenya population calculation and projection for the years 1969, 1979, 1989 and 1999 have been obtained using the statistical method. Other researchers that have utilized this method are Baltazar (1994), Zabba (1979). These researchers have attempted to investigate the effects of AIDS and fertility rates on population calculation and projection. There are many statistical methods that can be employed but over the years Kenya population calculation and projection has been done using the exponential formula given below.

$$p_n = p_o e^{rt} \qquad (1.3)$$

where p_n is the year on which projection is done
p_o is the base year's population
r is the intercensal growth rate

t is the time interval between the base year and the year for which the population is projected, respectively.

1.3 Dissertation Overview

This dissertation is on the use of numerical methods of population calculation and projections.

In chapter one we give a brief mention of the numerical methods and statistical methods to be used. We also mention the areas where numerical methods are often used i.e in engineering, mathematical application and science. The literature review and the dissertation overview are also given in the chapter.

In chapter two, we deal with interpolation methods and the operators used i.e. the forward, backward and central differences operators, denoted by Δ, ∇ and δ, respectively.

Also the equation which arise from these operators are given. Thus the forward, backward and central difference tables are generated. The Stirling and Bessel's formula are given. In chapter three, the discussion is on extrapolation method $a \leq x \leq b$ but values of $f(x)$ are need for say $x < a$ or $x > b$ then an extrapolation is required. Here mention is made on the use of Gregory-Newton forward and backward polynomial interpolation formula in population calculation and projection. Tabulation of values $\ln f(x)$ and $\ln(x)$ and hence the graph of $\ln f(x)$ against $\ln(x)$ is drawn, then the extrapolation of the values from the graph of $\ln f(x)$ against $\ln(x)$ is done.

In chapter four, we employ the three numerical formulae i.e Gregory-Newton forward and backward formula, Stirling formula and Bessel's formula to calculate and project Kenya population for the various years.

In chapter five, Kenya population is projected using 1989 as the base year and its corresponding population as the base population. In chapter six, Kenya population is also calculated and projected using 1999 as the base year and its corresponding population as the base population. In both chapter five and six, the projection is done for years 1990 to 2010.

In chapter seven, comparison is done on the values projected using the two methods, i.e the statistical method and numerical method when 1989 and 1999 are used as the base years. Bar graphs and line graphs are drawn to compare the results obtained using the two methods.

In chapter eight we give a detained discussion of the results when statistical and numerical methods are used to predict and project population. In chapter nine projection per province is done. In chapter ten the effect of mortality from other causes, fertility rates and mortality rates from AIDS are discussed.

CHAPTER 2

INTERPOLATION METHODS

2.0 Introduction

Interpolation is the process whereby untabulated values of a function tabulated only at certain values are estimated, on the assumption that the function behaves sufficiently smoothly between tabular points for it to be approximated by a polynomial of fairly low degree.

The knowledge of numerical methods enables one to select, modify and program a method suitable for any specific task. Interpolation methods here approximate the values for the given function. It must therefore be appreciated that approximation, far from always implying a sacrifice of accuracy, can provide the means whereby mathematics may be brought to bear on practical problems which could otherwise be out of reach. The main objectives of this chapter is to consider new ways of re-writing functions in an approximate form.

2.1 Generation of Difference Tables

In order to solve the problems presented in this dissertation, we shall employ three numerical approaches based on difference tables; i.e the forward difference scheme, a backward difference scheme and a central difference scheme. The operator notations used are Δ, ∇ and δ which denote the forward, backward and central differences, respectively.

Mathematically, the operators Δ, ∇ and δ applied on any function, say $f(x)$ give the following three equations

$$\Delta f(x_i) = f(x_{i+1}) - f(x_i) \tag{2.1}$$

$$\nabla f(x_i) = f(x_i) - f(x_{i-1}) \tag{2.2}$$

$$\delta f(x_i) = f(x_{i+\frac{1}{2}}) - f(x_{i-\frac{1}{2}}), \text{ see Hoffman (1993)} \tag{2.3}$$

Suppose a polynomial $P(s)$ approximates $f(x)$, where $s = \frac{x - x_o}{h}$. Where h is the interval between two consecutive values of x, i.e. the intercensal period. f_o is the initial value then, $f(x)$ can be given in terms of $p(s)$, Δ, and ∇as follows;

$$f(x) \simeq p(s) = f_o + s\Delta f_o + \binom{S}{2} \Delta^2 f_o + \binom{S}{3} \Delta^3 f_o + \text{---} + \binom{S}{n} \Delta^n f_o \tag{2.4}$$

and

$$f(x) \simeq p(s) = f_o + s\nabla f_o + \binom{S+1}{2}\nabla^2 f_o + \binom{S+1}{3}\nabla^3 f_o + \text{---} + \binom{S+n+2}{n}\nabla^n f_o \tag{2.5}$$

Here the notation $f_o \equiv f(x_o)$ is used where x_o is the value of x around which an approximate value lies.

To illustrate the use of these techniques, we consider discrete data showing Kenya's population in years, 1969, 1979, 1989, and 1999, see Kenya's population census 1999 (page xxvii).

For these range of years, the constant interval, h between the two consecutive census is 10, i.e the intecensal period is 10 see table 2.0 The years represented the x row and the population is represented by the function $f(x)$.

year x	***1969***	***1979***	***1989***	***1999***
$f(x)$ population	*10,942,705*	*15,327,061*	*21,448,774*	*28,686,607*

Table 2.0: Kenya's population for the years1969, 1979, 1989, and 1999

See Kenya population census 1999 (page xxvii)

2.1.1 Forward difference table

A forward difference table can be generated by taking forward differences at each point in x. For example, at $x = 1969$ the forward difference is given by

$\Delta f_{69} = f_{79} - f_{69} = 15{,}327{,}061 - 10{,}942{,}705 = 4{,}384{,}356$

where f_{69} and f_{79} have been used to denote the function values in the years 1969 and 1979, respectively, while the operator Δ denotes a forward difference operator. A forward difference table is shown in table 2.1

year *(x)*	*f(x)* population	Δf	$\Delta^2 f$	$\Delta^3 f$
1969	*10,942,705*			
		4,384,356		
1979	*15,327,061*		*1,737,357*	
		6,121,713		*-621,237*
1989	*21,448,774*		*1,116,120*	
		7,237, 833		
1999	*28,686,607*			

Table 2.1: Forward difference for Kenya's population for the years 1969, 1979, 1989 and 1999.

See Kenya population census 1999 (page xxvii)

From table 2.1 we see that the lower and upper halves of the table are not be filled. This is because one entry from both sides is lost in each column every time a new set of differences is taken. From this table, we see that a pyramid which is triangular like, in shape is generated.

2.1.2 Backward difference table

In a similar manner, a backward difference table can be generated by taking backward differences at each point in x. For example, for 1999 we have $\nabla f_{99} = f_{99} - f_{89} =$ 28,686,607 – 21,448,774 = 7,237,833 where f_{89} and f_{99} represent Kenya's population in the years 1989 and 1999, respectively, and ∇denotes a backward difference operator. Thus the backward difference table generated is shown in table 2.2.

year *(x)*	*f(x)* population	∇f	$\nabla^2 f$	$\nabla^3 f$
1969	*10,942,705*			
		4,384,356		
1979	*15,327,061*		*1,737,357*	
		6,121,713		*- 621,237*
1989	*21,448,774*		*1,116,120*	
		7,237,833		
1999	*28,686,607*			

Table 2.2: Backward differences for Kenya's population for the years 1969, 1979, 1989 and 1999.

See Kenya population census 1999 (page xxvii)

Again we note that both lower and upper halves of the table have not been filled. The entries in table 2.2 are the same as those in table 2.1, i.e. the x and *f(x)* values are unchanged.

However the resulting entries obtained when applying the forward Δ operator are different from those obtained when using a backward difference operator ∇.

2.1.3 Central difference table

A central difference table can be generated for Kenya population for the years 1969, 1979, 1989 and 1999 a central difference operator δ is used. The table 2.3 below shows the generated central differences.

year *(x)*	*f(x)* population	δf	$\delta^2 f$	$\delta^3 f$
1969	*10,942,705*			
		4,384,356		
1979	*15,327,061*		*1,737,357*	
		6,121,713		*- 621,237*
1989	*21,448,774*		*1,116,120*	
		7,237,833		
1999	*28,686,607*			

Table 2.3: A central differences for Kenya's population for the years 1969,1979,1989 and 1999.

See Kenya population census 1999 (page xxvii)

We now fill in the gaps in table 2.3 by taking the arithmetic mean of the values above and below each gap. These results are in table 2.4 below.

year *(x)*	*f(x)* population	δf	$\delta^2 f$	$\delta^3 f$
1969	*10,942,705*			
1974	*13,134,883*	*4,384,356*		
1979	*15,327,061*	*5,253,034.5*	*1,737,357*	
1984	*18,387,918*	*6,121,713*	*1,426,738.5*	*- 621,237*
1989	*21,448,774*	*6,679 773*	*1,116,120*	
1994	*25,067,691*	*7,237,833*		
1999	*28,686 607*			

Table 2.4: A central difference of Kenya's population with arithmetic means of the differences for the years 1969, 1979, 1989 and 1999.

See Kenya population census 1999 (page xxvii)

It is worth noting that the different tables 2.2, 2.3 and 2.4 show that a polynomial would fit the data points fairly well since successively higher differences become smaller in magnitude.

2.2 Interpolation with Forward and Backward Differences

The Taylor series expansion about the point x =1969 is given as

$$f(x)= f(1969)+xf^{/}(1969)+\frac{x^2}{2!}f^{//}(1969)+ \frac{x^3}{3!}f^{///}(1969)+\text{------------} \quad (2.6)$$

In this equation the values of the derivatives $f^{/}$, $f^{//}$ e.t.c. are known but finite difference expressions can be used to obtain $f^{/}(1969)$, $f^{//}(1969)$ e.t.c., see Smith (1996).

Recall that the numerical differentiation of a polynomial is given by

$$P_{(s)} \simeq f(x) \simeq \frac{1}{h}\left[\Delta fo+\frac{(2s-1)}{2}\Delta^2 f_o+\text{------}\right] \quad (2.7)$$

where h, s, Δ and f_o take their usual meanings see section 2.1 above.

From equation 2.6, suppose $f_o = f(1969)$, then we have equation 2.7 taking the form

$$f^{/}(1969)=\frac{\Delta f(1989)}{h}- \frac{h}{2}f^{//}(1969) + 6(h^2) \quad (2.8)$$

similarly, expressions for $f^{//}$ (1969), $f^{///}$ (1969), e.t.c can be obtained.

Substituting these expressions into equation 2.8 we obtain.

$$f(x)= f(1969)+\frac{x}{h}\Delta f(1969)+\frac{(x)(x-h)}{2!h^2}\Delta^2 f(1969)+\frac{x(x-h)(x-2h)}{3!h^3}\Delta^3 f(1969) \quad (2.9)$$

The remaining terms i.e terms with $f^{//}$ (1969) and higher derivatives can be obtained by induction. This formula is called the Gregory-Newton forward interpolation formula. The differences are of course obtained from the forward difference table. In the numerical differentiation using backward differences can also be performed and this leads to the backward Gregory–Newton formula which is written as

$$f(1969)=f(1969)+\frac{x}{h}\nabla f(1969)+\frac{(x)(x+h)}{2!h^2}\nabla^2 f(1969)+\frac{(x)(x+h)(x+2h)}{3!h^3}\nabla^3 f(1969) \quad (2.10)$$

generally equation 2.9 and 2.10 can be written as

$$f(x)= f_o+ \Delta f_o+\frac{x(x-1)}{2!}\Delta^2 f_o+\frac{x(x-1)(x-2)}{3!}\Delta^3 f_o \quad (2.11)$$

and

$$f(x)=f_o+x\nabla f_o+\frac{x(x+1)}{2!}\nabla^2 f_o+\frac{x(x+1)(x+2)}{3!}\nabla^3 f_o \quad (2.12)$$

Which are the Gregory – Newton forward and backward formulae

Now replacing x by s and f_o by y_o we have

$$f(s)=y_o+s\Delta y_o+\frac{s(s-1)}{2!}\Delta^2 y_o+\frac{s(s-1)(x-2)}{3!}\Delta^3 y_o \text{ and} \quad (2.13)$$

$$f(s)=y_o+s\nabla y_o+\frac{s(s+1)}{2!}\nabla^2 y_o+\frac{s(s+1)(x+2)}{3!}\nabla^3 y_o \quad (2.14)$$

where

$$s= \frac{x-x_o}{h}$$

and

h= interval between the evenly spaced x values. In our case h is 10 i.e interval between which census was carried out i.e intercensal period. We are now ready to consider methods of using a tabulated function and the difference tables to find values of the function between the tabulated points. Although the function is known only at the discrete tabulated points, we begin by assuming that the function is analytic over the entire range of values of interest. This assumption will help us to proceed and incase reasonable values are to be obtained of the function between the tabulated values then the actual function represented by the discrete points should be reasonably smooth and well behaved.

If function is analytic then it is possible to find the values of $f(x)$ at any point between the tabulated points by using Taylor series expansion of $f(x)$ about one of the tabulated points. We arbitrarily choose this point to be *x=1969* as given in equation 2.13.

Plotting *h* with respect to *f(x)* the function values gives

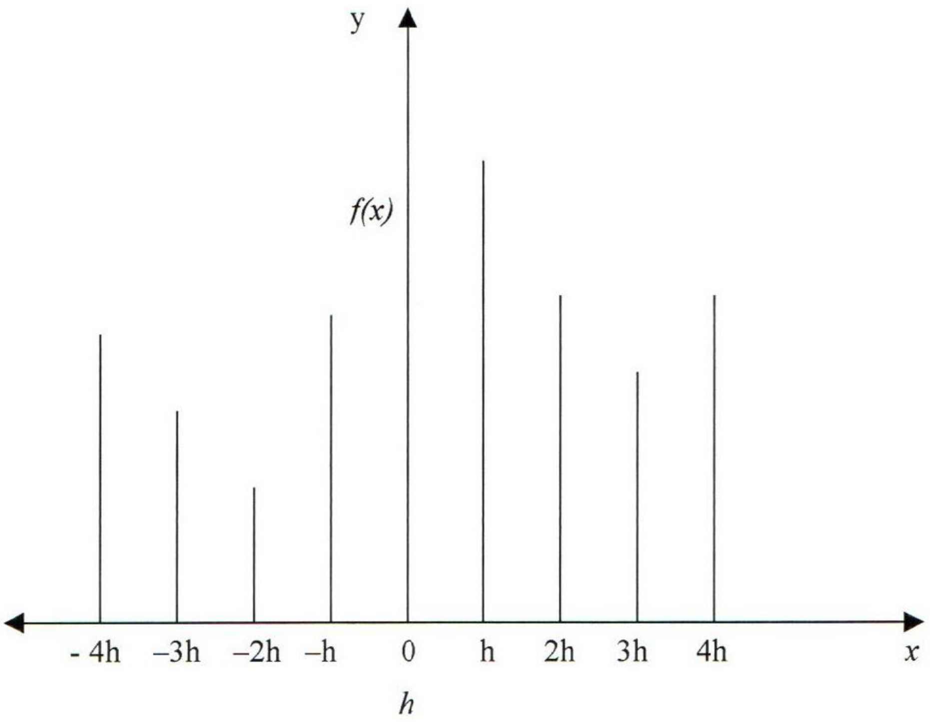

Figure 1.0: Plotting *h* with respect to *f(x)* for evenly spaced function

Note that this function is evenly spaced. We can now use the formula in equations 2.13 and 2.14 together with the difference tables discussed in section 2.0 to interpolate for intermediate values of $f(x)$. We use forward differences if interpolating value near the top of the table i.e. when x_I is close to x_o and backward difference if interpolating value is near the bottom i.e. when x_I is close to x_n. Here x_o, x_I and x_n are the initial, interpolating and final values respectively. Interpolation near the centre of the table is best achieved with central difference.

2.3 Interpolation with Central Differences

Interpolation near the centre of a set of evenly spaced tabulated values is best accomplished by using central differences. A central difference table is first generated as discussed in section 2.13 Then an appropriate interpolation formula is chosen. In this dissertation we employ two most commonly used interpolating formulae, namely.

(i) Stirling formula

This formula is written as

$$y_s = y_o + s(\delta y_o) + \frac{s^2}{2!}(\delta^2 y_o) + \frac{s(s-1)}{3!}\delta^3 y_o + \frac{s^2(s^2-1)}{4!}((\delta^2 y_o) \qquad (2.15)$$

Here the full lines are used as base

(ii) Bessel's formula

$$y_s = y_o + s(\delta y_o) + \frac{(s^2 - {}^1\!/_4)}{2!}\delta^2 y_o + \frac{s(s^2 - {}^1\!/_4)}{3!}\delta^3 y_o + \frac{(s^2 - {}^1\!/_4)(s^2 - {}^9\!/_4)}{4!}\delta^4 y_o \qquad (2.16)$$

Here half lines are used as base

In order to use this formulae the origin of s must be shifted to the base line. Since we have a choice as to whether the full lines or the half-lines of the difference table will be used as the base line s, should not be greater than $\pm$ 0.25. This value result in rapid "convergence" of the interpolation formula. This means that only a few terms are required to obtain a value which lies on the highest order interpolating polynomial available from the base line entries.

As with the forward and backward differences tables the empty regions of the central difference table can be filled out in order to provide additional entries in the base line, but it is often not necessary due to the rapid "convergence" of the central difference interpolation formula.

2.4 Conclusion

In the present investigations, we have no knowledge of the function *f(x)* to be used to generate values of *f(x)*. Therefore we cannot say what the values of Kenya's population was in the years 1974, 1984 and 1994 and other years or the error involved in the interpolated value.

Even in the well known function *f(x)*, we can only estimate the value of the function *f(x)* between the tabulated points based on numerical interpolating formulae and we shall get a reasonably accurate value. If the higher order difference of the tabulated function become small, then polynomial interpolation will be quite accurate.

In general, taking many terms in the Gregory –Newton interpolation formula increases the accuracy of the interpolated value.

CHAPTER 3

EXTRAPOLATION METHODS

3.0 Introduction

Sometimes an information or data may be required outside the interval under consideration. In this case interpolation technique will not be used to project or predict these values. Now we shall use extrapolation method to project the population for points outside the interval. In this case the points are the years after the census year.

3.1 Extrapolation

If a function $f(x)$ is known only on the interval say $a \leq x \leq b$ but values of $f(x)$ are needed for say $x < a$ or $x > b$, then an extrapolation method is required. Here we use this method to predict (estimate) values of $f(x)$ outside the given limits.

However, extrapolation method has some weaknesses. Unlike interpolation where the function is firmly anchored on both sides of the point where a value is to be obtained, extrapolation method requires that a function is fixed only on one side of the interval and is relatively free to move to the other side.

If the function is known at discrete evenly spaced points, then the Gregory – Newton forward and backward polynomial interpolation formulae are commonly employed for extrapolation with the last known point used as the base line. Note that the choice of either a forward or backward formula will of depend on whether $x > b$ *or* $x < a$.

In order to obtain meaningful results using extrapolation, it is important that the function be well defined and therefore suited to polynomial interpolation that is the higher order differences in the difference table must approach zero. Otherwise, we resort plotting the graph of $lnf(x)$ against $ln(x)$, for if a function to be extrapolated cannot be well approximated by a polynomial, then a useful approach is to plot $f(x)$ verses x on log – log graph paper, this reduces large variety of functions to straight lines or to smooth curves which are easy to extrapolate. The numerical equivalent of this graphical procedure is to tabulate $lnf(x)$ and $ln(x)$ and then plot $lnf(x)$ against $ln(x)$ and then carry out polynomial extrapolation. The main objective of using these graphs is to find the relationship between the function $f(x)$ and x, see Hornbeck (1975).

And since we are taking natural logarithms, it means that, the function in use is an exponential one and may be expressed in the form

$$p_n = p_o e^{rt}$$

where p_n is the year on which projection is done (3.0)
p_o is the base year's population
r is the intercensal growth rate
t is the time interval between the base year and the year for which the population is projected, respectively.

This equation gives rise to the statistical method, which is usually employed to calculate and project a country's population, e.g Kenya's population presented in this dissertation.

3.2 Projection of Population from Extrapolation Graph

year *(x)*	***ln(x)***	Population ***f(x)***	***lnf(x)***
1969	*7.585*	*10,942,705*	*16.208*
1979	*7.590*	*15,327,061*	*16.545*
1989	*7.595*	*21,448,774*	*16.881*
1999	*7.600*	*28,686,607*	*17.172*

Table 3.0: Generation of *ln(x)* and *lnf(x)* from given years and population, respectively.

Information in table 3.0, can be used to find the relation of *lnf(x)* and *ln(x)* and this is seen in figure 3.0.

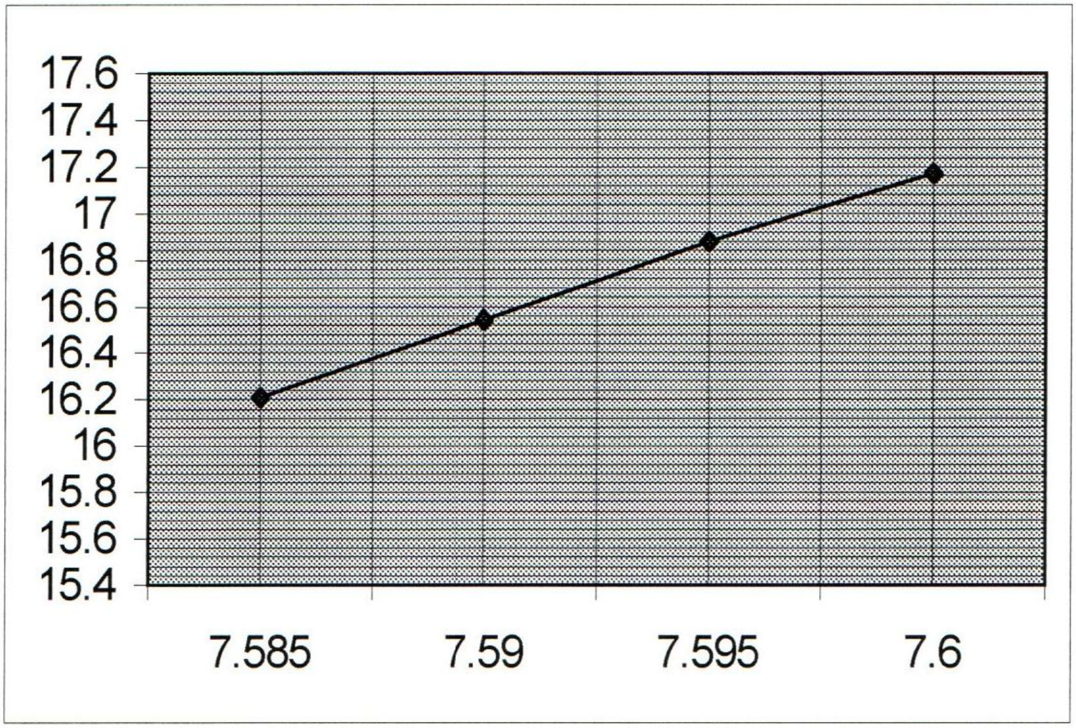

Figure 3.0: Plot of *lnf(x)* and *ln(x)*

It can be observed from this figure that, if the line is extended or produced carefully and accurately, then the population for the given years can be projected or predicted and conversely, the years for projected population can be read from the graph, thus we see that the extrapolation technique plays a crucial role in predicting/projecting population.

3.3 Extrapolation for the years 2000 to 2010

A table is generated for the values of *x* in the range *x*=2000 to *x*=2010 and corresponding values of *ln(x)* are calculated. Figure 3.1 can be used to obtain corresponding values of *lnf(x)* for the years 2000 to 2010, respectively, by just extrapolating (extending) the line outside the interval given. These values are shown in table 3.2.

year ***(x)***	***ln(x)***	Corresponding value of ***lnf(x)***	Extrapolated population
2000	*7.6009*	*17.18*	*28,918,729*
2001	*7.6010*	*17.20*	*29,502,925*
2002	*7.6020*	*17.25*	*31,015,573*
2003	*7.6024*	*17.35*	*34,277,509*
2004	*7.6029*	*17.38*	*35,321,415*
2005	*7.6034*	*17.40*	*36,034,955*
2006	*7.6039*	*17.45*	*37,882,506*
2007	*7.6043*	*17.46*	*38,263,232*
2008	*7.6049*	*17.48*	*39,036,200*
2009	*7.6054*	*17.49*	*39,428,521*
2010	*7.6059*	*17.50*	*39,824,784*

Table 3.1: Extrapolation for the years 2000 - 2010

3.4 Conclusion

The extrapolation method can be used to project and predict the population for example natural logarithm for the year 2000 is 7.6010 see figure 3.0 and its corresponding natural logarithm of *f(x)* is 17.20. This gives population of year 2000 as 29,502,925. Thus the same procedure can be used to project or predict the population or the years for the given population. The values obtained are closer to the values obtained when statistical and numerical methods are used, see projections for the years 2000 to 2010 when both 1989 and 1999 are used as the base years and corresponding populations as the base populations, respectively. Therefore this method is reliable and accurate in predicting population.

CHAPTER 4

APPROXIMATION OF KENYA'S POPULATION USING THE INTERPOLATION METHOD

4.0 Introduction

Here we apply the numerical method to project the population for the given years, i.e. the years half way the intercensal period. Intecensal period is the period between the two successive census years i.e. $1999 - 1989 = 10$

4.1 Interpolation Methods

Using the data of table 2.0, the population for the years other than 1969, 1979, 1989 and 1999 can be obtained. Suppose we consider the years, say (a)1974 (b) 1984 (c) 1994.

Then we can use numerical techniques to project their corresponding values, here we maintain the constant interval to be 10 . The interpolation techniques used are

(i) Gregory – Newton forward and backward formula

(ii) Stirling formula

(iii) Bessel's formula

4.2 Kenya's Population in 1974

4.2.1 Gregory-Newton Forward formula

Since 1974 is closer to 1969, we will use the Gregory-Newton forward formula, i.e.

$$f(x) \simeq P_{(s)} = y_o + s\Delta y_o + s(s-1)\frac{\Delta^2 y_o}{2!} + \frac{s(s-1)(s-2)\,\Delta^3 y_o}{3!} \quad (4.0)$$

For derivation see Hornback (page 39)

Using table 2.0, then x= 1974, x_o=1969 and here s= $\frac{x-x_o}{h}$,

where h=10 and thus

$$s = \frac{1974 - 1969}{10} = \frac{5}{10} = 0.5$$

Therefore

$$P_{(1974)} = 10{,}942{,}705 + (.5)\,(4{,}384{,}356) + \frac{(.5)\,(-.5)\,(1{,}737{,}357)}{2!} + \frac{(.5)(-.5)(-1.5)\,(-621{,}237)}{3!}$$

$$10{,}942{,}705 + 2{,}192{,}178 - 217{,}169.63 - 38{,}827.313 = 12{,}878{,}886$$

The Gregory-Newton formula approximates Kenya's population in 1974 to be

12,878,886.

4.3 Kenya's Population in 1984

4.3.1 Stirling formula

Since 1984 is centrally placed in table 2.4, it suggest that we use the central difference, and apply Stirling formula and Bessel's formula. The second formula namely the Stirling formula is given by

$$f(x) \simeq P_{(s)} = y_o + s\delta y_o + \frac{s^2\ \delta^2 y_o}{2!} + \frac{s(s-1)\delta^3 y_o}{3!} \quad (4.1)$$

see Hornbeck (1975)

For x = *1984*, x_o=*1979*, and here $s = \frac{x-x_o}{h}$ where $h=10$ and thus $s = \frac{1984-1979}{10} = 0.5$

Thus equation 4.1 now becomes

$$P(1984) = 15{,}327{,}061 + \frac{0.25}{2!}\,(5{,}253{,}034.5) + \frac{.25}{3!}(1{,}737{,}357) = 18{,}170748$$

This formula approximates Kenya's population in 1984 to be 18,170748.

4.3.2 Bessel's formula

The Bessel's formula is stated as

$$y_p = y_o + s\delta y_o + \frac{(s^2 - \frac{1}{4})}{2!}\delta^2 y_o + \frac{s(s^2 - \frac{1}{4})}{3!}\delta^3 y_o \qquad (4.2)$$

For $x = 1984$, $x_o = 1979$ and $h = 10$, hence $s = \frac{x - x_o}{h}$, $s = \frac{1984 - 1979}{10} = \frac{5}{10} = 0.5$

and y_o = the initial value i.e 1979 population.

Thus we have

$$Y_{(1984)} = 15,327,061 + (0.5)\ 6,121,713 + (0) + (0) = 18,387,918$$

We see that the Bessel's formula gives Kenya's population for 1984 to be 18,387,918 and Stirling formula gave 18,170,748 which indicates that the two formulae give reasonably good results though numerical errors exist.

4.4 Kenya's Population in 1994

4.4.1 Gregory – Newton backward formula

Since the year 1994 is closer to the last census year, i.e. 1999, see table 2.4, then it suggests the application of the Gregory-Newton backward formula. This formula is given by

$$y_p = y_n + s\nabla y_n + \frac{s(s+1)}{2!}\nabla^2 y_n + \frac{s(s+1)(s+2)}{3!}\nabla^3 y_n \qquad (4.3)$$

Here $x = 1994$, $x_n = 1999$, $h = 10$

hence $s = \frac{x - x_o}{h}$,

$y_n = 28,686,607$

therefore $s = \frac{x - x_n}{h} = \frac{1994 - 1999}{10} = \frac{-0.5}{10} = -0.5$

thus

$$P_{(1994)} = 28{,}828{,}607 + (-0.5)(7{,}237{,}833) + \frac{(-0.5)(0.5)(1\ 116{,}120)}{2!} + \frac{(-0.5)(0.5)(1.5)(-621{,}237)}{3!}$$

$$28{,}686{,}607 - 3{,}618{,}916.5 - 139{,}515 + 38{,}827.313 = 24{,}967{,}003$$

Thus the Gregory-Newton backward formula approximates the population to be 24,967,003. Now we proceed to compare the values we have calculated using the above formulae and the values we obtained in table 2.4 as the result of calculating the arithmetic mean of the values above and below each gap. The comparison is as in table 4.0 below.

year	*The calculated value A*	*Arithmetic mean B*	*Difference B-A*
1974	*12,878,886*	*13,134,883*	*255,997*
1984	*18,387,918*	*18,387,918*	*0*
1994	*24,967,003*	*25,067,691*	*100,688*

Table 4.0: comparison of calculated value and arithmetic mean values

Now we express the differences as the percentage of the arithmetic mean value as shown below

For the year 1974 we have $\frac{255997}{13{,}134{,}883} \times 100 = 1.9\ \%$

For the year 1984 we have $\frac{0}{18{,}387{,}918} \times 100 = 0\%$

For 1994 we have $\frac{100{,}688}{25{,}067{,}691} \times 100 = 0.4\%$

Analysis of this percentages indicate that the differences are small or negligible thus Gregory-Newton backward formula and Bessel's formula are more accurate since the percentages are 0% and 0.4% as seen above. We can conclude from this observation that if we keep on calculating arithmetic means for the years with intercensal period of ten years or any other intercensal period we shall be getting Kenya's population for the averaged years as the arithmetic means.

Since this involves the generation of a central difference table, and arithmetic means obtained from the values above and below each gap it supports the use of interpolation methods in population calculation and projection.

CHAPTER 5

PROJECTION OF KENYA'S POPULATION FOR THE YEARS 1990-2010 USING NUMERICAL METHOD

5.0 Introduction

Projection is done using the years 1989 and 1999 as the base years and their corresponding population values as the base population and numerical methods are employed.

Here we want to project the population for the calendar years 1990 to 2010. We project these populations using the 1989 census, enumerated population as the base year and also using 1999 census, enumerated population as the base year populations. Then we shall compare the results with the projections obtained when using the statistical methods.

5.1 Projection Using 1989 as Base Year

x_n is fixed through in this section to be 1989 and h is fixed at 10

Then $P(x)=y_n+s\nabla y_n+\frac{s(s+1)}{2!}\nabla^2 y_n+\frac{s(s+1)(s+2)}{3!}\nabla^3 y_n$ (5.0)

here x=1990, x_n=1989, h=10, y_n = 21,448,774 and thus

$$s=\frac{x-x_n}{h}=\frac{1990-1989}{10}=\frac{1}{10}=0.1$$

therefore

$$P(1990) = 21,448,774+(0.1)(6,121,713)+(0.1)(1.1)(1,737,357)$$

$$= 21,44,8,774+612,171.3+95,554.635 = 22,156,500$$

for $x = 1991$, we have $s = \frac{x-x_n}{h} = \frac{1991-1989}{10} = 0.2$

$$P(1991) = 21,448,774+(0.2)(6,121,713)+(0.2)\frac{(1.2)(1,737,357)}{2!}$$

$$=21,448,774+1,224,342.6+208,482.84$$

$$= 22,881,599$$

for $x = 1992$, we have $s = \frac{x-x_n}{h} = \frac{1992-1989}{10} = 0.3$

$$P(1992) = 21448774+(0.3)(6121713+(0.3)\frac{(1.3)(1737357)}{2!}$$

$$= 21,448,774+1,836,513.9+338,784.62$$

$$= 23,624,073$$

for $x = 1993$, we have $s = \frac{x-x_n}{h} = \frac{1993-1989}{10} = 0.4$

$$P(1993) = 21,448,774+(0.4)(6,121,713+(0.4)\frac{(1.4)(1,737,357)}{2!}$$

$$= 21,448,774+2,448,685.2+486,459.96$$

$$= 24,383,919$$

for $x = 1994$, we have $s = \frac{x-x_n}{h} = \frac{1994-1989}{10} = 0.5$

$$P(1994) = 21,448,774+(0.5)(6,121,713)+\frac{(0.5)(1.5)(1,737,357)}{2!}$$

$$= 21,448,774+3,060,856.5+651,508.88$$

$$= 25,161,139$$

for $x = 1995$, we have $s = \frac{x-x_n}{h} = \frac{1995-1989}{10} = 0.6$

$$P(1995) = 21,448,774+(0.6)(6,121,713)+(0.6)\frac{(1.6)(1,737,357)}{2!}$$

$$= 2,1448,774+3,673,027.8+333,931.36$$

$$= 25,955,733$$

for $x = 1996$, we have $s = \frac{x-x_n}{h} = \frac{1996-1989}{10} = 0.7$

$P(1996) \quad = 21,448,774+(0.7)(6,121,713)+(0.7)\frac{(1.7)(1,737,357)}{2!}$

$= 21,448,774+4,285,199.1+1033,727.4$

$= 26,767,701$

for $x = 1997$, we have $s = \frac{x-x_n}{h} = \frac{1997-1989}{10} = 0.8$

$P(1997) \quad = 21,448,774+(0.8)(6,121,713)+(0.8)\frac{(1.8)(1,737,357)}{2!}$

$= 21,448,774+4,897,370.4+1,250,897$

$= 27, 597, 041$

for $x = 1998$, we have $s = \frac{x-x_n}{h} = \frac{1998-1989}{10} = 0.9$

$P(1998) \quad = 21,448,774+(0.9)(6,121,713)+(0.9)\frac{(1.9)(1,737,357)}{2!}$

$= 21,448,774+5, 509, 541.7+1,485,440.2$

$= 28, 443, 756$

for $x = 1999$, we have $s = \frac{x-x_n}{h} = \frac{1999-1989}{10} = 1.0$

$P(1999) \quad = 21,448,774+(1)(6,121,713)+(1)\frac{(2)(1,737,357)}{2!}$

$= 21,448,774+612,1713+1,737,357$

$= 29,307, 844$

for $x = 2000$, we have $s = \frac{x-x_n}{h} = \frac{2000-1989}{10} = 1.1$

$P(2000) \quad = 21,448,774+(1.1)(6,121,713)+(1.1)\frac{(2.1)(1,737,357)}{2!}$

$= 21,448,774+6,733,888.3+2,006,647.3$

$= 30,189,306$

for $x = 2001$, we have $s = \frac{x-x_n}{h} = \frac{2001-1989}{10} = 1.2$

$P(2001) \quad =21,448,774+(1.2)(6,121,713)+\frac{(1.2)(2.2)(1,737,357)}{2!}$

$= 21,448,774+73,460,55.6+229,334.2$

$= 31,088, 141$

for $x = 2002$, we have $s = \frac{x-x_n}{h} = \frac{2002-1989}{10} = 1.3$

$$P(2002) = 21{,}448{,}774 + (1.3)(6121713) + \frac{(1.3)(2.3)(1{,}737{,}357)}{2!}$$

$$= 21{,}448{,}774 + 7{,}958{,}226.9 + 259{,}348.7$$

$$= 32{,}004{,}350$$

for $x = 2003$, we have $s = \frac{x-x_n}{h} = \frac{2003-1989}{10} = 1.4$

$$P(2003) = 21{,}448{,}774 + (1.4)(6{,}121{,}713) + \frac{(1.4)(2.4)(1{,}737{,}357)}{2!}$$

$$= 21{,}448{,}774 + 8{,}570{,}398.2 + 2{,}918{,}759.8$$

$$= 32{,}937{,}932$$

for $x = 2004$, we have $s = \frac{x-x_n}{h} = \frac{2004-1989}{10} = 1.5$

$$P(2004) = 21{,}448{,}774 + (1.5)(6{,}121{,}713) + \frac{(1.5)(2.5)(1{,}737{,}357)}{2!}$$

$$= 21{,}448{,}774 + 9{,}182{,}569.5 + 3{,}257{,}544.4$$

$$= 33{,}888{,}888$$

for $x = 2005$, we have $s = \frac{x-x_n}{h} = \frac{2005-1989}{10} = 1.6$

$$P(2005) = 21{,}448{,}774 + (1.6)(6{,}121{,}713) + \frac{(1.6)(2.6)(1{,}737{,}357)}{2!}$$

$$= 21{,}448{,}774 + 9{,}794{,}740.8 + 3{,}613{,}702.6$$

$$= 34{,}857{,}217$$

for $x = 2006$, we have $s = \frac{x-x_n}{h} = \frac{2006-1989}{10} = 1.7$

$$P(2006) = 21{,}448{,}774 + (1.7)(6{,}121{,}713) + \frac{(1.7)(2.7)(1{,}737{,}357)}{2!}$$

$$= 21{,}448{,}774 + 10{,}406{,}912 + 39{,}872{,}34.3$$

$$= 35{,}842{,}920$$

for $x = 2007$, we have $s = \frac{x-x_n}{h} = \frac{2007-1989}{10} = 1.8$

$$P(2007) = 21{,}448{,}774+(1.8)(6121713)+\frac{(1.8)(2.8)(17{,}37{,}357)}{2!}$$

$$= 21{,}448{,}774+11{,}090{,}83+4{,}378{,}139.6$$

$$= 36{,}845{,}\ 997$$

for $x = 2008$, we have $s = \frac{x-x_n}{h} = \frac{2008-1989}{10} = 1.9$

$$P(2008) = 2{,}1448{,}774+(1.9)(6{,}121{,}713)+\frac{(1.9)(2.9)(1{,}737{,}357)}{2!}$$

$$= 21{,}448{,}774+11{,}631{,}255+478{,}418.5$$

$$= 37{,}\ 866{,}448$$

for $x = 2009$, we have $s = \frac{x-x_n}{h} = \frac{2009-1989}{10} = 2.0$

$$P(2009) = 21{,}448{,}774+(2)(6{,}121{,}713)+\frac{(2)(3)(1{,}737{,}357)}{2!}$$

$$= 21{,}448{,}774+12{,}243{,}426+5{,}212{,}071$$

$$= 38{,}904{,}271$$

for $x = 2010$, we have $s = \frac{x-x_n}{h} = \frac{2010-1989}{10} = 2.1$

$$P(2010) = 21{,}448{,}774+(2.1)(6{,}121{,}713)+\frac{(2.1)(3.1)(1{,}737{,}357)}{2!}$$

$$= 21{,}448{,}774+12{,}855{,}597+5{,}655{,}099$$

$$= 39{,}959{,}468$$

5.2 Projection Using 1999 as Base Year

For this case x_o is fixed at 1999 and h is fixed at *10*

y_o = 28,686,607 (enumerated), see table 2.1. The formula to be used is given as

$$P_{(x)} = y_o + s\nabla y_o + \frac{s(s+1)}{2!}\nabla^2 y_n + \frac{s(s+1)(s+2)}{3!}\nabla^3 y_n \qquad (5.1)$$

where

$x_o = 1999$, $x = 2000$, $h = 10$ and $y_o = 28{,}686{,}607$

$$s = \frac{x-x_o}{h} = \frac{1}{10} = 0.1$$

for $x = 2000$, we have $s = \frac{x-x_n}{h} = \frac{2000-1999}{10} = 0.1$

$$P(2000) = 28,686,607+(0.1)(7,237,833)+(0.1)(1.1)\frac{(1,116,120)}{2!}+(0.1)\frac{(1.1)(2.1)(-621,237)}{3!}$$

$$= 28,686,607+23,783.3+61,386.6-23,917.625$$

$$= 29,447,859$$

for $x = 2001$, we have $s = \frac{x-x_n}{h} = \frac{2001-1999}{10} = 0.2$

$$P(2001) = 28,686,607+(0.2)(7,237,833)+(0.2)(1.2)\frac{(1,116,120)}{2!}+(0.1)\frac{(1.2)(2.2)(-621,237)}{3!}$$

$$= 28,686,607+1,447,566.6+133,934.4-54,668.856$$

$$= 30,213,439$$

for $x = 2002$, we have $s = \frac{x-x_n}{h} = \frac{2002-1999}{10} = 0.3$

$$P(2002) = 28,686,607+(0.3)(7,237,833)+(0.3)(1.3)\frac{(1,116,120)}{2!}+(0.3)\frac{(1.3)(2.3)(-621,237)}{3!}$$

$$= 28,686,607+2,171,349.9+217,643.4-92,874.932$$

$$= 30,982,725$$

for $x = 2003$, we have $s = \frac{x-x_n}{h} = \frac{2003-1999}{10} = 0.4$

$$P(2003) = 28,686,607+(0.4)(7,237,833)+(0.4)(1.4)\frac{(1,116,120)}{2!}+(0.4)\frac{(1.4)(2.4)(-621,237)}{3!}$$

$$= 28,686,607+2,895,133.2+312,513.6-139,157.09$$

$$= 31,755,079$$

for $x = 2004$, we have $s = \frac{x-x_n}{h} = \frac{2004-1999}{10} = 0.5$

$$P(2004) = 28,686,607+(0.5)(7,237,833)+(0.5)(1.5)\frac{(1,116,120)}{2!}+(0.5)\frac{(1.5)(2.5)(-621,237)}{3!}$$

$$= 28,686,607+3,618,916.5+418,545-194,136.56$$

$$= 32,529,932$$

for $x = 2005$, we have $s = \frac{x-x_n}{h} = \frac{2005-1999}{10} = 0.6$

$$P(2005) = 28,686,607 + (0.6)(7,237,833) + (0.6)(1.6)\frac{(1,116,120)}{2!} + (0.6)\frac{(1.6)(2.6)(-621,237)}{3!}$$

$$= 28,686,607 + 4,342,699.8 + 535,737.6 - 25,843,459$$

$$= 33,306,610$$

for $x = 2006$, we have $s = \frac{x - x_n}{h} = \frac{2006 - 1999}{10} = 0.7$

$$P(2006) = 28,686,607 + (0.7)(7,237,833) + (0.7)(1.7)\frac{(1,116,120)}{2!} + (0.7)\frac{(1.7)(2.7)(-621,237)}{3!}$$

$$= 28,686,607 + 5,066,483.1 + 664,091.4 - 232,572.41$$

$$= 34,084,509$$

for $x = 2007$, we have $s = \frac{x - x_n}{h} = \frac{2007 - 1999}{10} = 0.8$

$$P(2007) = 28,686,607 + (0.8)(7,237,833) + (0.8)(1.8)\frac{(1,116,120)}{2!} + (0.8)\frac{(1.8)(2.8)(-621,237)}{3!}$$

$$= 28,686,607 + 5,790,266.4 + 803,606.4 - 417,471.26$$

$$= 34,863,029$$

for $x = 2008$, we have $s = \frac{x - x_n}{h} = \frac{2008 - 1999}{10} = 0.9$

$$P(2008) = 28,686,607 + (0.9)(7,237,833) + (0.9)(1.9)\frac{(1,116,120)}{2!} + (0.9)\frac{(1.9)(2.9)(-621,237)}{3!}$$

$$= 28,686,607 + 6,514,049.7 + 954,282.6 - 513,452.38$$

$$= 35,641,487$$

for $x = 2009$, we have $s = \frac{x - x_n}{h} = \frac{2009 - 1999}{10} = 1.0$

$$P(2009) = 28,686,607 + (1)(7,237,833) + (1)(2)\frac{(1,116,120)}{2!} + (1)\frac{(2)(3)(-621,237)}{3!}$$

$$= 28,686,607 + 7,237,833 + 1,116,120 - 621,237$$

$$= 36,419,323$$

for $x = 2010$, we have $s = \frac{x - x_n}{h} = \frac{2010 - 1999}{10} = 1.1$

$$P(2010) = 28,686,607 + (1.1)(7,237,833) + (1.1)(2.1)\frac{(1,116,120)}{2!} + (1.1)\frac{(2.1)(3.1)(-621237)}{3!}$$

$$= 28,686,607 + 7,961,616.3 + 1,289,118.6 - 741,446.36$$

$$= 37,195,896$$

5.3 Calculation of Population for the Years 1990 to 1998

Calculation of population using both the Gregory-Newton forward and backward difference tables. Since 1990-1993 are near the top of the table, we employ the methods

i) Gregory-Newton forward difference

ii) Gregory-Newton backward difference

5.3.1 Gregory-Newton forward difference

x_o is fixed at 1989 and h is fixed at 10.
Thus we have $s = \frac{x-x_o}{h}$

for $x = 1990$, $x_o = 1989$, $h=10$, $s = \frac{x-x_o}{h} = \frac{1990-1989}{10} = \frac{1}{10} = 0.1$,

$y_o = 21,448,774$

Then using table 2.1

$$P(x) = y_o + s\Delta y_o + \frac{s(s-1)}{2!}\Delta^2 y_n + \frac{s(s-1)(s-2)}{3!}\Delta^3 y_o \quad (5.2)$$

$$P(1990) = 21,448,774+(0.1)(7,237,833)+(0)+(0)$$

$$= 21,448,774+723,783.3$$

$$= 22,172,557$$

for $x = 1991$

$$s = \frac{x-x_o}{h} = \frac{1991-1989}{10} = 0.2$$

$$P(1991) = 21,448,774+(0.2)(7,237,833)$$

$$= 21,448,774+1,447,566.6$$

$$= 22,896,341$$

for x= *1992*

$$s = \frac{x-x_o}{h} = \frac{1992-1989}{10} = 0.3$$

$$P(1992) = 21,448,774+(0.3)(72,237,833)$$

$$= 21,448,774+2,171,349.9$$

$$= 23,620,124$$

for x= *1993*

$$s = \frac{x-x_o}{h} = \frac{1993-1989}{10} = 0.4$$

$$P(1993) = 21,448,774+(0.4)(72,237,833)$$

$$= 2,1448,774+2,895,133.2$$

$$= 24,343,907$$

P(1994) was calculated using the central different table 2.4 as the arithmetic mean and the value was 25,067,691.

5.3.2 Gregory-Newton backward difference

x_o is fixed at 1999 and h at *10*.

Now for us to calculate the population for the years 1995-1998 inclusive we employ the Gregory-Newton backward difference formula and use table 2.2 thus the formula is written as.

$$P(x) = y_n+s\nabla y_n+\frac{s(s+1)}{2!}\nabla^2 y_n+\frac{s(s+1)(s+2)}{3!}\nabla^3 y_n \qquad (5.3)$$

for x= *1995*

$\frac{x-x_o}{h} = s$, , x_o=1999, $s=\frac{1995-1999}{10} = -0.4$, $y_n = 28,686,607$

$$P(1995)=28,686,607+(-0.4)(7,237,833)+\frac{(-0.4)(0.6)(1,116,120)}{2!}+\frac{(-0.4)(0.6)(1.6)(-621,237)}{3!}$$

$$= 28,686,607-2,895,133.2-133,934.4+39,759.168$$

$$= 25,697,299$$

for $x = 1996$

$s=\frac{x-x_{o,}}{h}$, $s=\frac{1996-1999}{10} = -0.3$, $y_n = 28,686,607$

$$P(1996)=28,686,607+(-0.3)(7,237,833)+\frac{(-0.3)(0.7)(1,116,120)}{2!}+\frac{(-0.3)(0.7)(1.7)(-621,237)}{3!}$$

$$= 28,686,607-2,171,349.9-117,192.6+36,963.602$$

$$= 26,435,028$$

for $x = 1997$

$s=\frac{x-x_{o,}}{h}$, $s=\frac{1997-1999}{10} = -0.2$, $y_n = 28,686,607$

$$P(1997)=28,686,607+(-0.2)(7,237,833)+\frac{(-0.2)(0.8)(1,116,120)}{2!}+\frac{(-0.2)(0.8)(1.8)(-621,237)}{3!}$$

$$= 28,686,607-1,447,566.6-89,289.6+29,819.376$$

$$= 27,179,570$$

for $x = 1998$

$s=\frac{x-x_{o,}}{h}$, $s=\frac{1998-1999}{10} = -0.1$, $y_n = 28,686,607$

$$P(1998)=28,686,607+(-0.1)(7,237,833)+\frac{(-0.1)(0.9)(1,116,120)}{2!}+\frac{(-0.1)(0.9)(1.9)(-621,237)}{3!}$$

$$= 28,686,607-7,237,833.3-50,225.4+17,705.255$$

$$= 28,930,304$$

5.4 Conclusion

From the preceding calculations, we see that Numerical Methods can be used to predict population for a given country for years with no corresponding enumerated values.

CHAPTER 6

PROJECTION OF KENYA'S POPULATION FOR THE YEAR 1990-2010 USING STATISTICAL METHOD

6.0 Introduction

Statistical approaches (methods) have been used to project and calculate population for many years, see table 2.1 Kenya population for the years 1969, 1979, 1989 and 1999. This technique produces reliable results and therefore we use it to project Kenya population for the years 1990 to 2010 using the year 1999 as the base year and its corresponding population as the base population. The results will then be compared with those projected using numerical technique.

6.1 Using 1999 as the base year

Here the formula used to project the population is

$$p_n = p_o e^{rt} \tag{6.0}$$

where p_n is the year on which projection is done

p_o is the base year's population

r is the intercensal growth rate

t is the time interval between the base year and the year for which the population is projected, respectively.

This formula is exponential due to the presence of e and it is called the natural logarithm, or Napierian logarithm and it has the great virtue that it emphasizes that the base of the logarithm is *e***.** However in recent years *ln(x)* has been universally adopted as the standard abbreviation.

$$r = \frac{1}{t} \ln \frac{(1999\ population)}{(1989\ population)} \qquad (6.1)$$

can be refereed to as intercensal growth rate and therefore $t=10$

From table 2.0 Kenya's population in the years 1989 and 1999 is 21,448,774 and 28,686,607 respectively thus applying equation 6.2 we have

$$r = \frac{1}{10} \ln \frac{(28,686,607)}{(21,448,774)}$$

$$r = \frac{1}{10} \ln 1.3374474$$

$$r = \frac{1}{10} \ X \ 0.2907628$$

$$= 0.02908$$

$$= 0.029 \qquad (6.2)$$

Therefore equation 6.1 becomes

$$p_n = p_o e^{0.029t} \qquad (6.3)$$

This equation can now be used to project Kenya's population for the years 1990 to 2010, i.e. taking 1999 as the base year.

Projection for the year 1990

$t = 1990\text{-}1999 = -9$

$P(1990) = 28,686,607xe^{0.029(-9)}$

$= 28,686,607 \ x \ e^{-0.0261}$

$= 28,686,607x0.7702809$

$= 22,096,746$

Projection for the year 1991

$t = 1991 - 1999 = -8$

$P(1991) = 28{,}686{,}607 \times e^{0.029(-8)}$

$= 28{,}686{,}607 \times e^{-0.232}$

$= 28{,}686{,}607 x 0.7929461$

$= 22{,}746{,}934$

Projection for the year 1992

$t = 1992 - 1999 = -7$

$P(1992) = 28{,}686{,}607 \times e^{0.029(-7)}$

$= 28{,}686{,}607 \times e^{-0.203}$

$= 28{,}686{,}607 X 0.8162782$

$= 23{,}416{,}253$

Projection for the year 1993

$t = 1993 - 1999 = -6$

$P(1993) = 28{,}686{,}607 \times e^{0.029(-6)}$

$= 28{,}686{,}607 \times e^{-0.174}$

$= 28{,}686{,}607 x\ 0.8402969$

$= 24{,}105{,}267$

Projection for the year 1994

$t = 1994 - 1999 = -5$

$P(1994) = 28{,}686{,}607 \times e^{0.029(-5)}$

$= 28{,}686{,}607 \times e^{-0.145}$

$= 28{,}686{,}607 X 0.8650222$

$= 24{,}814{,}555$

Projection for the year 1995

$t = 1995 - 1999 = -4$

$P(1995) = 28,686,607 \times e^{0.029(-4)}$

$= 28,686,607 \times e^{-0.116}$

$= 28,686,607x\ 0.8904752$

$= 25,544,712$

Projection for the year 1996

$t = 1996 - 1999 = -3$

$P(1996) = 28,686,607 \times e^{0.029(-3)}$

$= 28,686,607 \times e^{-0.087}$

$= 28,686,607X0.9166771$

$= 26,296,356$

Projection for the year 1997

$t = 1997 - 1999 = -2$

$P(1997) = 28,686,607 \times e^{0.029(-2)}$

$= 28,686,607 \times e^{-0.058}$

$= 28,686,607x0.9436499$

$= 27,070,115$

Projection for the year 1998

$t = 1998 - 1999 = -1$

$P(1998) = 28,686,607 \times e^{0.029(-1)}$

$= 28,686,607 \times e^{-0.029}$

$= 28,686,607x0.9714164$

P(1999) = 27,866,642 enumerated during the census period

Projection for the year 2000

$P = 2000 - 1999 = 1$

$P(2000) = 28,686,607 \text{ x } e^{0.029(1)}$

$= 28,686,607x1.0294246$

$= 29,530,699$

Projection for the year 2001

$t = 2001 - 1999 = 2$

$P(2001) = 28,686,607 \text{ x } e^{0.029x2}$

$= 28,686,607 \text{ x } e^{0.058}$

$= 28,686,607x1.059715$

$= 30,399,628$

Projection for the year 2002

$t = 2002 - 1999 = 3$

$P(2002) = 28,686,607 \text{ x } e^{0.029x3}$

$= 28,686,607 \text{ x } e^{0.087}$

$= 28,686,607x1.0908967$

$= 31,294,124$

Projection for the year 2003

$t = 2003 - 1999 = 4$

$P(2003) = 28,686,607 \text{ x } e^{0.029x4}$

$= 28,686,607 \text{ x } e^{0.116}$

$= 28,686,607x1.1229959$

$= 32,214,941$

Projection for the year 2004

$t = 2004 - 1999 = 5$

$P(2004) = 28,686,607 \times e^{0.029x5}$

$= 28,686,607 \times e^{0.145}$

$= 28,686,607x1.1560396$

$= 33,162,853$

Projection for the year 2005

t = 2005 – 1999 = 6

$P(2005) = 28,686,607 \times e^{0.029x6}$

$= 28,686,607 \times e^{0.174}$

$= 28,686,607x1.1900556$

$= 34,138,656$

Projection for the year 2006

$t = 2006 - 1999 = 7$

$P(2006) = 28,686,607 \times e^{0.029x7}$

$= 28,686,607 \times e^{0.203}$

$= 28,686,607x1.2250725$

$= 35,143,172$

Projection for the year 2007

$t = 2007 - 1999 = 8$

$P(2007) = 28,686,607 \times e^{0.029x8}$

$= 28,686,607 \times e^{0.232}$

$= 28,686,607x1.2611197$

$= 36,177,246$

Projection for the year 2008

$t = 2008 - 1999 = 9$

$P(2008) = 28,686,607 \times e^{0.029x9}$

$= 28,686,607 \times e^{0.261}$

$= 28,686,607 \times 1.2982277$

$= 37,241,747$

Projection for the year 2009

$t = 2009 - 1999 = 10$

$P(2009) = 28,686,607 \times e^{0.029 \times 10}$

$= 28,686,607 \times e^{0.29}$

$= 28,686,607 \times 1.3364275$

$= 38,337,570$

Projection for the year 2010

$t = 2010 - 1999 = 11$

$P(2010) = 28,686,607 \times e^{0.029 \times 11}$

$= 28,686,607 \times e^{0.319}$

$= 28,686,607 \times 1.3757513$

$= 39,465,638$

6.2 Conclusion

In this chapter, we see that Statistical Methods can be used to predict population for all years whose corresponding enumerated values are not available.

CHAPTER 7

COMPARISON OF PROJECTED VALUES FROM NUMERICAL AND STATISTICAL METHODS

7.0 Introduction

Various columns are drawn to represent various values, the first column shows various years, second column shows values obtained by statistical method, third column shows values obtained when numerical method is used. the fourth column shows the differences between the values obtained by statistical method and the one obtained by numerical method. The fifth column shows the difference as the percentage of the respective statistical values. The sixth column shows the enumerated population for the year 1989 and 199 respectively.

7.1 The Projected Values

The projected values for the Kenya's population using 1999 as the base year are summarised in the following table, i.e. table 7.0 below and their corresponding bar graphs and line graphs are shown in figures 7.0 and 7.2, respectively.

Year	*Statistical Method (Base Year 1989)* A	*Numerical Methods (base year 1989)* B	*Difference (absolute values)* A-B= C	*Difference as a percentage of statistical value* $\frac{C}{A}x100 = D$	*Enumerated*
1990	*23,715,000*	*22,157,000*	*1,558,000*	*6.570*	
1991	*24,477,000*	*22,882,000*	*1,595,000*	*6.516*	
1992	*25,240,000*	*23,624,000*	*1,616,000*	*6.403*	
1993	*26,002,000*	*24,384,000*	*1,618,000*	*6.222*	
1994	*26,762,000*	*25,161,000*	*1,601,000*	*5.982*	
1995	*27,519,000*	*25,956,000*	*1,563,000*	*5.680*	
1996	*28,267,000*	*26,768,000*	*1,499,000*	*5.303*	
1997	*29,010,000*	*27,597,000*	*1,413,000*	*4.871*	
1998	*29,746,000*	*28,444,000*	*1,302,000*	*4.377*	
1999	*30,473,000*	*29,308,000*	*1,165,000*	*3.823*	*28,686,607*
2000	*31,189,000*	*30,189,000*	*1,000,000*	*3.206*	
2001	*31,896,000*	*31,088,000*	*808,000*	*2.533*	
2002	*32,588,000*	*32,004,000*	*584,000*	*1.792*	
2003	*33,264,000*	*32,938,000*	*326,000*	*0.9800*	
2004	*33,922,000*	*33,889,000*	*33,000*	*0.0973*	
2005	*34,561,000*	*34,857,000*	*296,000*	*0.0856*	
2006	*35,178,000*	*35,843,000*	*665,000*	*0.0189*	
2007	*35,773,000*	*36,846,000*	*1,073,000*	*0.0300*	
2008	*36,344,000*	*37,866,000*	*1,522,000*	*4.188*	
2009	*36,890,000*	*38,904,000*	*2,014,000*	*5.459*	
2010	*37,408,000*	*39,959,000*	*2,551,000*	*6.819*	

Table 7.0: Comparison of values obtained by the use of statistical and numerical methods. 1989 used as base year.

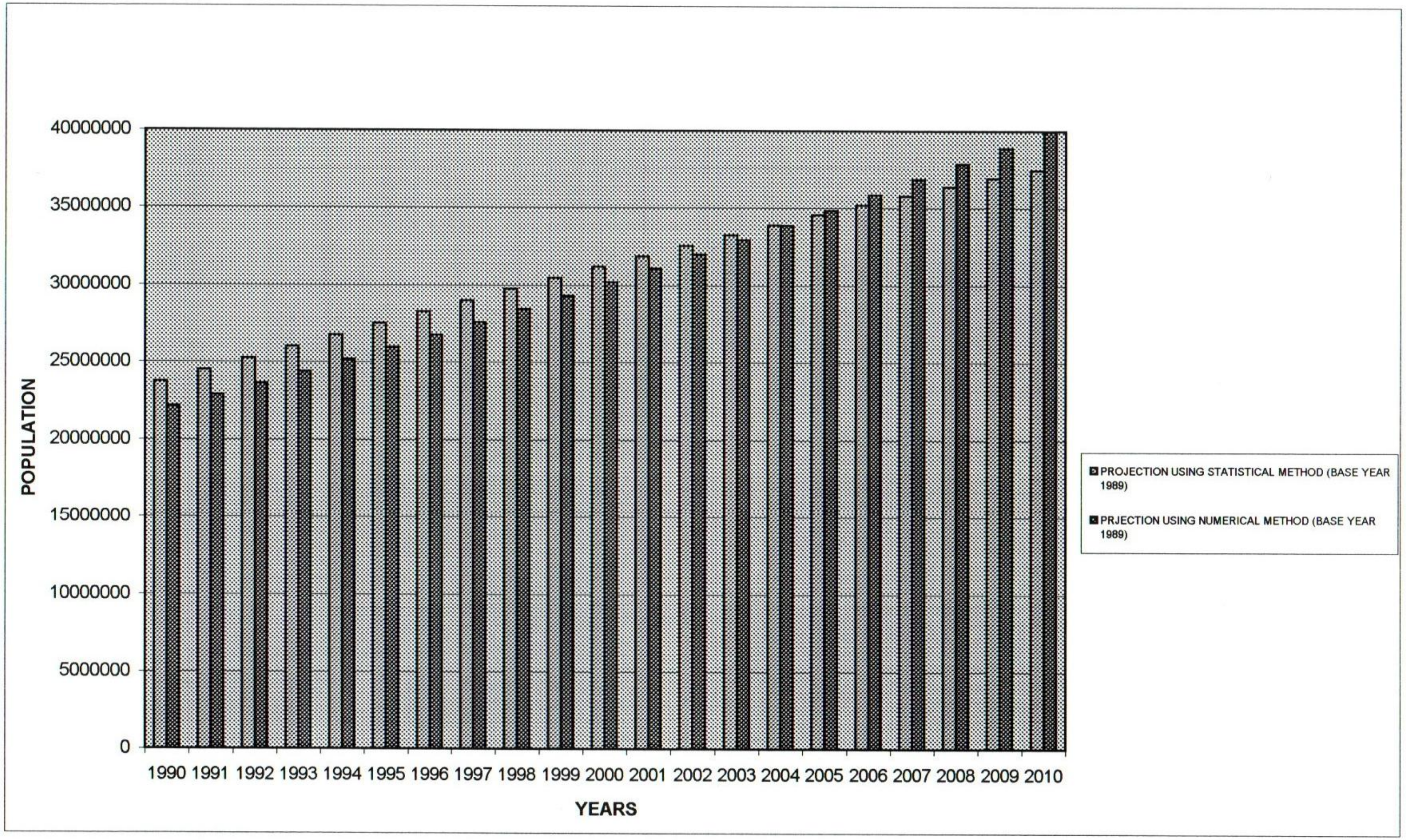

Figure 70: Bar graphs for projected population using 1989 as the base year

7.2 Comparison of the Two Methods Using 1999 as the Base Year

The projected population using the two methods i.e statistical and numerical methods are summarised in table 7.1 below and their corresponding bar graphs and line graphs are as seen in figures 7.1 and 7.3 respectively, here the year 1999 is used as the base year and its corresponding population as the base population.

	Base Year 1999 E	***Base Year 1999 F***	***(absolute value)***		
Year	***Projection using statistical method***	***Projection using numerical method***	***Difference* E-F=G**	***Differences as percentage of E*** $\frac{G \times 1000}{E} = H$	***Enumerated***
1990	*22,097,000*	*22,173,000*	*76,000*	*0.3439*	
1991	*22,747,000*	*22,896,000*	*149,000*	*0.6550*	
1992	*23,416,000*	*23,620,000*	*204,000*	*0.8712*	
1993	*24,105,000*	*24,344,000*	*239,000*	*0.9915*	
1994	*24,815,000*	*25,161,000*	*346,000*	*1.394*	
1995	*25,545,000*	*25,697,000*	*152,000*	*0.5950*	
1996	*26,296,000*	*26,435,000*	*139,000*	*0.5286*	
1997	*27,070,000*	*27,180,000*	*110,000*	*0.4064*	
1998	*27,867,000*	*28,930,000*	*1,063,000*	*3.8145*	
1999	*28,686,917*	*29,546,290*	*85,9373*	*2.996*	
1999	*28,686,607*	*29,798,000*	*1,111,393*	*3.874*	*28,686,607*
2000	*29,531,000*	*29,448,000*	*83,000*	*0.2811*	
2001	*30,400,000*	*30,213,000*	*187,000*	*0.6151*	
2002	*31,294,000*	*30,983,000*	*311,000*	*0.9938*	
2003	*32,215,000*	*31,755,000*	*46,000*	*1.428*	
2004	*33,163,000*	*32,530,000*	*633,000*	*1.9088*	
2005	*34,139,000*	*33,307,000*	*832,000*	*2.437*	
2006	*35,143,000*	*34,085,000*	*1,058,000*	*3.0106*	
2007	*36,177,000*	*34,863,000*	*1,314,000*	*3.632*	
2008	*37,242,000*	*35,641,000*	*1,601,000*	*4.2989*	
2009	*38,338,000*	*36,419,000*	*1,919,000*	*5.005*	
2010	*39,466,000*	*37,196,000*	*2,270,000*	*5.7518*	

Table 7.1: Comparison of values obtained by the use of statistical and numerical methods. 1999 used as base year

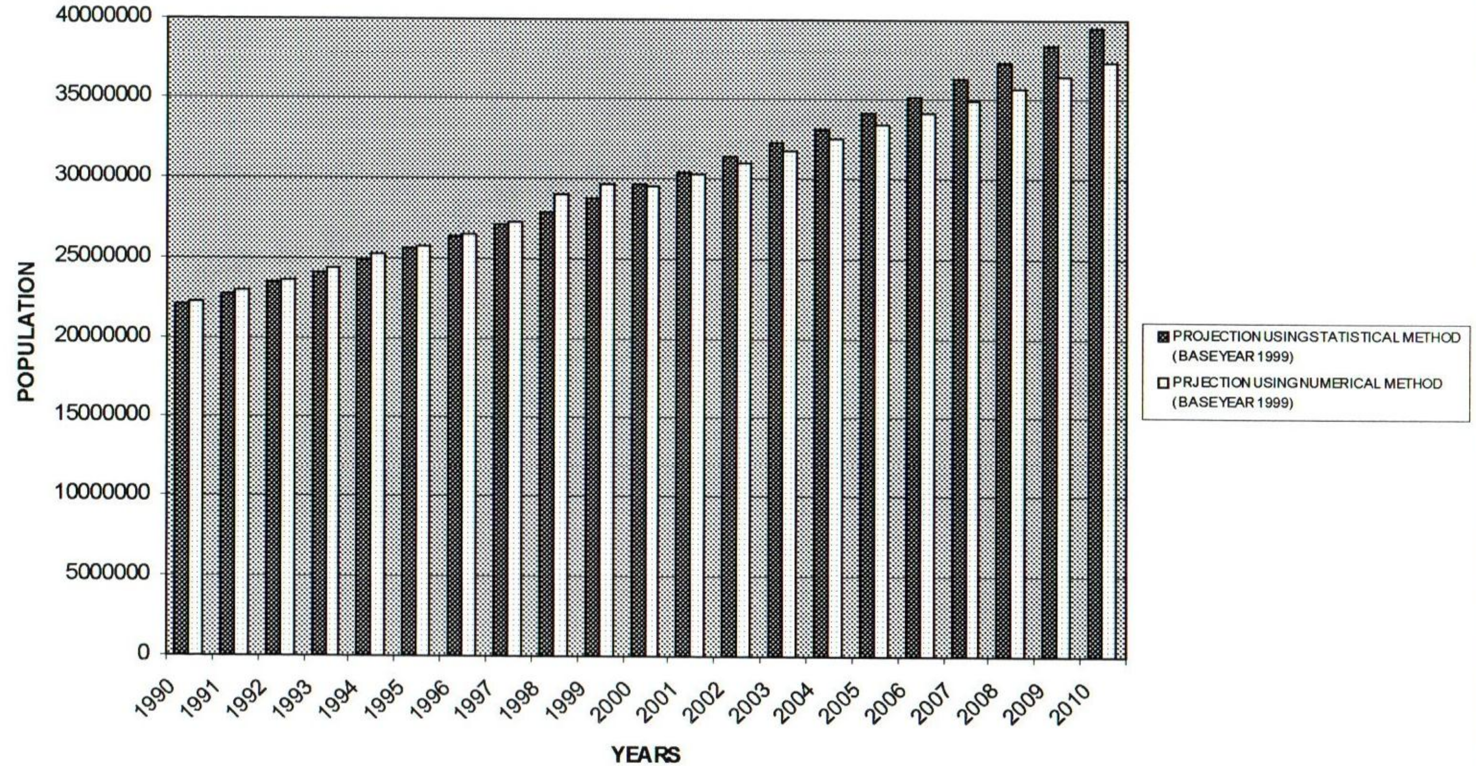

Figure 7.1: (A2) Bar graphs for projected population using 1999 as the base year

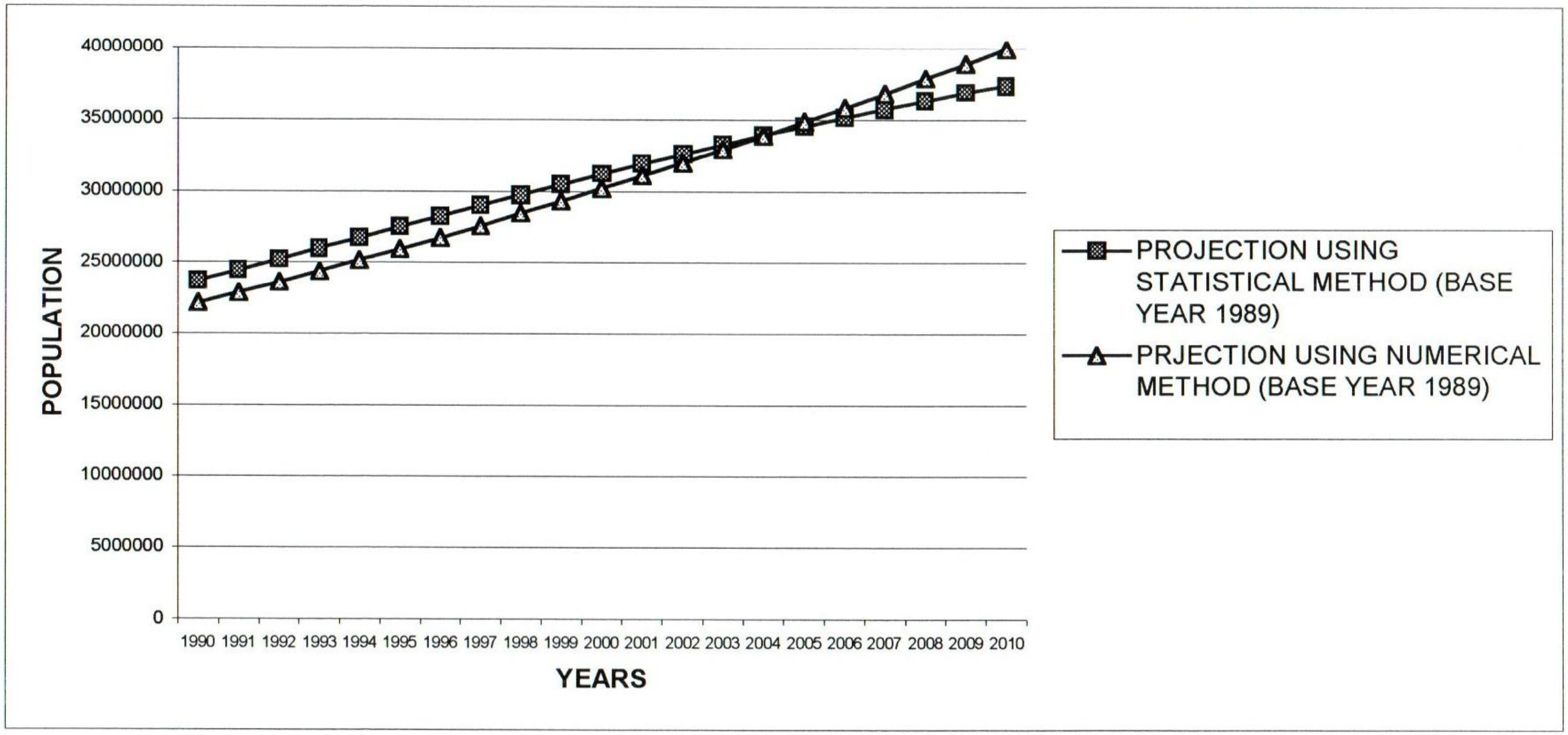

Figure 7.2: Line graph for projected population using 1989 as the base year

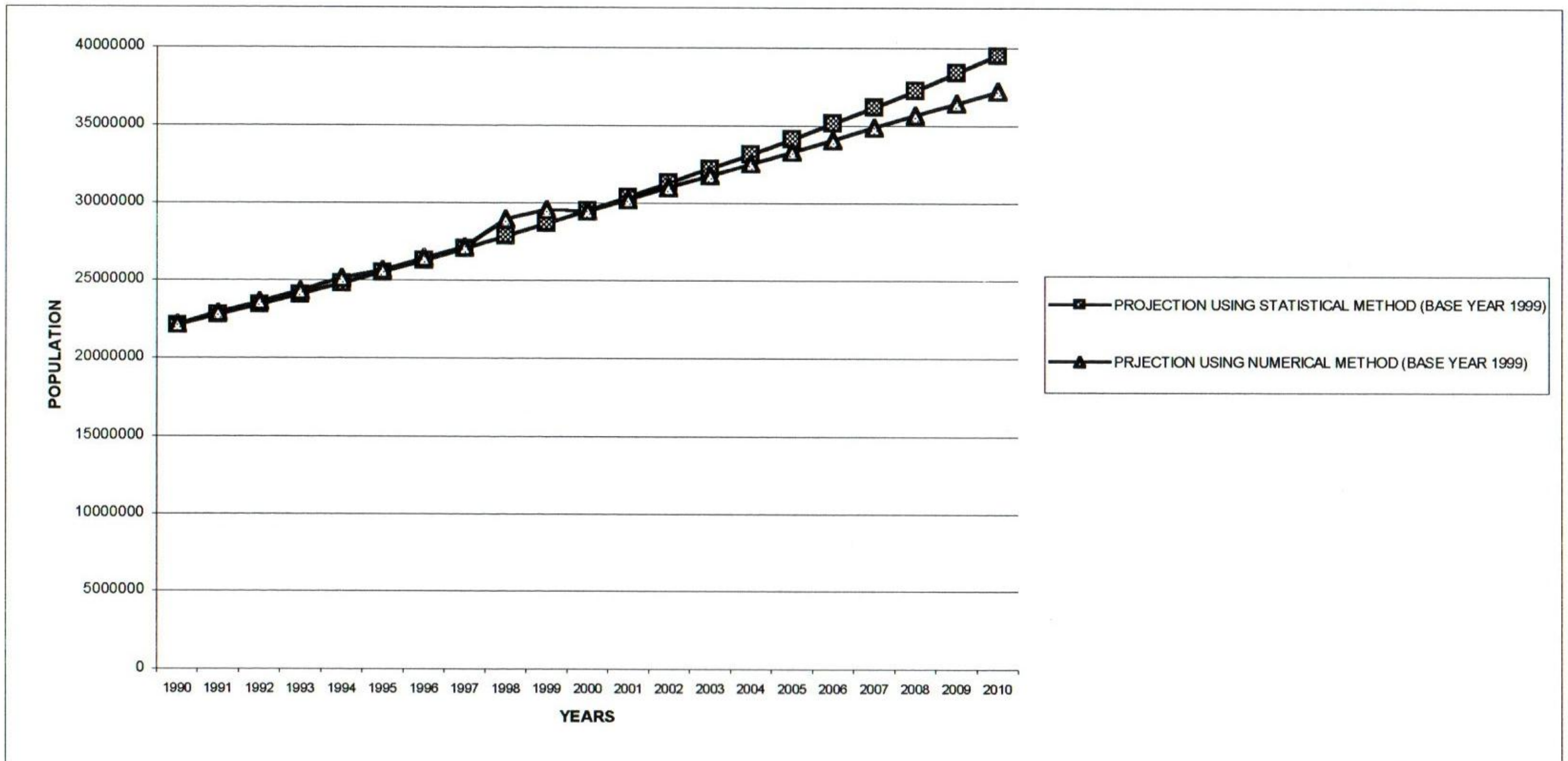

Figure 7.3: (B2) Line graph for projected population using 1999 as the base year

7.3 Discussion of Results

It can be observed from both the bar graphs and the line graphs that both methods give fairly accurate values when compared to the enumerated population. However, the trend of the bar graphs when 1999 population is used as the base population numerical method seems to be more accurate than the statistical method. This can be attributed to deaths from AIDS, mortality rate is high among the children, fertility decline among women due to family planning, working groups, education levels e.t.c. It can also be observed from the line graph figure 7.3 that the trend in population growth is low. Thus numerical method is accurate.

From table 7.0 it is seen that the differences expressed as the percentage of a statistical values ranges between 0.0189% and 6.8%. The numerical method gave a value of 29,308,000, statistical method gave a value of 30,473,507, while the numerated value was 28,686,607. Thus it is seen that numerical methods gives a fairly accurate value when compared to enumerated values.

When we calculated the difference between the enumerated values we find that the numerical method give better results. The difference is smaller than that of statistical method. This is clearly seen in the in the following comparison.

Statistical	***Enumerated***	***Difference***
30,473,000	28,686,607	1,786,393
Numerical	***Enumerated***	***Difference***
29,308,000	28,686,607	621,393

It can be observed that the difference is almost one third of the statistical method. This therefore suggests that the numerical method is a better method to use in the projection of the population. When 1999 was used as the base year the difference between the projected and enumerated value during the year 1999 are as follows

Statistical	***Enumerated***	***Difference***
28,686,917	28,686,607	310
Numerical	***Enumerated***	***Difference***
29,546,290	28,686,607	859,683

It can be observed that still the figure are close to the enumerated value. The much difference in this can be accounted for as the errors incurred as no errors term is considered for subtraction or addition. When we look at the inter year population growth rate for both projections using both 1989 and 1999 as the base years, we find that the growth rate is almost the same, which is almost 1.

Statistical Projection

$\frac{1991 \text{ (value)}}{1990 \text{ (value)}}$

$= \frac{24,477,000}{23,715,000}$

$= 1.03221316$

Numerical projection

$\frac{1991 \text{ (value)}}{1990 \text{ (value)}}$

$\frac{22,882,000}{22,157,000} = 1.0327$

1.03272

To one or two decimal places the inter year growth rate is the same for statistical and numerical method. Therefore the numerical method can also be used to project Kenya's population.

When 1999 was used as base year.

$\frac{1991 \text{ (value)}}{1990 \text{ (value)}}$

Statistical projection

$\frac{22,747,000}{22,097,000}$

$= 1.0294158$

Numerical projection

$\frac{22,896,000}{22,173,000}$

$= 1.0326072$

Again correct to two decimal places the inter year population growth rate is 1.03

Therefore we can now conclude that even the numerical methods give better approximation for the population projection.

Also it can be observed that the extrapolated values are closer to the numerical values and the statistical values.

CHAPTER 8

USE OF DIFFERENCE TABLE AND AVERAGING METHOD TO PROJECT THE POPULATION

8.0 Introduction

Other methods of population projection and calculation are discussed and this involves averaging the values below and above the gaps of difference tables see table 2.4

8.1 Averaging Method

Here we construct a central difference table as in table 2.4. We leave space between each line of data and define the lines containing the original data as full lines and the lines between the full lines as half-lines.

We then take the difference and fill the gaps in the table by finding arithmetic mean of the values above and below each gap. Let us re-construct the table again and proceed as explained above.

Year x	Population $f(x)$	δf	$\delta^2 f$	$\delta^3 f$
1969	*10,942,705*			
1971.5	*12,038,794*			
1974	*13,134,883*	*4,384,356*		
1976.5	*14,230,972*	*4,818,695*		
1979	*15,327,061*	*5,253,034*	*1,737,357*	
1981.5	*16,857,490*	*5,687,373.5*	*1,582,047.8*	
1984	*18,387,918*	*6,121,713*	*1,426,738.5*	*-621,*
1986.5	*19,918,346*	*6,400,743*	*1,271,421.3*	
1989	*21,448,774*	*6,679,773*	*1,116,120*	
1991.5	*23,258,233*	*6,958,803*		
1994	*25,067,691*	*7,237,833*		
1996.5	*26,877,149*			
1999	*28,686,627*			

Table 8.0 : Population projection by finding arithmetic means of values above and below the gaps.

Now we proceed again to leave space between each line and find the difference and the arithmetic mean. The result is as in table 8.0 above.

Now we calculate the growth rate after every five years.

1) $\frac{1974 \text{ (value)}}{1969 \text{ (value)}} = \frac{13,134,883}{10,942,705} = 1.2003324 \cong 1.2$

2) $\frac{1979 \text{ (value)}}{1974 \text{ (value)}} = \frac{15,327,061}{13,134,883} = 1.1668974 \cong 1.2$

3) $\frac{1984 \text{ (value)}}{1979 \text{ (value)}} = \frac{18,387,918}{15,327,061} = 1.1997028 \cong 1.2$

4) $\frac{1989 \text{ (value)}}{1984 \text{ (value)}} = \frac{21,448,774}{18,387,918} = 1.1664602 \cong 1.2$

5) $\frac{1994 \text{ (value)}}{1989 \text{ (value)}} = \frac{25,067,691}{21,448,774} = 1.1687237 \cong 1.2$

6) $\frac{1999 \text{ (value)}}{1994 \text{ (value)}} = \frac{28,886,607}{25,067,691} = 1.1443657 \cong 1.1$

We can now observe that the growth rate every after five years has been 1.2 to two decimal places. Now due other factors such as AIDS, mortality rates and fertility decline we can see it has declined to 1.144 to four significant figures.

We can now then project the population for the year 2004 as 286,866,077x1.144 = 32,817,478 and that of 2009 as 32,817,478x1.144 = 37,543,195. From table 8.0, we can see that the figures are almost the same but since all these are projections and we have not considered the errors made then we can conclude that the difference table and the averaging method which is a numerical method can be used to project Kenya's population.

It can also be observed that if we continue finding the differences and the arithmetic mean we shall be in the position to calculate the population every after 2 months or 6 months.

CHAPTER 9

POPULATION PROJECTIONS PER PROVINCES FOR YEARS 2004 AND 2009

9.0 Introduction

Projection and calculation of population per province is done and the sum of the various values for all provinces is obtained and compared with entire population projections for the whole country.

For each and every province we construct three tables, namely a forward difference table, a backward difference table and a central difference table. However, for the projections for the said years, i.e. 2004 and 2009, we only construct a backward difference table and since we are extrapolating we use the Gregory-Newton backward formula.

9.1 Projection for Coast Province

Backward difference table

Year x	Population f(x)	∇f	$\nabla^2 f$	$\nabla^3 f$
1969	944,082			
1974	1,143,438			
1979	1,342,794	398,712		
1984	1,584,277.5	440,839.5		
1989	1,825,761	482,967	84,255	
1994	2,156,512.5	572,236	131,395.5	
1999	2,487,264	661,503	178,536	94,281

Table 9.0: Backward differences for the years 1969, 1979, 1989 and 1999 for Coast Province

It can be noted that, if we take the arithmetic mean above and below the given data, we can get the population for the years 1974, 1984, and 1994 as indicated in the table 9.0. Now for the year 2004, we have $x= 2004$, $x_o=1999$, $h=10$, y_n is the census enumerated value i.e 1999 population.

$$s=\frac{x-x_o}{h} = \frac{2004-1999}{10} = \frac{5}{10} = 0.5$$

$$P(2004)= y_n+s\nabla y_n+\frac{s(s+1)}{2!}\nabla^2 y_n+\frac{s(s+1)(s+2)}{3!}\nabla^3 y_n \qquad (9.0)$$

$$= 2,487,264+(0.5)(661,503)+\frac{(0.5)(1.5)(178,536)}{2!}+\frac{(.5)(1.5)(2.5)(94,281)}{3!}$$

$$= 2,487,264+330,751.5+66,951+29,462.813$$

$$= 2,914,429$$

For $x = 2009$, $x_o= 1999$, $h=10$, $y_n = 28,686,607$

$$S = \frac{x-x_o}{h} = \frac{2009-1999}{10} = 1$$

$$P(2004)= 2,487,264+(1.0)(661,503)+\frac{(1)(2)(178,536)}{2!}+\frac{(1)(2)(3)(94,281)}{3!}$$

$$= 2,487,264+661503+178,536+94,281$$

$$= 3,421,584$$

9.2 Projection for North Eastern Province

Backward difference table

year *x*	population f(*x*)	∇f	$\nabla^2 f$	$\nabla^3 f$
1969	*245,757*			
1974	*309,772*			
1979	*373,787*	*128,030*		
1984	*372,589*	*62,817*		
1989	*371,391*	*-2,396*	*-130,426*	
1994	*666,767*	*294,178*	*231,361*	
1999	*962,143*	*590,752*	*593,148*	*723,574*

Table 9.1: Backward differences for the years 1969, 1979, 1989 and 1999 for North Eastern Province.

As usual the arithmetic mean of the value above and below the given data gives the populations for the years 1974, 1984 and 1994 as indicated in the table 9.1.

$$P(2004) = y_n + s\nabla y_n + \frac{s(s+1)}{2!}\nabla^2 y_n + \frac{s(s+1)(s+2)}{3!}\nabla^3 y_n \qquad (9.1)$$

Here $x = 2004$, $x_o = 1999$, $h = 10$. $y_n = 28{,}686{,}607$

$$s = \frac{x - x_o}{h} = \frac{2004 - 1999}{10} = \frac{5}{10} = 0.5$$

$$P(2004) = 962{,}143 + (0.5)(590{,}752) + \frac{(0.5)(1.5)(593{,}148)}{2!} + \frac{(0.5)(1.5)(2.5)(223{,}574)}{3!}$$

$$= 962{,}143 + 295{,}376 + 222{,}430.5 + 226{,}116.88$$

$$= 1{,}706{,}066$$

for $x = 2009$, $x_o = 1999$, $h = 10$, $y_n = 28{,}686{,}607$

$$s = \frac{x - x_o}{h} = \frac{2009 - 1999}{10} = 1$$

$$P(2009) = 962{,}143 + (1)(590{,}752) + \frac{(1)(2)(593148)}{2!} + \frac{(1)(2)(3)(723{,}574)}{3!}$$

$$= 962{,}143 + 590{,}752 + 593{,}148 + 723{,}574$$

$$= 2{,}869{,}617$$

9.3 Projection for Nairobi Province

Backward difference table

year x	population $f(x)$	∇f	$\nabla^2 f$	$\nabla^3 f$
1969	509,286			
1974	668,530.5			
1979	827,775	318,489		
1984	1,076,172.5	407,642		
1989	1,324,570	496,795	178,306	
1994	1,733,912	657,739.5	250,097.5	
1999	2,143,254	818,684	321,889	143,583

Table 9.2: Backward differences for the years 1969, 1979, 1989 and 1999 for Nairobi Province.

The Kenya's population for the years 1979, 1984 and 1994 are as gotten above as the arithmetic means of values above and below each gap of the difference table.

For $x_o = 1999,\ x = 2004,\ h = 10,\quad y_n = 28{,}686{,}607$

$$s = \frac{x - x_o}{h} = \frac{2004 - 1999}{10} = 0.5$$

$$P(2004) = y_n + s\nabla y_n + \frac{s(s+1)}{2!}\nabla^2 y_n + \frac{s(s+1)(s+2)}{3!}\nabla^3 y_n \qquad (9.2)$$

$$= 2{,}143{,}254 + (0.5)(818{,}684) + \frac{(0.5)(1.5)(321{,}889)}{2!} + \frac{(0.5)(1.5)(2.5)(143{,}583)}{3!}$$

$$= 21{,}43{,}254 + 409{,}342 + 120{,}708.38 + 44{,}869.688$$

$$= 2{,}718{,}174.1$$

for $x = 2009,\ x_o = 1999,\ h = 10,\quad y_n = 28{,}686{,}607.$

$$s = \frac{x - x_o}{h} = \frac{2009 - 1999}{10} = 1$$

$$P(2009) = 2{,}143{,}254 + (1)(818{,}684) + \frac{(1)(2)(321{,}889)}{2!} + \frac{(1)(2)(3)(143{,}583)}{3!}$$

$$= 2{,}143{,}254 + 818{,}684 + 321{,}889 + 143{,}583$$

$$= 3{,}427{,}410$$

9.4 Projection for Central Province

Backward difference table

year x	population $f(x)$	∇f	$\nabla^2 f$	$\nabla^3 f$
1969	1,675,647			
1974	2,010,740			
1979	2,345,833	670,186		
1984	2,728,544	717,804		
1989	3,111,255	765,422	95,236	
1994	3,417,707	689,163	- 28,641	
1999	3,724,159	612,904	- 152,518	- 247,754

Table 9.3 Backward differences for the years 1969, 1979, 1989 and 1999 for Central Province.

For the years 1974, 1984 & 1994 we find the arithmetic means as explained earlier to give Kenya's population.

For x_o=1999, x=2004, h= 10, y_n = 28,686,607

$$s = \frac{x - x_o}{h} = \frac{2004 - 1999}{10} = 0.5$$

$$P(2004) = y_n + s\nabla y_n + \frac{s(s+1)}{2!}\nabla^2 y_n + \frac{s(s+1)(s+2)}{3!}\nabla^3 y_n \qquad (9.3)$$

$$= 3,724,159 + (0.5)(612,904) + \frac{(0.5)(1.5)(-152,518)}{2!} + \frac{(0.5)(1.5)(2.5)(-247,754)}{3!}$$

$$= 37,241,59 + 306,452 - 57,194.25 - 77,423.125$$

$$= 3,895,993.6$$

For x_o=1999, x=2009, h= 10, y_n = 28,686,607

$$s = \frac{x - x_o}{h} = \frac{2009 - 1999}{10} = 1$$

$$P(2009) = 3,724,159 + 612,904 - 152,518 - 247,754$$

$$= 3,936,791$$

9.5 Projection for Rift Valley Province

Backward difference table

year x	population $f(x)$	∇f	$\nabla^2 f$	$\nabla^3 f$
1969	*2,210,289*			
1974	*2,725,345.5*			
1979	*3,240,402*	*1,030,113*		
1984	*4,078,976.5*	*1,353,631*		
1989	*4,917,551*	*1,677,149*	*647,036*	
1994	*5,952,293.5*	*1,873,317*	*519,686*	
1999	*6,987,036*	*2,169,485*	*392,336*	*- 254,700*

Table9.4 Backward differences for the years 1969, 1979, 1989 and 1999 for Rift Valley Province

Arithmetic means give Kenya's population for the years 1974, 1984 and 1994 as in table 9.4.

For x_o=1999, x=2004, h= 10, y_n = 28,686,607

$$s = \frac{x - x_o}{h} = \frac{2004 - 1999}{10} = 0.5$$

$$P(2004) = y_n + s\nabla y_n + \frac{s(s+1)}{2!}\nabla^2 y_n + \frac{s(s+1)(s+2)}{3!}\nabla^3 y_n \quad (9.4)$$

$$= 6,987,036 + (0.5)(2,069,485) + \frac{(0.5)(1.5)(392,336)}{2!} + \frac{(0.5)(1.5)(2.5)(-254,700)}{3!}$$

$$= 6,987,036 + 1034,742.5 + 147,126 - 79,593.75$$

$$= 8,089,310.8$$

For x_o=1999, x=2004, h= 10, y_n = 28,686,607

$$s = \frac{x - x_o}{h} = \frac{2004 - 1999}{10} = 0.5$$

$$P(2009) = 6,987,036 + 2,069,485 + 392,336 - 254,700$$

$$= 9,194,157$$

9.6 Projection for Nyanza Province

Difference table

year x	population $f(x)$	∇f	$\nabla^2 f$	$\nabla^3 f$
1969	*2,122,045*			
1974	*2,383,000.5*			
1979	*2,643,956*	*521,911*		
1984	*3,075,558*	*692,557.5*		
1989	*3,507,160*	*863,204*	*341,293*	
1994	*3,949,678*	*874,120*	*181,562.5*	
1999	*4,392,196*	*885,036*	*21,832*	*- 319,461*

Table 9.5: Backward differences for the years 1969, 1979, 1989 and 1999 for Nyanza Province.

Averages as usual give Kenya's population for the years 1974, 1984, and 1994 as in table 9.5

For x_o=1999, x=2004, h= 10

$$s = \frac{x - x_o}{h} = \frac{2004 - 1999}{10} = 0.5$$

$$P(2004) = y_n + s\nabla y_n + \frac{s(s+1)}{2!}\nabla^2 y_n + \frac{s(s+1)(p+2)}{3!}\nabla^3 y_n \qquad (9.5)$$

$$P(2004) = 4,392,196 + (0.5)(885,036) + \frac{(0.5)(1.5)(21,832)}{2!} + \frac{(0.5)(1.5)(2.5)(-319,461)}{3!}$$

$$= 4,392,196 + 442,518 + 8,187 - 99,831.563$$

$$= 4,743,069.4$$

For x_o=1999, x=2009, h= 10, $y_n = 28,686,607$

$$s = \frac{x - x_o}{h} = \frac{2009 - 1999}{10} = 1.0$$

$$P(2009) = 4,392,196 + 885,036 + 21,832 - 319,461$$

$$= 4,976,603$$

9.7 Projection for Western Province

Backward difference table

year x	Population $f(x)$	∇f	$\nabla^2 f$	$\nabla^3 f$
1969	*1,328,298*			
1974	*1,580,480.5*			
1979	*1,832,663*	*544,365*		
1984	*2,227,530*	*647,049.5*		
1989	*2,622,397*	*789,735*	*285,369*	
1994	*2,990,586.5*	*763,056.5*	*116,007*	
1999	*3,358,776*	*736,379*	*- 53,355*	*- 338,724*

Table 9.6 : Backward differences for the years 1969, 1979, 1989 and 1999 for Western Province.

Arithmetic means give Kenya's population for the years 1974, 19

as in table 9.6 above.

For x_o=1999, x=2004, h= 10, $y_n = 28,686,607$

$$s = \frac{x - x_o}{h} = \frac{2004 - 1999}{10} = 0.5$$

$P(2004) = y_n + s\nabla y_n + \frac{s(s+1)}{2!}\nabla^2 y_n + \frac{s(s+1)(s+2)}{3!}$ … (9.7)

$= 3,358,776 + (0.5)(736,37$ … $(-422,036)$ … $3!$

$= 3,358,776 + 368$

$= 3,601,10$

For $x_o = 1999$

$s = \frac{x - x_o}{h} = \frac{200}{1}$

$P(2009) = 3,358,776$

$= 3,703,076$

9.7 Projection for Western Province

Backward difference table

year x	Population $f(x)$	∇f	$\nabla^2 f$	$\nabla^3 f$
1969	*1,328,298*			
1974	*1,580,480.5*			
1979	*1,832,663*	*544,365*		
1984	*2,227,530*	*647,049.5*		
1989	*2,622,397*	*789,735*	*285,369*	
1994	*2,990,586.5*	*763,056.5*	*116,007*	
1999	*3,358,776*	*736,379*	*- 53,355*	*- 338,724*

Table 9.6 : Backward differences for the years 1969, 1979, 1989 and 1999 for Western Province.

Arithmetic means give Kenya's population for the years 1974, 1984 and 1994 as in table 9.6 above.

For x_o=1999, x=2004, h= 10, $y_n = 28,686,607$

$$s = \frac{x - x_o}{h} = \frac{2004 - 1999}{10} = 0.5$$

$$P(2004) = y_n + s\nabla y_n + \frac{s(s+1)}{2!}\nabla^2 y_n + \frac{s(s+1)(s+2)}{3!}\nabla^3 y_n \qquad (9.6)$$

$$= 3,358,776 + (0.5)(736,379) + \frac{(0.5)(1.5)(-53,355)}{2!} + \frac{(0.5)(1.5)(2.5)(-338,724)}{3!}$$

$$= 3,358,776 + 368,189.5 - 20,008.125 - 105,851.25$$

$$= 3,601,106.1$$

For $x_o=1999$, $x=2009$, $h=10$, $y_n = 28,686,607$

$$s = \frac{x - x_o}{h} = \frac{2009 - 1999}{10} = 1.0$$

$$P(2009) = 3,358,776 + 736,379 - 53,355 - 338,724$$

$$= 3,703,076$$

9.8 Projection for Eastern Province

Backward difference

year x	Population $f(x)$	∇f	$\nabla^2 f$	$\nabla^3 f$
1969	*1,907,301*			
1974	*2,313,576*			
1979	*2,719,851*	*812,550*		
1984	*3,244,270*	*930,394*		
1989	*3,768,689*	*1,048,838*	*236,288*	
1994	*4,200,234*	*955,964*	*25,270*	
1999	*4,631,779*	*863,090*	*- 185,748*	*- 422,036*

Table 9.7: Backward differences for the years 1969, 1979, 1989 and 1999 for Eastern Province.

Arithmetic means give Kenya's population for the years 1974, 1984 and 1994 as in table 9.7 above.

For $x_o=1999,\ x=2004,\ h=10,\ \ y_n = 28,686,607$

$$s = \frac{x-x_o}{h} = \frac{2004 - 1999}{10} = 0.5$$

$$P(2004) = y_n + s\nabla y_n + \frac{s(s+1)}{2!}\nabla^2 y_n + \frac{s(s+1)(s+2)}{3!}\nabla^3 y_n \qquad (9.7)$$

$$= 4,631,779 + (0.5)(86,090) + \frac{(0.5)(1.5)(-185,748)}{2!} + \frac{(0.5)(1.5)(2.5)(-422,036)}{3!}$$

$$= 4,631,779 + 431,545 - 69,655.5 - 131,886.25$$

$$= 4,861,782.3$$

For $x_o=1999,\ x=2009,\ h=10,\ \ y_n = 28,686,607$

$$s = \frac{x-x_o}{h} = \frac{2009 - 1999}{10} = 1.0$$

$$P(2009) = 4,631,779 + 8,63,090 - 185,748 - 422,036$$

$$= 4,887,085$$

9.9 Comparison of the Sum for Individual Provinces and for Kenya

Here we sum all the projections for each and every province for the year 2004 and 2009 and compare when we projected for Kenya.

Year 2004

$=2,914,429+1,706,066+2,718,174.1+3,895,993.6+8,089,310.8+4,734,069.4+3,601,106.1+4861782.3 = 32,529,931$

Year 2009

$= 3,421,584+2,869,617+3,427,410+3,936,791+9,194,157+4,979,603+3,703,076+4,887,085 = 36,419,323$

As it can be seen from the previous projections for Kenya for the years 2004 and 2009 the figures are the same.

9.10 Conclusion

From the comparison of the projected and the enumerated values, we observe that the error is negligible. Thus these results suggest that numerical methods cal also be used to project our population and avoid the burden of preparing for the census. Also since census is carried out every ten years, we can project the population for each calendar years so that our planning or the objectives of census can be achieved. However, our projections will always differ slightly from the enumerated values because the method has not taken into account the mortality rates, fertility rates AIDS, outbreak of diseases such as Ebola, deaths resulting from road accidents and other forms of accidents, migration to other countries e.t.c.

CHAPTER 10

OTHER FACTORS THAT AFFECT POPULATION GROWTH

10.0 Introduction

Other factors that affect population growth have been discussed by other researchers who have used statistical method. In this dissertation we have overlooked this factors because the projections are based on academic interests. That is to say other factors which affect population growth were not considered when projecting the population.

10.1 Mortality and Fertility Projections

Mortality in Kenya is projected using the logit model life table system. For the purpose of the projections, the mortality is divided into two components. AIDS and other causes.

10.2 Mortality from Other Causes

This component is assumed to continue to decline upto 2010. However, the validity of this assumption is questionable on the various grounds such as health services, have been deteriorating in many parts of the country and even the KEPI program has been faltering; malaria has been on the increase, poverty and unemployment have also been increasing among some sections of the population.

However, there are counter arguments of these allegations, education of mothers which is one of the determinant of child mortality has continued to improve in Kenya. Birth order, family size and child spacing have also been shown to reduce mortality rate. Determinant of child mortality, thus the decline in fertility and increased use of contraception for child spacing can be expected to promote further decline in child mortality. Thus with all these consideration in mind, the assumption of mortality rate is supported.

10.3 Mortality from AIDS

This component is projected using a model developed by Basia Zaba (1994). In order to use the model to predict AIDS mortality, it is necessary to make estimates of HIV prevalence in Kenya for each year from the start of the epidemic until the year 2010.

10.4 Fertility Projections

Total fertility rates (TFRs) have been assumed at about 7.00 and 6.2 during the early and late 1980s, respectively. For the purpose of intercensal projections, it therefore seems probable that the total fertility rate will have fallen to below 6 in the early 1990s. For the purpose of the projections, a simple linear decline in total fertility rate is assumed see Kenya population census 1989, Analytical Report, April 1996 volume viii (pages 5,7 and 12)

10.5 Conclusion

In conclusion, we see that fairly accurate figures for Kenya's population for a particular year can be obtained after taking into account the mortality rates from AIDS, the fertility rate and other causes.

These projections were made without the AIDS component in the mortality calculations, and without the three different assumptions about the rate of fertility declined as well as projection on what would happen if fertility did not decline at all, and the Total Fertility Rates (TFR) remained constant at 6 births per woman.

However it is emphasized here that the calculations without AIDS have been made purely on academic interest, i.e. projection was done without taking into consideration other factors that might affect population growth rate. However, any realistic forecast of Kenya's population must include AIDS as other cause of mortality.

REFERENCES

Atkinson L.V., Harley P.J., Hudson J.D.: Numerical Methods with Fortran 77 Addison – Wesley Publishing Company, 1994.

Baltazar G.M. Estimating National HIV Prevalence from Sentinel Surveillance Data. Unpublished Note, 1994.

Brass B. Disciplining Demographic Data International Population Conference London, 1969.

Hoffman J.: Numerical Methods for Engineers and Scientists, London McGraw-Hill, 1993.

Hornbeck W. R.: Numerical Methods, New York, Quantum Publishers Inc. , 1975.

Hosking, Joe, Joyce and Turner: First Steps in Numerical Analysis, Arnold a member of the Hodder Headline Group 338 Euston Road London NW1 13 BH, 1996.

ICDDRB 1995 – International Centre for Diarrhoea Research, Bangladesh (1995). Demographic Surveillance System, Early indicators; Matlab, Dhaka Indicators, Matlab 1994, Dhaka.

Jaques I. and Judd C.: Numerical Analysis Chapman and Hall London New York, 1987.

Kenya's population census 1989 Analytical Report Volume vii (population projections' April 1996.

Kenya's population census 1999 Volume 1 printed by the Government Printer.

Kenya's population census 1999 Analytical Report printed by the Government Printer.

National AIDS Control Programme (NACP) "AIDS in Kenya Background Projections, Impact Interventions", 1979 Population Census Volume II. Analytical Report (PP. 155-122).

Smith, G.D. Numerical Solution of Partial Differential Equations, Finite Difference Methods, Clareuden Press, Oxford, 1996.

Zaba The Demographic Impact of AIDS some stable population, population simulation results London school of Hygiene and Tropical Medicine University of London, 1994.