APPROXIMATION OF FINITE POPULATION TOTALS USING LAGRANGE POLYNOMIAL

LAMIN KABAREH

MASTER OF SCIENCE
(Mathematics-Statistics Option)

Pan African University Institute for Basic Sciences, Technology and Innovation

2018

APPROXIMATION OF FINITE POPULATION TOTALS USING LAGRANGE POLYNOMIAL

LAMIN KABAREH

MS300-0007/16

A thesis submitted to Pan African University Institute of basic Sciences, Technology and Innovation in partial fulfillment of the requirements for the award of the degree of Master of Science in Mathematics (Statistics Option)

2018

DECLARATION

This thesis is my own work and has not been presented for a degree award in any other university.

Signature .... Date 12 /04/2018

LAMIN KABAREH

Declaration by supervisors.

This thesis has been submitted for examination with my approval as university supervisor.

Signature .... Date..

Dr Thomas Mageto,

Jomo Kenyatta University of Agriculture and Technology

Signature .... Date..

Dr Benjamin Muema,

Jomo Kenyatta University of Agriculture and Technology

ACKNOWLEDGEMENTS

This work has been possible through the help of many individuals who gave their support. I thank the Almighty God for granting me the knowledge to carry out this study. I also thank African Union for the financial support through PAUISTI during my period of study in Kenya. My Sincere thanks go to JKUAT and AICAD the host institutions of PAUISTI for allowing us use their facilities without any restriction. I can't overlook the love and support from my family and friends who encouraged me throughout my studies.

Special thanks to my supervisors Dr. Thomas Mageto and Dr. Benjamin Muema for accepting to supervise this work despite their busy schedules. I do appreciate their unlimited constructive advice, suggestions, ideas and comments towards the work. I highly appreciate their moral and academic support as well.

DEDICATION

This work is dedicated to my lovely mother, Adama and father, Augustos for their love, care, motivations and prayers in making this work a success.

Contents

LISTS OF FIGURES

LISTS OF ABBREVIATIONS

KNBS Kenya National Bureau of Statistics

CMVT Cauchy Mean Value Theorem

BAP Best Approximating Polynomial

LCQR Local Composite Quantile Regression

LPR Local Polynomial Regression

MSE Mean Square Error

HT Horvitz-Thompson

SRSWOR Simple Random Sampling Without Replacement

LISTS OF NOMENCLATURES

P_m Polynomial of degree m

U_N finite population

π inclusion probability

π_i first inclusion probability

π_{ij} first and second inclusion probabilities

$\prod$ Product

$\sum$ Summation

$\propto$ Proportional to

n Sample size

N Population size

s Sample

ξ Xi

e Residual

∞ infinity

ζ Zeta

$\in$ belongs to

η Eta

ω Omega

μ mu

ρ rho

$C^n[a, b]$ continuously differentiable bounded interval

(a, b) Open interval

ABSTRACT

In this thesis, attempt to study the effects of extreme observations on one approximator of finite population total is made. We are particularly approximating the finite population totals using the Lagrange polynomial of finite population totals given different finite populations. The study revealed that both the classical and the non parametric estimator based on the local linear polynomial give good outcomes when the auxiliary and the study variables are highly correlated. It is however realized that in the presence of outliers the local linear polynomial performs better with respect to design mean square error. However, this approach relies entirely on the bandwidth selection in order to attain a better precision. The Lagrange polynomial has been proposed which does not put emphasis on choice of bandwidth selection. The study developed a Lagrange polynomial and showed how to obtain the error term. However, the asymptotic properties are also determined with the use of the Karl Weierstrass theorem. This revealed that, the linear polynomial is the best approximating polynomial which can converge faster than other higher degree polynomials with high precision. Finally, the empirical analysis showed a good outcome which is in conformity with what the theorem revealed and gave a good projection of the population total for the coming census in Kenya in 2019.

CHAPTER ONE

1 INTRODUCTION

1.1 Background information

Survey sampling often provides information about a study variable only for sampled elements. However, auxiliary information is often provided for the whole population. The relationship of the auxiliary information with the study variable across the sample permits inferences about the non-sampled portion of the population. Thus, the use of auxiliary information at the estimation stage of a survey increases the accuracy of the estimates parameters studied. One approach to using this auxiliary information in estimation is to assume a working model describing the relationship between the study variable of interest and the auxiliary variables. Estimators are then derived on the basis of this model. Previous studies have shown how this can be done to estimate the population total.

However, other literature have also shown the non parametric technique used in increasing the precision of our outcome in estimating the finite population totals. Usually a parametric approach is used to represent the relationship between the auxiliary variables and the study variable. But in some situations, the parametric model is not appropriate, and the resulting estimators do not achieve any efficiency gain over pure estimators. A natural alternative was first proposed by (KUO, 1988) for the distribution function, that adopts a nonparametric approach, which does not place any barriers on the relationship between the auxiliary data and the study variable. Other significant works in this topic are (Chambers, 1986) and(Martanez et al., 2011).

(Johnson et al., 2008) used the traditional local polynomial regression estimator for the unknown regression function $m(x)$. They assumed that $m(x)$ is a smooth function of x and obtained an asymptotically design-unbiased and consistent estimator of the finite population total. The local polynomial regression estimator has the nature of the generalized regression estimator, but is based on a non parametric super population model applicable to a much larger class of functions. (Breidt, 2005) considered a related nonparametric model-assisted regression estimator, replacing local polynomial smoothing with penalized splines. (Johnson et al., 2008) extended the local polynomial non parametric regression estimation to two-stage sampling, in which a probability sample of clusters is selected, and then sub-samples of elements within each selected cluster are obtained. In this paper, we are concerned with the estimation of the finite population total in the presence of one auxiliary variable using the local polynomial regression. General effects of outliers: (Barnett and Moore, 1997) suggested accommodation and transformation as methods of dealing with outliers in a set of data. They explored the use of nonparametric methods in accommodation of outlying observations and further suggested transformations such as the use of square roots or natural logarithms when data points are non-negative to pull outliers into proximity with the rest of the data. Finally they suggested that deletion of outliers may be necessary if they are found to be errors that cannot be corrected.

(Welsh and Ronchetti, 1998) investigated the effects of outliers on a regression line.In their work, a high leverage point that does not conform to the linear relationship between the variables in the question is influential and would considerably change p-values from significance tests. (Osborne and Overbay, 2004) discussed effects of both deterministic and random outliers. His work considered the effects of outliers on sample means and variances. He suggested the use of visual aids, dot plots, scatter plots for identification of outliers before one proceeds with the analysis of a given set of data. He further explored a non-parametric or distribution free approach to detect outliers based on computing medians.(Cellmer, 2014) investigated effects of outliers on mean square curves and variance. (Webster et al., 2006) outlines the effects of outlying observations on regression analysis.

Outlier Robust Estimation: (Cassel et al., 1976) and (Rao et al., 1980) considered the generalized regression estimators which feature great robustness to model misspecification. Aspects of the ratio and local polynomial regression estimators of finite population total considered in this document have been discussed by various researchers. (Cochran, 1946) constructed a modified ratio estimator corrected for bias. (Barnett and Moore, 1997) showed that the ratio estimator makes use of parametrically specified models and that it is applicable as an estimator in a bivariate set of data where the two population characteristics are highly correlated. (Breidt, 2005) considered estimation of finite population totals in the presence of auxiliary information based on the local polynomial regression. Design-based approaches to dealing with outliers in survey estimation have been described by (Searls, 1964). (Chambers, 1986) developed model-based outlier robust techniques for sample surveys.
This research work is using an approximation technique to approximate the finite population total called the Lagrange polynomial that doesn't require any selection of bandwidth as in the case of local polynomial regression estimator. The Lagrange polynomials are used for polynomial interpolation and extrapolation. For a given set of distinct points x_j and numbers y_j, the Lagrange polynomial of lowest degree that assumes at each point x_j the corresponding value y_j(ie the functions coincide at each point).

1.2 Statement of problem

In the recent years, census has become a serious challenge to the growing economy and have received a great deal of interest from various stakeholders including academicians, statisticians and policy makers. There has been considerable amount of research on estimation of finite population totals using different approaches in estimation of finite population totals; however, no single model has performed well without gaps. (Kai et al.,2010) proposed the local composite quantile regression (LCQR) as an alternative to the local polynomial regression (LPR) known to be better than the existing estimators. (Breidt, 2005) also proposed using penalized splines instead of local polynomial smoothing.Which simply means, there is still room for improvement in this area. Prior to any survey taking place, the survey administrator needs to identify the target population to who questions would be asked (Fraenkel J.and Wallen, N.E., 2006). In order to do this, the survey administrator must take into account every one who could possibly be represented by the survey; and must not include anyone who could not be affected by the survey (Foddy, 1993). After the target population has been identified, the mode of data collection is considered (Briggs, 1986); (Lavrakas, 1993) and (Shuy, 2002). This largely depends upon financial resources, human resources, time frame, accessibility, and other issues that can act as catalysts or deterrents to collecting survey responses (Arleck and Settle, 2004); (Fowler, 2002); (Singleton,R and Straits, B. C., 2002). In view of these problems, it is very difficult to have data that can correlate with another data collected from a survey to satisfy the conditions of the ratio estimator and if the sample size collected is very small, that will also have an impact in obtaining a smooth curve along the knots, since the presence of outliers may be limited. That is why, this study has proposed the Lagrange polynomial approximator.

1.3 General objective

The main objective of the study is to approximate finite population totals using Lagrange polynomial.

Objectives of the study

1. To develope a Lagrange polynomial to approximate the finite population total
2. To determine the convergence properties of the Lagrange polynomial approximate of finite population totals.
3. To perform an empirical study using the Lagrange polynomial to extrapolate the population totals at a specified year (2019).

1.4 Significance of the study

Most oftenly, the population structure of the study variable is mostly not known and stakeholders like government departments and other private sectors are always interested in knowing the population total in order to help in budgeting and other macroeconomic engagements. That is why governments spend a lot of money in census. This issue of spending money on counting the entire population provoked the thinking of researchers, politicians and academicians on how to minimize cost and not compromising accuracy. However, this phenomenon brought about the various techniques to be used to solve it, but still none is completely accurate. That is why, this study has proposed an approximation technique known as the Lagrange polynomial to help address this problem.

Theoretical results of the study would improve the estimation techniques of the finite population total. Given the wide application of other estimators used earlier on in estimating the finite population totals, better estimators are still required to help improve the precision with less cost especially for governments around the world in conducting censuses and other related surveys that are allocated by governments and other stakeholders high amount of money. Empirical results of the study would help inform various users of such data on certain macroeconomic plans that are efficient and money saving. This will also help emerging economies to save money that could have been used to conduct surveys.

CHAPTER TWO

2 LITERATURE REVIEW

2.1 Estimators of finite population total

Ratio estimator

(Deville and Sarndal, 1992) in the context of using auxiliary information from survey data to estimate the population total defined $U_1, U_2 ... U_N$ as the set of labels for the finite population. Letting $(y_i,\ x_i)$ be the respective values of the study variable y and the auxiliary variable x attached to i^{th} unit.Of interest is the estimation of population total $Y_t = \sum_{i=1}^{N} y_i$ effectively using the known population totals $X_t = \sum_{i=1}^{N} x_i$ at the estimation stage. If we let $s_1, s_2 ..., s_n$ be the set of sampled units under a general sampling design p, and let $\pi_i = p \quad (i \in s)$ be the first order inclusion probabilities, then the conventional calibration estimator for total Y_t is defined by

$$\hat{Y} = \sum_{i \in s} \frac{y_i}{\pi_i} ....(2.0)$$

In 1946, Cochran made an important contribution to the modern sampling theory by suggesting methods of using the auxiliary information for the purpose of estimation in order to increase the precision of the estimates (Cochran and Goulden, 1940). He developed the ratio estimator to estimate the population mean or the total of the study variable y. The ratio estimator of population $\bar{Y}$ is of the form

$$\bar{y}_r = \frac{\bar{y}}{\bar{x}} \bar{X}; \quad \bar{x} \neq 0 ....(2.1)$$

The aim of this method is to use the ratio of sample means of two characters which would be almost stable under sampling fluctuations and, thus, would provide a better estimate of the true value. It has been well-known fact that $\bar{y}_r$ is most efficient than the sample mean estimator $\bar{y}$, where no auxiliary information is used, if ρ_{yx}, the coefficient of correlation between y and x is greater than half the ratio of coefficient of variation of x to that of y, that is, if

$$\rho_{yx} > \frac{1}{2}(\frac{C_x}{C_y}) ....(2.2)$$

Thus, if the information on an auxiliary variable is either already available or can be obtained at no extra cost and it has a high positive correlation with the main character, one would certainly prefer ratio estimator to develop more and more superior techniques to reduce bias and also to obtain unbiased estimators with greater precision by modifying either the sampling schemes or the estimation procedures or both. (Cochran, 1946) further extended the work of (Madow and Madow, 1944) on systematic sampling. (Searls, 1964) also dealt with the problem of estimation using some a priori-information. Contrary to the situation of ratio estimator, if variables y and x are negatively correlated then the product estimator of population mean $\bar{Y}$ is of the form

$$\bar{y}_q = \frac{\bar{y}}{\bar{X}}\bar{x} ....(2.3)$$

;

$$\bar{X} \neq 0 ....(2.4)$$

This was proposed by (Robson, 1957). It has been observed that the product estimator gives higher precision than the sample mean estimator $\bar{y}$ under the condition that is if

$$\rho_{yx} < -\frac{1}{2}(\frac{C_x}{C_y}) .....(2.5)$$

The expressions for bias and mean square errors of $\bar{y}_r$ and $\bar{y}_q$ have been derived by (Cochran and Goulden, 1940), which are also available in the books by (Sukhatme, 1984).

(Hansen et al., 1953) made use of known value of $\bar{X}$ for defining the difference estimator

$$\bar{y}_d = \bar{y} + \beta(\bar{X} - \bar{x}) ....(2.6)$$

where β is a constant. The best choice of β which minimizes the variance of the estimator is seen to be

$$\beta = \frac{S_{yx}}{S_x^2}........(2.7)$$

which is the population regression coefficient of y on x. Since, β is generally unknown in practice, therefore, it is estimated by sample regression coefficient

$$b = \frac{s_{yx}}{s_x^2}.........(2.8)$$

Using sample regression coefficient (i.e. b),(Watson, 1964) defined simple linear regression estimator as

$$\bar{y}_{1r} = \bar{y} + b(\bar{X} - \bar{x})........(2.9)$$

This estimator is biased, the bias being negligible for large samples.

The most common way of defining a more efficient class of estimators than usual ratio (product) and sample mean estimator is to include one or more unknown parameters in the estimators whose optimum choice is made by minimizing the corresponding mean square error or variance. Sometimes, such modifications or generalizations are made by mixing two or more estimators with unknown weights whose optimum values are then determined which generally depend upon population parameters. In order to propose efficient classes of estimators, (Singh et al., 1994) suggested a one-parameter family of factor-type (F-T) ratio estimators defined as

$$\bar{y}_f = \bar{y}[\frac{(A+C)\bar{X} + fB\bar{x}}{(A+fB)\bar{X} + C\bar{x}}]........(2.10)$$

where A=(d-1)(d-2), B=(d-1)(d-4), C=(d-2)(d-3)(d-4), $d \geq 0$, $f = \frac{n}{N}$. In some situations of practical importance, the information on more than one auxiliary character correlated with the study variable is available. To cope with such situations, (Olkin, 1958) proposed a weighted multivariate ratio estimator. (Srivastava, 1983) extended his work for positive correlation in the population while

(Rao and Mudholkar, 1967), (Singh, 1967) and (Ray and Sahai, 1980) proposed similar estimators using multi-auxiliary characters for negative correlation in the population. For positive correlation in the population,(Sukhatme, 1962) developed multivariate ratio-type estimators.

Another way of using multi-auxiliary information in double sampling is chaining of estimators. If the population mean of the main auxiliary variable is unknown, it may be estimated more efficiently with the help of another auxiliary variable whose population mean is known, using ratio, product or regression type estimators as the case may be. The process was termed as 'chaining' by (Lu and Yan, 2014) who used ratio estimator in the first-phase sample to estimate the population mean of the main auxiliary variable. Later on (Ray and Sahai, 1980), (Mukerjee et al., 1987) extended his work and proposed ratio-in-regression and regression-in-regression chain type estimators. Further, (Singh et al., 1994),(Swain et al., 2013), (Singh and Espejo, 2000), (Singh and Vishwakarma, 2008), (Dash and Mishra, 2011), (Choudhury and Singh, 2012), (Khan et al., 2014) and (Solanki et al., 2014) among others proposed various chain type estimators of population mean of study variable. Successive (rotation) sampling resembles two-phase sampling, hence, there is a greater scope to consider the chain-type estimators in successive sampling over different occasions.

Local polynomial regression estimator

Parametric regression finds the set of parameter estimates that fit the data best for a predetermined family of functions.In many cases, this method yields easily interpretable models that do a good job of explaining the variation in the data. However, the chosen family of functions can be overly-restrictive for some types of data. (Fan and Gijbels, 1996) present examples in which even a 4th-order polynomial fails to give visually satisfying fits. Higher order fits may be attempted, but this leads to numerical instability. An alternative method is desired.

One early method for overcoming these problems was the (Nadaraya, 1964) and (Watson, 1964). To find an estimate for some function, $m(x)$, we take a simple weighted average, where the weighting function is typically a symmetric probability density and is referred to as a Kernel function. (Gasser and Muller, 1984) proposed a similar estimator. Given n observations, (X_i,Y_i)

$$\hat{m}(x) = \sum_{i=1}^{n} Y_i \int_{s_{i-1}}^{s_i} K(u-x)du..........(2.11)$$

where $s_i = (X_i + X_{i+1})/2$, $s_0 = -\infty$ and $s_{n+1} = \infty$. This estimator is able to pick up local features of the data because only points within a neighborhood of x are given positive weight by K. However, the fit is constant over each interval, (s_i,s_{i+1}), and a constant approximation may be insufficient to accurately represent the data. A more dynamic modeling framework is desired.

The concept of nonparametric models within a model assisted framework was first introduced by (Johnson et al., 2008) in estimating population parameters like population total and mean. The estimator was based on local polynomial smoothing. For a population of size N and where values for y are fully observed, they proposed the following estimator for population total of the variable y.

$\hat{Y}_{gen} = \sum_{i \in s}(\frac{y_i - \hat{\mu}(x_i)}{\pi_i}) + \sum_{j=1}^{N} \hat{\mu}(x_j).........(2.12)$

Where j=1,2,...,N and i=1,2,...,n. $\hat{\mu}(x_i)$ were obtained using local polynomial, a kernel nonparametric method. π_i is the inclusion probability into the sample. $\hat{\mu}(x_i)$ is a smooth function of a single variable x. The first term in (2.12) is an adjustment for bias while the second is an estimator of population total. The estimator could also be written as

$$\hat{Y}_{gen} = \sum_{i \in s} \frac{y_i}{\pi_i} + (\sum_{j=1}^{N} \hat{\mu}(x_j) - \sum_{i \in s} \frac{\hat{\mu}(x_i)}{\pi_i}).......(2.13)$$

The first term in (1.13) is a design estimator while the second is model component. Therefore, when the sample comprises of the whole population, the model component reduces to zero since $\pi_i = 1$ and $s = N$. We therefore have the actual population total. (Wu and Sitter, 2001) proposed more complex models and generalized the calibration procedure by means of model calibration. In particular, they considered generalized linear models and nonlinear parametric regression models for the super population model ξ, such that $E_\xi(y_i) = \mu(x_i)$ where $\mu(x_i)$ is a known function of x_i. They proposed model calibration estimator for population total Y_t to be $\tilde{Y} = \sum_{i \in s} \frac{y_i}{\pi_i}$.

In local polynomial regression, a lower-order weighted least squares (WLS) regression is fit at each point of interest, x using data from some neighborhood around x. Following the notation from (Fan and Gijbels, 1996), let the (X_i, Y_i) be ordered pairs such that

$$Y_i = m(X_i) + \sigma(X_i)\epsilon_i, ...(2.14)$$

where ϵ $N(0,1)$, $\sigma^2(X_i)$ is the variance of Y_i at the point X_i, and X_i comes from some distribution, f . In some cases, homoskedastic variance is assumed, so we let $\sigma^2(X) = \sigma^2$. It is typically of interest to estimate $m(x)$. Using Taylor's expansion:

$$m(x) \approx m(x_0) + m'(x_0)(x - x_0) + ... + \frac{m^n(x_0)}{n!}(x - x_0)^n........(2.15)$$

We can estimate these terms using weighted least squares by solving the following for β:

$$\sum_{i=1}^{n}[Y_i - \sum_{j=0}^{q}\beta_j(X_i - x_0)^j]^2 K_h(X_i - x_0)........(2.16)$$

In (1.16), h controls the size of the neighborhood around x_0, and $K_h(.)$ controls the weights, where $K_h(.) \equiv \frac{K(\frac{\cdot}{h})}{h}$, and K is a kernel function. Denote the solution to (2.16) as $\hat{\beta}$. Then estimated $m^v(x_0) = v!\hat{\beta}_v$. It is often simpler to write the weighted least squares problem in matrix form. Therefore, denote X as the design matrix centered at x_0:

$$X = \begin{bmatrix} 1 & x_1 - x_0 & . & . & (x_1 - x_0)^p \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ 1 & x_n - x_0. & .. & . & (x_n - x_0)^p \end{bmatrix} \text{..............................(2.17)}$$

Let W be a diagonal matrix of weights such that

$$W_{jj} = K_h(X_i - x_0)$$

Then the minimization problem

$$argmin_{\beta}(y - X\beta)^T W(y - X\beta)\text{........(2.18)}$$

is equivalent to (2.16), and $\hat{\beta} = (X^T WX)^{-1} X^T Wy$ (Fan and Gijbels, 1996). We can also use this representation to express the conditional mean and variance of $\hat{\beta}$:

$$E(\hat{\beta}|X) = \beta + (X^T WX)^{-1} X^T Ws\text{...........(2.19)}$$

$$var(\hat{\beta|}X) = (X^T WX)^{-1}(X^T \Sigma X)(X^T WX)^{-1}\text{.............(2.20)}$$

where $s = (m(X_1),.....,m(X_2)) - X\beta$ and $\sigma = diag[K_h^2(X_i - x_0)\sigma^2(X_i)]$. There are critical parameters whose choice can have an effect on quality of the fit. These are the bandwidth, h, the order of the local polynomial being fit, p, and the kernel or weight function, K (often denoted K_h to emphasize its dependence on the bandwidth). While we focus mainly on estimation of $m(x)$, many of these results can be used for estimating the r^{th} derivative of $m(x)$ with slight modification.

Horvitz-Thompson(HT) Estimator

This method of estimating the finite population totals doesn't make use of the auxiliary information x_i but instead uses only the study variable y_i to obtain the population totals.

Consider the population of size N with units $y_1, y_2, y_3, ...y_N$. Suppose we want to select sample s of size n_s, define an indicator variable:

$$I_i = \begin{cases} 1 & if\ i \in s \\ 0 & otherwise \end{cases}$$

Let π_i be the probability of including i^{th} unit of the population in sample s. This is called the inclusion probability or first order inclusion probability of i^{th} unit in the sample.

Let π_{ij} be the probability of including i^{th} and j^{th} units in the sample. This is called the joint inclusion probability or second order inclusion probability. The first order inclusion probabilities satisfy

$$\sum_{i=1}^{N} \pi_i = n...............(2.21)$$

where n is the sample size. If the sampling design is a fixed-size sampling design such that

$$V(n) = 0..................(2.22)$$

$$\sum_{i=1}^{N} \pi_{ij} = n\pi_j..................(2.23)$$

Given the sample index set s, define the following indicator function

$$I_i = \begin{cases} 1 & if \quad i \in s \\ 0 & otherwise \end{cases}$$

In this case, I_i is a random variable with $E(I_i) = \pi_i$ and $E(I_i I_j) = \pi_{ij}$. Furthermore, by the definition of sample size n,

$$\sum_{i=1}^{N} I_i = n\text{..................}(2.24)$$

Thus, taking expectations of both sides of (2.24), we can get (2.21). Also, multiplying both sides of (2.24) by π_j and taking expectations again, we get (2.23) When the sample is obtained from a probability sampling design, an unbiased estimator for the total $Y = \sum_{i=1}^{N} y_i$ is given by

$$\hat{Y}_{HT} = \sum_{i=1}^{N} \frac{y_i}{\pi_i} = \sum_{i=1}^{N} y_i \pi_i^{-1}\text{......................}(2.25)$$

This often called Horvitz- Thompson (HT) estimator, which is originally discussed by (Horvitz and Thompson,1952). This method doesn't make use of the auxiliary information x_i but instead uses only the study variable y_i to obtain the population total.

The Horvitz -Thompson estimator, given by (2.25), satisfies the following properties:

$E(\hat{Y}) = Y$ where π_i is the inclusion probability. Next, we observe that

$$E[\hat{Y}_{HT}]E[\sum_{i=1}^{N} \frac{y_i}{\pi_i}] = E[\sum_{i=1}^{N} [\frac{I_i y_i}{\pi_i}]]\text{.......}(2.26)$$

$$\sum_{i=1}^{N}\{\frac{y_i}{\pi_i}E[I_i]\}.................(2.27)$$

$$\sum_{i=1}^{N}\frac{y_i}{\pi_i}\pi_i = \sum_{i=1}^{N} y_i = Total = Y.............(2.28)$$

ie

$$E[\hat{Y}_{HT}] = \sum_{i=1}^{N} y_i = T = Y......................(2.29)$$

$$\hat{Y}_{HT} \text{ is unbiased of T}$$

ie $\hat{Y}_{HT}$ is unbiased under design based approach $(SainiandKumar,\ 2016)$

Variance

$$V(\hat{Y}_{HT}) = \sum_{i=1}^{N}\sum_{j=1}^{N}(\pi_{ij} - \pi_i\pi_j)\frac{y_i y_j}{\pi_i\pi_j}....................(2.30)$$

Similarly, we have

$$V(\hat{Y}_{HT}) = \sum_{i=1}^{N}\sum_{j=1}^{N}\frac{y_i}{\pi_i}\frac{y_j}{\pi_j}(\pi_{ij} - \pi_i\pi_j)............................(2.31)$$

The variance of this estimator can be minimized when $\pi_i \propto y_i$. That is, if the first order inclusion probability is proportional to y_i, the resulting HT estimator under this sampling design will have zero variance. However, in practice, we can't construct such design because we don't know the value of y_i in the design stage. If there is a good auxiliary variable x_i which is believed to be closely related with

y_i, then a sampling design with $\pi_i \propto x_i$ can lead to very efficient sampling design.

2.2 Asymptotic properties

Statisticians in fields such as demography sometimes insist on benchmarking over lots of variables to match the known totals from a census at the risk of worsening the efficiency of the estimations. On the other hand, if complete auxiliary information $x_1, x_2, ...x_N$ is known which is usually the case in most survey problems, a very compelling question to ask would be; What is the best calibration equation to be used in the construction of the calibration estimator?

By noting that it is the relationship between y and x hopefully captured by the working model that determines how well the auxiliary information should be used. Research in the theory of sampling for surveys has been concerned with the development of more efficient sampling systems, the system including both the sample design and the method of estimation. One sampling system is said to be more efficient than the other if the variance or mean square error of the estimate with the first system is less than that of the second, provided the cost of obtaining the data and results is the same for both. The development of stratified, multi-stage, multiphase, cluster, systematic, and other sample designs beyond simple or unrestricted random sampling, as well as alternative methods of estimation, have all resulted in increased efficiency in specific circumstances. As indicated above, the appropriate use of variable probabilities for the selection of the sample elements can lead to gains in efficiency over systems using equal probabilities of selection. It is well know that if samples of size one are with probabilities proportionate to the exact measure of the characteristic under observation, unbiased estimates of means or totals for the population exist which have zero sampling error (Horvitz and Thompson,1952).

One of the most useful and well-known classes of functions mapping the set of real numbers into itself is algebraic polynomials, the set of functions of the form

$$P_n(x) = a_n x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0 ....(2.32)$$

where n is a non negative integer and $a_0......a_n$ are real constants. One reason for their importance is that they uniformly approximate continuous functions. By this we mean that given any function, defined and continuous on a closed and bounded interval, there exists a polynomial that is as - close - to the given function as desired (Burden and Faires, 2001)

Ratio estimator

The literature on survey sampling describes a great variety of techniques of using auxiliary information to obtained more efficient estimators. Ratio, product, and regression methods of estimation are good examples in this context . In the situation where the relation between the study variate Y and the auxiliary variate X is a straight line and passing through the origin, the usual ratio and product estimators have efficiencies equal to the usual regression estimator. But in many practical situations the line does not pass through the origin, and in such circumstances the usual ratio and product estimators do not perform equally well as the regression estimator. Keeping this fact in view, a large number of authors have paid their attention toward the formulation of modified ratio and product estimators using information on an auxiliary variate, for instance, see (Solanki et al., 2014) and (Singh et al., 2016).

Suppose n is large and $MSE(\hat{R}) = Var(\hat{R})$. We assume that $\bar{x}$ and $\bar{X}$ are quite close such that

$$\hat{R} - R = \frac{\bar{y} - R\bar{x}}{\bar{x}} = \frac{\bar{y} - R\bar{x}}{\bar{X}} ....(2.33)$$

so that the bias of $\hat{R}$ becomes quite small.

Local polynomial regression estimator

(Montanari and Ranalli, 2003) proposed to use nonparametric method to obtain $\mu(.)$.We note that any nonparametric method such as kernel methods can be used to recover the fitted values for the non sampled units. Such estimators are however challenging to employ in cases of multiple covariates and when data is sparse. Another challenge is how to incorporate categorical covariates. It is therefore necessary to consider other methods to recover the fitted values such as splines.

The term spline originally referred to a tool used by draftsmen to draw curves. According to (Keele, 2008), splines are piecewise regression functions we constrain to join at points called knots. In their simplest form,splines are regression models with a set of dummy variables on the right hand side of the model that are used to force the regression line to change direction at some point along the range of auxiliary variable x.A higher degree polynomial yields a smoother $\hat{\mu}(.)$ but worsens the boundary variance (Lairez, 2016)

Like local polynomial regression, the analyst must make several modeling decisions with splines. With splines, one must choose the degree of polynomial for the piecewise regression functions, the number of knots and the location of knots, (Breidt, 2005). For some types of splines, the number of knots will control the amount of smoothing, while for other types of splines, a smoothing parameter controls the smoothing (Breidt, 2005).

Piecewise polynomials offer two advantages; First, piecewise polynomial regression functions ensure that the first derivatives are defined at knots which guarantees that the spline estimate will not have sharp corners. A spline with two

knots will be linear and globally smooth since there is only one piecewise function. Increasing the number of knots increases the number of piecewise functions fit to the data allowing for greater flexibility. If one selects a large enough number of knots, the spline model will interpolate between the data points, since more knots shrink the amount of data used for each piecewise function. The number of knots effectively acts as a span parameter for splines. If one uses a small number of knots, the spline estimate will be overly smooth with little variability but may be biased. Using a high number of knots implies little bias but increases variability in the fit and may result in over fitting, (Breidt, 2005) but this can be solved.
The first order linearization is widely used in survey practice, but that in general it is very difficult to evaluate the quality of approximation analytically. Therefore, simulations are presented that show reasonable results (Al-Jirarha and Bataineh, 2014).

The basic idea of a population decomposition is the expression of the y_i in terms of a sum of several components, usually a linear or quadratic function of x_i plus a residual term (Deng and Chhikara, 1990). Given a finite population (y_i,x_i), i=1,2,....,N, we can write the population in terms of a fitted regression line as follows:

$$y_i = A + Bx_i + e_i, i = 1, 2, ...N..................(2.34)$$

where

$$B = \frac{\sum_{i=1}^{N}(x_i - \bar{X})(y_i - \bar{Y})}{\sum_{i=1}^{N}(x_i - \bar{X})^2}...........(2.35)$$

and

$$A = \bar{Y} - B\bar{X}......................(2.36)$$

It is easy to see that the residual e_i satisfies the following

$$\sum_{i=1}^{N} e_i = 0. \quad \sum_{i=1}^{N} e_i x_i = 0................(2.37)$$

(Fan and Gijbels, 1992) established some asymptotic properties for the estimator described in (2.38). In particular, they gave an expression for the conditional variance of the estimator $\hat{m}(x)$ found by minimizing:

$$\sum_{j=1}^{n}(Y_n - \beta_0 - \beta_1(x - X_j))^2 \omega(X_j) K(\frac{x - X_j}{h_n}\omega(X_j))...(2.38)$$

This model is slightly more complex than (2.15), as it allows for a variable bandwidth control by $\omega(X_j)$. Note that the linear (q=1) case of (2.15) is an equivalent to (2.38) when $\omega(X_j) = 1$. The conditional bias and variance are important because they allow us to look at the conditional mean squares errors, which is important for choosing the bandwidth.

The results from (Fan and Gijbels, 1992) are limited to the case where the $X_i's$ are univariate. (Ruppert and Wand, 1994) gave results for multivariate data proposing the following model:

$$Y_i = m(X_i) + \sigma(X_i)\epsilon_i, i = 1, ..., n........(2.39)$$

where $m(x) = E(Y|X = x)$, $x \in \Re^d$, ϵ_i are iid with mean 0 and variance 1, and $\sigma^2(x) = var(Y|X = x) < \infty$. A solution to the problem comes from slightly modifying (2.17).

(Kai et al., 2010) proposed an alternative to local polynomial regression(LPR) in the form of local composite quantile regression(LCQR). While LPR is the best linear smoother, CQR is not a linear estimator, so it may still be an improvement. Indeed, for many common error distribution, this method appears to be more efficient asymptotically than LPR. LCQR can also be applied to derivative estimation.

2.3 Comparative empirical studies

Research literature have revealed that the ratio estimator performs better than the local linear polynomial estimator when the population is linear no matter which variance is used. The local linear polynomial regression estimator becomes a better estimator when the population used is either quadratic or exponential. The relative mean square errors (MSE) increases as the bandwidths increase as well which shows robustness of the local linear polynomial regression estimator when the quadratic Kernel is with a smaller bandwidth. This is also true with an increase in the sample size which increases the likelihood of outliers in the sample, the non parametric estimator performs better (Cochran, 1946).

CHAPTER THREE

3 METHODOLOGY

3.1 Approximation of finite population totals

In this chapter, an approximator is being introduced, that is the Lagrange polynomial approximate of the finite population totals.

Lagrange Polynomial

Consider a finite population $U = \{U_1, U_2, ..., U_N\}$ of N units. Let (y, x) be the (total,year) variables taking non negative real values (y_i,x_i) respectively, on the unit $U_i (i = 1, 2, ...N)$. From the population U, a simple random sample of size n is drawn without replacement. Then, the Lagrange interpolating polynomial is the polynomial $p(x)$ of degree $\leq (n-1)$ that passes through the n points (x_1,y_1=f(x_1)),(x_2,y_2=f(x_2)),....,(x_n,y_n=f(x_n)) and is given by:
$p(x) = \sum_{j=1}^{n} p_j(x)$, where $p_j(x) = y_j \prod_{k=1}^{n} \frac{x-x_k}{x_j-x_k}$ written explicitly,

$$p(x) = \frac{(x-x_2)(x-x_3)...(x-x_n)}{(x_1-x_2)(x_1-x_3)...(x_1-x_n)} y_1 + \frac{(x-x_1)(x-x_3)...(x-x_n)}{(x_2-x_1)(x_2-x_3)...(x_2-x_3)...(x_2-x_n)} y_2 + ...$$
$$+ \frac{(x-x_1)(x-x_2)...(x-x_{n-1})}{(x_n-x_1)(x_n-x_2)...(x_n-x_{n-1})} y_n$$

The problem of determining a polynomial of degree one that passes through the distinct points (x_0, y_0) and (x_1, y_1) is the same as approximating a function f for which $f(x_0) = y_0$ and $f(x_1) = y_1$ by means of a first-degree polynomial interpolating, or agreeing with, the values of f at the given points. Using this polynomial for approximation within the interval given by the endpoints is called polynomial interpolation. Also, using this polynomial for approximation outside the interval given by the endpoints is called polynomial extrapolation. On the other hand, extrapolation is the process of estimating a value of $f(x)$ that lies outside the range of the known base points, $x_1, x_2, x_3, ..., x_n$ (Burden and Faires, 2001).

Define the functions

$$L_0 = \frac{x - x_1}{x_0 - x_1}$$

and

$$L_1 = \frac{x - x_0}{x_1 - x_0}$$

The linear Lagrange interpolating polynomial through (x_0,y_0) and (x_1,y_1) is

$$p(x) = L_0(x)f(x_0) + L_1(x)f(x_1) = \frac{x - x_1}{x_0 - x_1}f(x_0) + \frac{x - x_0}{x_1 - x_0}f(x_1)$$

Note that

$$L_0(x_0) = 1, L_0(x_1) = 0, L_1(x_0) = 0, \quad and \quad L_1(x_1) = 1,$$

which implies that

$$p(x_0) = 1.f(x_0) + 0.f(x_1) = f(x_0) = y_0$$

and

$$p(x_1) = 0.f(x_0) + 1.f(x_1) = f(x_1) = y_1$$

So p is the unique polynomial of degree at most one that passes through (x_0, y_0) *and* (x_1, y_1).

3.1 **Theorem:** (Burden and Faires, 2001)

If $x_0, x_1, ..., x_n$ are $n+1$ distinct numbers and f is a function whose values are given at these numbers, then a unique polynomial $p(x)$ of degree at most n exists with

$$f(x_k) = p(x_k), for \quad each \quad k = 0, 1, ..., n,$$

This polynomial is given by

$$p(x) = f(x_0)L_{n,0}(x) + ..., +f(xn)L_{n,n}(x) = \sum_{k=0}^{n} f(x_k)L_{n,k}(x).....(3.11)$$

where, for each k=0,1,.....,n

$$L_{n,k}(x) = \frac{(x-x_0)(x-x_1)(x-x_2).....(x-x_{k-1})(x-x_{k+1})....(x-x_n)}{(x_k-x_0)(x_k-x_1)(x_k-x_2).....(x_k-x_{k-1})(x_k-x_{k+1})....(x_k-x_n)}$$

We will write $L_{n,k}(x)$ simply as $L_k(x)$.

3.2 **Theorem:** (Burden and Faires, 2001)

Suppose $x_0, x_1, ..., x_n$ are distinct numbers in the interval $[a, b]$ and $f \in C^{n+1}[a, b]$. Then, for each x in $[a, b]$, a number $\zeta(x)$(generally unknown) between $x_0, x_1, ..., x_n$, and hence in (a, b), exists with

$$f(x) = p(x) + \frac{f^{n+1}(\zeta(x))}{(n+1)!}(x-x_0)(x-x_1)....(x-x_n).......(3.12)$$

where $p(x)$ is the interpolating polynomial given in (3.11).

Proof of theorem 3.2: Note first that if $x = x_k$, for any $k = 0, 1, ..., n$, then $f(x_k) = p(x_k)$, and choosing $\zeta(x_k)$ arbitrarily in (a,b) yields (3.12)

If $x \neq x_k$, for all $k = 0, 1, ..., n$, define the function g for t in $[a, b]$ by

$$g(t) = f(t) - p(t) - [f(x) - p(x)]\frac{(t - x_0)(t - x_1).....(t - x_n)}{(x - x_0)(x - x1).....(x - x_n)}$$

$$g(t) = f(t) - p(t) - [f(x) - p(x)]\prod_{i=0}^{n}\frac{(x_k - x_i)}{(x - x_i)} = [f(x) - p(x)].0 = 0$$

.

Moreover,

$$g(x) = f(x) - p(x) - [f(x) - p(x)]\prod_{i=0}^{n}\frac{(x - x_i)}{(x - x_i)} = f(x) - p(x) - [f(x) - p(x)] = 0$$

.

Thus $g \in C^{n+1}[a, b]$, and g is zero at the $n+2$ distinct numbers $x_0, x_1, ..., x_n$. By generalized Rolle's theorem, there exists a number ζ in (a, b) for which $g^{n+1}(\zeta) = 0$.

So

$$0 = g^{(n+1)}(\zeta) = f^{(n+1)}(\zeta) - p^{n+1}(\zeta) - [f(x) - p(x)]\frac{d^{n+1}}{dt^{n+1}}[\prod_{i=0}^{n}\frac{(t - x_i)}{(x - x_i)}]_{t-\zeta} \quad ......(3.13)$$

However, $p(x)$ is a polynomial of degree at most n, so the $(n+1)st$ derivative, $p^{(n+1)}(x)$, is identically zero. Also $\prod_{i=0}^{n}\frac{(t-x_i)}{(x-x_i)}$ is a polynomial of degree $(n+1)$, so

$$\prod_{i=1}^{n}\frac{(t - x_i)}{(x - x_i)} = [\frac{1}{\prod_{i=0}^{n}}(x - x_i)]t^{n+1} + (lower \ \ degree \ \ terms \ \ in \ \ t)$$

and

$$\frac{d^{n+1}}{dt^{n+1}}\prod_{i=0}^{n}(\frac{t - x_i}{x - x_i}) = \frac{(n+1)!}{\prod_{i=0}^{n}(x - x_i)}$$

. Equation (3.13) now becomes

$$0 = f^{(n+1)}(\zeta) - 0 - [f(x) - p(x)]\frac{(n+1)!}{\prod_{i=0}^{n}(x - x_i)}$$

and upon solving for f(x), we have

$$f(x) = p(x) + \frac{^{(n+1)(\zeta)}}{(n+1)!}\prod_{i=0}^{n}(x - x_i)$$

The Lagrange polynomial of degree n uses information at the distinct numbers $x_0, x_1, x_2, ..., x_n$ and, in place of $(x - x_0)^n$, its error formula uses a product of the n+1 terms $(x - x_0)(x - x_1)....., (x - x_n)$:

$$\frac{f^{(n+1)}(\zeta)}{(n+1)!}(x - x_0)(x - x_1).....(x - x_n)$$

Convergence properties of polynomial approximations

Let $x_1, x_2, ..., x_n$ be different real numbers and let $f_1, f_2, ..., f_n$ be the corresponding values of a function f. The approximating polynomial is denoted by p. The whole of the methods for polynomial approximations and interpolations can be put into 4 groups but only 3 are considered in this study as follows:

1. The method of Lagrange interpolation determines the unique polynomial p of least degree, which has the values $f_1, ..., f_n$ at the points $x_1, ..., x_n$, i.e., for which $p(x_i) = f_i$, $1 \leq i \leq n$. We shall see later that, in some tricky cases, Lagrange interpolation polynomials p based on n interpolation points will not converge to a continuous target function f as n increases, despite the Weierstrass theorem which states that, "every bounded sequence has a convergent subsequence". He further stated that, "if f is continuous on the unit interval, then there exists a

sequence of polynomials P_n that converges uniformly to f". That is

$$\max_{0\leq x\leq 1} |P_n(x) - f(x)| \longrightarrow 0 \;\; as \;\; n \longrightarrow \infty$$

In many practical applications we wish to approximate the function f by a polynomial P_m of degree m much less than n, the number of data points. In this case the problem is as follows: find the polynomial or polynomials of degree m that are "closest" to the function values at the given distinct points. The construction of the required approximating polynomial depends on the selected measure of distance between given functions and their approximating polynomials. Points 2 and 3 below give the most popular distance measures in polynomial approximation applications.

2. Let us define the distance between the given function and an m^{th}-degree polynomial P_m by the discrete uniform distance formula

$$\max_{1\leq i\leq n} |f_i - P_m(x_i)| ....(3.14)$$

In this case, we choose the polynomial of given degree m for which the quantity (3.14) has a minimum. Such a polynomial is called a best-approximating polynomial (BAP) of degree m based on the n points $x_1, ..., x_n$.

3. In the least squares method we measure the distance between the given function and polynomial P_m by the value of

$$[\sum_{i=1}^{n}[f_i - P_m(x_i)]^2]^{\frac{1}{2}} ....(3.15)$$

and the least squares approximation of the function is defined by the polynomial of degree m for which the value (3.15) has a minimum.

One of the most useful and well-known classes of functions mapping the set of real numbers into itself is the algebraic polynomials, the set of functions of the form

$$P_n(x) = a_n x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0$$

, where n is a nonnegative integer and $a_0, ..., a_n$ are real constants. One reason for their importance is that they uniformly approximate continuous functions. By this we mean that given any function, defined and continuous on a closed and bounded interval, there exists a polynomial that is as "close" to the given function as desired. This result is expressed precisely in the Weierstrass approximation Theorem as shown below. The simplest polynomial is a constant, i.e.

$$P(x) = K = f(0)$$

3.3 **Intermediate Value Theorem for derivatives:** (Burden and Faires, 2001) Let f be differentiable on $[a, b]$, and let K be a number between $f'(a)$ and $f'(b)$. Then there is a $c \in (a, b)$ such that $f'(c) = K$.

Proof of theorem 3.3:

Suppose $f'(a) < f'(b)$. Then $f'(a) < K < f'(b)$. Let $g(x) = f(x) - Kx$. Then $g' = f'(x) - K$. So $g'(a) = f'(a) - K < 0$ and $g'(b) = f'(b) - K > 0$. Since g is continuous on a closed interval, it must have a minimum on that interval. Then minimum cannot be at either endpoint. So the minimum has to be at some point $c \in (a, b)$, so $g'(c) = 0$. Then $f'(c) - K = 0$, so $f'(c) = K$

3.4 **Rolle's Theorem:** (Burden and Faires, 2001)

If f is continuous on the interval $[a, b]$ and differentiable on (a, b) and if $f(a) = f(b)$, then there is a $c \in (a, b)$ such that $f'(c) = 0$.

Proof of theorem 3.4:

Since f is continuous on $[a, b]$ it must have both a minimum and a minimum on $[a, b]$. If both the maximum and the minimum are at the endpoints, then the function must be constant. Then the derivative would be 0 everywhere. If the maximum is not at an endpoint, but at an interval point c, then $f'(c) = 0$.

If the minumum is not at an endpoint, but at an interior point c, then $f'(c) = 0$. In every case, $f'(c) = 0$ at one point (or more)

3.5 **Cauchy Mean Value Theorem (CMVT):** (Burden and Faires, 2001)

Let f and g be functions continuous on $[a, b]$ and differentable on (a, b). Then there is a $c \in (a, b)$ such that

$$[f(b) - f(a)]g'(c) = [g(b) - g(a)]f'(c)$$

What if $g(x) = x$?

$$[f(b) - f(a)]1 = [b - a]f'(c),$$

so

$$\frac{f(b) - f(a)}{b - a} = f'(c)$$

Written in fraction form:

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)}$$

But we can only write it this way if we know that the denominators are not zero.

Proof of theorem 3.5:

Let $h(x) = [f(b) - f(a)]g(x) - [g(b) - g(a)]f(x)$. This h is continuous on $[a, b]$ and differentiable (a, b)

$h(a) = [f(b) - f(a)]g(a) - [g(b) - g(a)]f(a) = f(b)g(a) - g(b)f(a)$

$h(b) = [f(b) - f(a)]g(b) - [g(b) - g(a)]f(b) = -f(a)g(b) + g(a)f(b)$

So $h(a) = h(b)$

So by Rolle's theorem, there is a $c \in (a, b)$ such that $h'(c) = 0$

$h'(x) = [f(b) - f(a)]g'(x) - [g(b) - g(a)]f'(x)$

$h'(c) = [f(b) - f(a)]g'(c) - [g(b) - g(a)]f'(c) = 0$

$$[f(b) - f(a)]g'(c) = [g(b) - g(a)]f'(c)$$

$$\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(c)}{g'(c)}$$

or

$$\frac{f(b)-f(a)}{b-a} = f'(c)$$

Here, we assume f is once differentiable. The mean value theorem is a special case of Taylor's theorem. If we assume f to be $n+1$ times differentiable we get the Taylor's theorem shown below:

3.6 **Taylor's Theorem:** (Burden and Faires, 2001)

Suppose $f \in C^n[a,b]$, $f^{(n+1)}$ exists on [a,b], and $x_0 \in [a,b]$. For every $x \in [a,b]$, there exists a number $\eta(x)$ between x_0 and x with

$$f(x) = p_n(x) = f(x_0) + f\prime(x_0)(x-x_0) + \frac{f\prime\prime(x_0)}{2!}(x-x_0)^2 + ... + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$$

$$= \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k$$

and

$$R_n(x) = \frac{f^{(n+1)}(\eta(x))}{(n+1)!}(x-x_0)^{n+1}$$

.

$p_n(x)$ is called the n^{th} Taylor polynomial for f about x_0, and $R_n(x)$ is referred to as the remainder term (or truncation error) related to $p_n(x)$. Since the number $\eta(x)$ in the truncation error $R_n(x)$ relies on the value of x at which the polynomial $p_n(x)$ is being evaluated, it is a function of the variable x. However, we should not expect to be able to explicitly determine the function $\eta(x)$. Taylor's Theorem simply ensures that such a function exists and that its value lies between x and x_0. In fact, one of the common problems in numerical methods is to try to determine a realistic bound for the value of $f^{(n+1)}(\eta(x))$ when x is in some specified interval. The infinite series obtained by obtaining the limit of $p_n(x)$ as n tends to infinity is called the Taylor series for f about x_0. Where $x_0 = 0$, the Taylor polynomial is referred to as a Maclaurin polynomial, and the Taylor series is called a Maclaurin series.

3.7 **Polynomial Approximation of Functions:**

(1) Within a given neighbourhood of a value

Note that $p(x) = K = f(0)$ is the simplest polynomial. Therefore, $p(0) = f(0)$. Also, $p'(0) = f'(0)$ Suppose we have a linear polynomial as

$$P(x) = f(0) + f'(0)x.........(3.16)$$

substituting $x = 0$ we have

$$P(0) = f(0)$$

If we find the derivative of (3.16) we have

$$P'(x) = 0 + f'(0)$$

We do it with a quadratic equation

$$p(x) = f(0) + f'(0)x + f''(0)\frac{x^2}{2}........(3.17)$$

At x=0 we have

$$p(0) = f(0) + 0 + 0$$

If we find the derivative of (3.17) and substitute $x = 0$ we have

$$p'(x) = 0 + f'(0) + f''(0)x$$

$$P'(0) = 0 + f'(0) + 0$$

Similarly, we do it for cubic equation

$$p(x) = f(x) + f'(x)x + \frac{f''(x)x^2}{2} + \frac{f'''(x)x^3}{3!}........(3.18)$$

If we find the second derivative of (3.18) and substitute $x = 0$ we will have

$$p''(x) = f''(x) + f'''(x)x$$

$$P''(0) = f''(0) + 0 = f''(0)$$

Therefore,

$$p(x) = f(0) + f'(0)x + \frac{f''(0)x^2}{2} + \frac{f'''(0)x^3}{3!} + ...$$

In general;

$$p(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)x^n}{n!}$$

This is called the Maclaurin series which is a special case of the Taylor series. Suppose we want to approximate $f(x)$ using a polynomial at the point c, we have

$$p(x) = f(c) + f'(c)(x-c) + \frac{f''(c)}{2!}(x-c)^2 + \frac{f'''(c)}{3!}(x-c)^3 + ..., ...(3.19)$$

If $x = c$ we have

$$p(c) = f(c)$$

, then

$$p'(c) = f'(c)$$

Therefore, finding the derivative of (3.19) we have

$$p'(x) = f'(c) + f''(c)(x-c) + \frac{f'''(c)}{2!}(x-c)^2$$

The advantage of Taylor's polynomial is that it will be a good approximation in a particular neighbourhood of a value and not in the whole interval. This takes us to Karl Weierstrass Theorem.

(2)Within the whole interval

3.8 **Weierstrass Theorem:** (Burden and Faires, 2001)

$f : [a, b] \longrightarrow \Re$ continuous

Then there exists a sequence of polynomials $P_n(x)$ such that $||f - P_n||_\infty = max_{x\in[a,b]}|f(x) - P_n(x)| \longrightarrow 0$ as $n \longrightarrow \infty$

Proof of theorem 3.8:

$f : [a, b] = [0, 1] \longrightarrow \Re$ continuous.

$$P_n(x) = B_n(f)(x) = \sum_{k=0}^{n} (\frac{n!}{k!(n-k)!}) f(\frac{k}{n}) x^k (1-x)^{n-k}$$

(Bernstein Polynomial)

$$||f - P_n||_\infty \longrightarrow 0 \quad as \quad n \longrightarrow \infty$$

We are going to consider three functions: $f(x) = 1$, $f(x) = x$ and $f(x) = x^2$ and show convergence.

$$B_n(f)(x) = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} f(\frac{k}{n}) x^k (1-k)^{n-k}$$

$f(x) = 1$

$$B_n(f)(x) = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} x^k (1-k)^{n-k}$$
$$= (x + 1 - x)^n = 1, n \geq 0$$

Hence

$$||f - B_n(f)||_\infty = 0$$

Also,

$f(x) = x$

$$B_n(f)(x) = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} \frac{k}{n} x^k (1-k)^{n-k}$$
$$= \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-k)!} x^k (1-x)^{n-k}$$

Let k-1=L

$$= x \sum_{L=0}^{n-1} \frac{(n-1)!}{L!(n-1-L)!} x^L (1-x)^{n-1-L}$$
$$= x\{x + 1 - x\}^{n-1} = x, n \geq 1$$

Hence

$$||B_n(f) - f||_\infty = 0, n \geq 1$$

$f(x) = x^2$

$$B_n(f)(x) = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} \frac{k^2}{n^2} x^k (1-k)^{n-k}$$
$$= \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-k)!} \frac{k-1+1}{n} x^k (1-x)^{n-k}$$
$$= \sum_{k=2}^{n} \frac{(n-1)!}{(k-2)!(n-k)!} \frac{1}{n} x^k (1-x)^{n-k} + \frac{1}{n} \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-k)!} x^k (1-x)^{n-k}$$
$$B_n(f)(x) = \frac{(n-1)}{n} x^2 \sum_{k=2}^{n} \frac{(n-2)!}{(k-2)!(n-k)!} x^{k-2} (1-x)^{n-k} + \frac{1}{n} x (x+1-x)^{n-1}$$
$$B_n(f)(x) = \frac{n-1}{n} x^2 + \frac{1}{n} x$$
$$= x^2 + \frac{1}{n} x(1-x)$$
$$|f(x) - B_n(f)(x)| = \frac{1}{n} |x(1-x)|, \ for \ \ n \geq 2$$
$$||f - B_n(f)||_\infty = \frac{1}{4n} \longrightarrow 0 \ \ as \ \ n \longrightarrow \infty$$

In order to obtain a best approximating polynomial that has less error, one needs to choose a linear interpolating points that is closest to the target point

CHAPTER FOUR

4 RESULTS AND DISCUSSIONS

4.1 Results

In this chapter, data from the Kenya National Bureau of Statistics (KNBS) was obtained from 1969 to 2009 to carryout the empirical analysis. A simple random sampling technique without replacement (SRSWOR) was used to obtain a sample size of two as clearly stated by the Karl Weierstrass theorem in order to fit in polynomials to approximate the function (i.e. the population trend line in green). However, this investigative approach will reveal the Best Approximating Polynomial (BAP) that can be used to extrapolate the population total in 2019 census as shown below:

Data Exploration

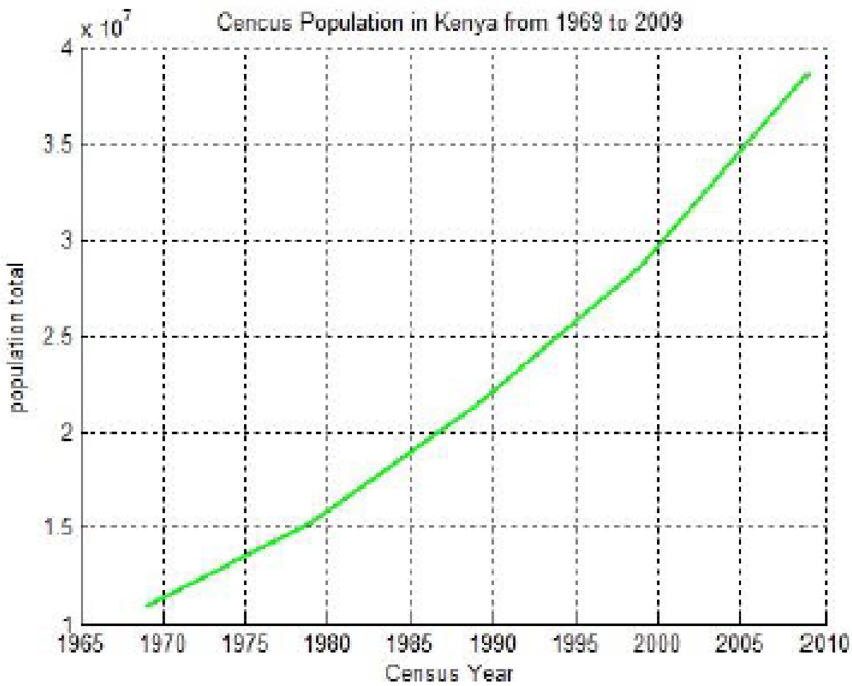

Figure 1: Kenya population Census data from 1969 to 2009.

The Kenya population census data since 1969 to 2009 were plotted in Figure 1 above to see the behaviour of the data as soon shown. The plot showed an upward growth in the population of Kenya.

However, we aimed at selecting a sample size of two from 1969 to 2009 population census using a technique of simple random sampling without replacement making a sample total of ten. A pair of linear samples selected were plotted on the same charts to approximate the function $f(x)$ in green colour as shown below for each.

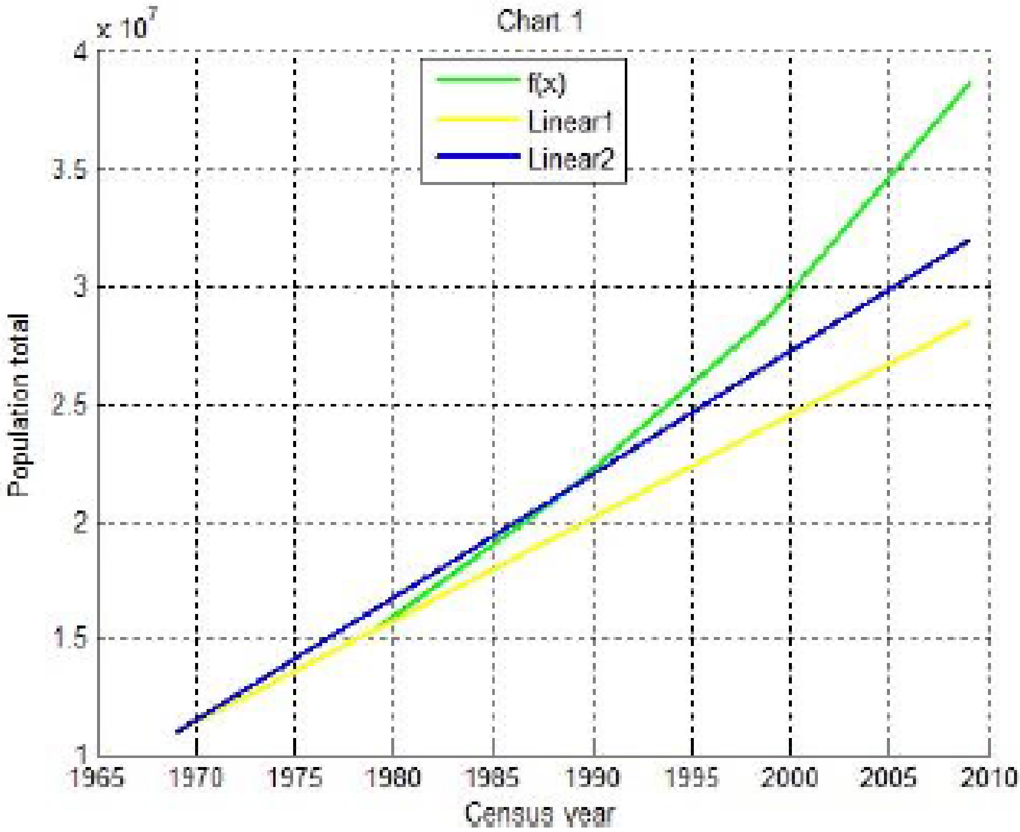

Figure 2: Polynomials in [1969,1979] and [1969,1989] approximating the function.

This chart was obtained from a set of data ranging from [1969,1979] in yellow to [1969,1989] in blue shown in Figure 2 above in order to give a better approximate to the population total in 2019. As can be seen, the two linear plots are not showing any better approximate of the function $f(x)$ in green in order to help us extrapolate the population total in 2019.

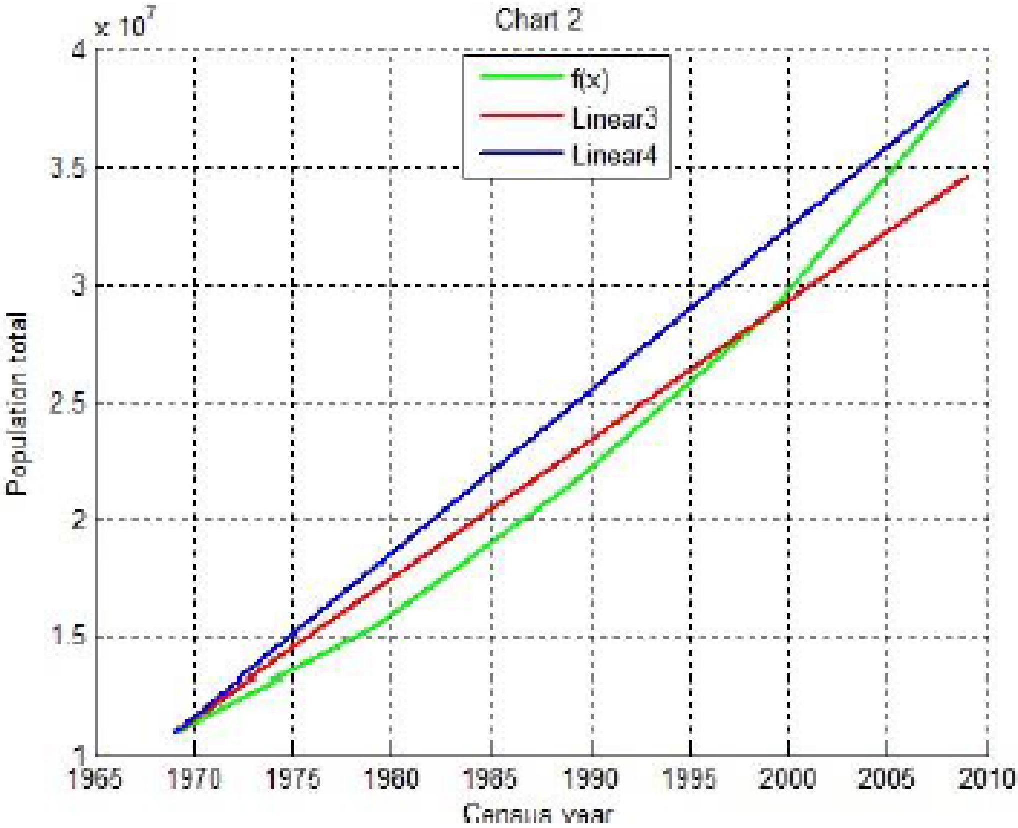

Figure 3: Polynomials in [1969,1999] and [1969,2009] approximating the function.

This chart shown in Figure 3 was obtained from a set of data ranging from [1969,1999] in red to [1969,2009] in blue which was used to approximate the function $f(x)$ in green so as to help us extrapolate the population total in 2019. This was clearly seen to have obtained high variation in the approximation. The blue line appeared to be better than the red at the end point.

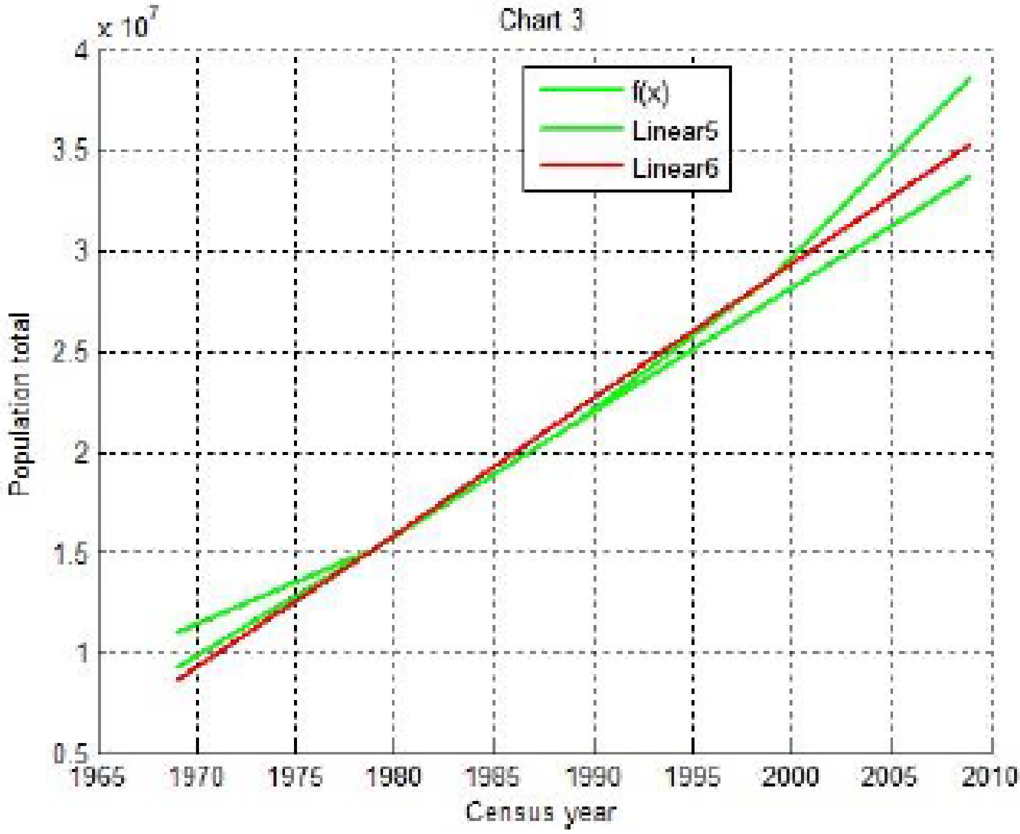

Figure 4: Polynomials in [1979,1989] and [1979,1999] approximating the function.

This chart shown in Figure 4 was obtained from a set of data ranging from [1979,1989] in green dotted line to [1979,1999] in red as a way to help us approximate better the function $f(x)$ in green. Unfortunately, the two approximating lines are not suitable to help extrapolate the population total in 2019.

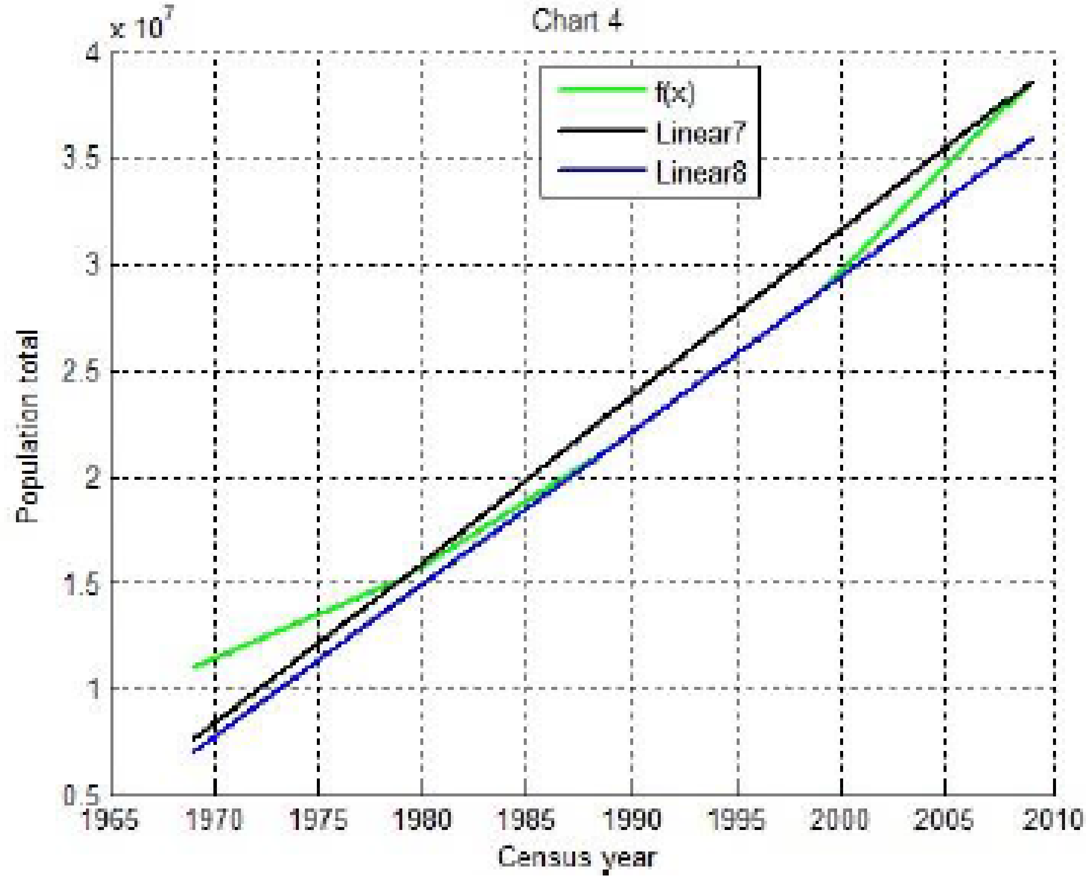

Figure 5: Polynomials in [1979,2009] and [1989,1999] approximating the function.

The chart in Figure 5 above was obtained from a set of data ranging from [1979,2009] in black to [1989,1999] in blue to help us approximate the function $f(x)$ in green representing the trend of the entire population. As seen on the chart, the black line appeared to perform better at the end point than the blue but showed some variations.

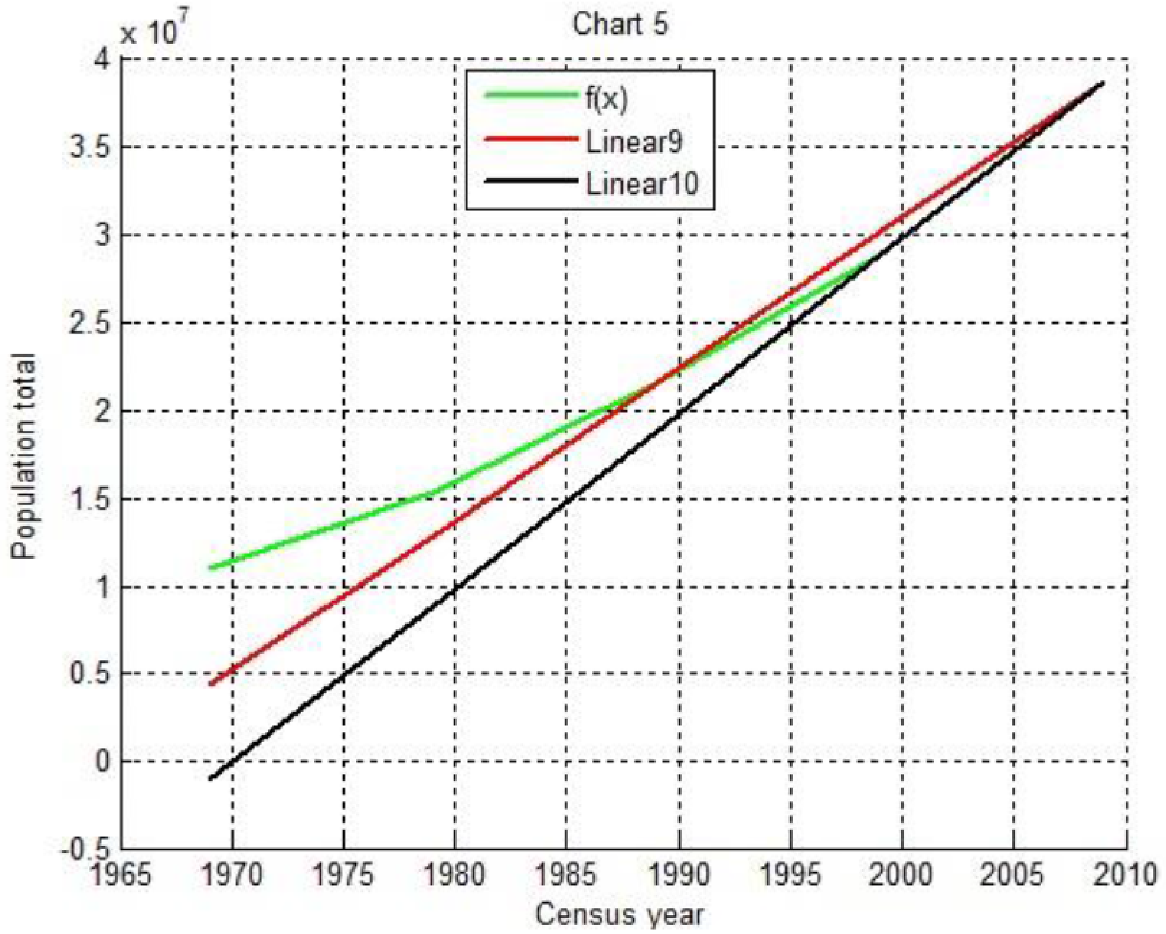

Figure 6: Polynomials in [1989,2009] and [1999,2009] approximating the function.

The chart in Figure 6 above was obtained from a set of data ranging from [1989,2009] in red dotted line to [1999,2009] in black dotted line as an approach to help us approximate the function $f(x)$ representing the total population trend per each year. The chart has clearly shown that, the black dotted line depicted the best approximate on its entire interval which is [1999,2009] as the place for the Best Approximating Polynomial (BAP) to approximate the function $f(x)$ uniformly to any degree of accuracy.

Calculating missing values via interpolation

x[1]=[1999] and y[1]=[28,686,607] x[11]=[2009] and y[11]=[38,610,097]

Columns 1 through 8

28686607 29678956 30671305 31663654 32656003 33648352 34640701 35633050

Columns 9 through 11

36625399 37617748 38610097

$$y[i] = y[i-1] + (y[11] - y[i-1])/h$$

where $i \geq 2$ and h= annual step size

Approximation of population total in 2009

$$x[11] = [2009] \quad and \quad y[11] = [38,610,097] \quad given$$

$$x[10] = [2008] \quad and \quad y[10] = [37,617,748] \quad approximated$$

$$x[9] = [2007] \quad and \quad y[9] = [36,625,399] \quad approximated$$

$$L9 = (x[11] - x[10])/(x[9] - x[10]) * y[9]$$

$$L10 = (x[11] - x[9])/(x[10] - x[9]) * y[10]$$

Approximated value=L9+L10

Approximated population total=38,610,097

Error=0

Extrapolation of 2019 population total

$$x[11] = [2009] and y[11] = [38,610,097]$$

$$x[10] = [2008] and y[10] = [37, 617, 748]$$

$$L19 = (2019 - x[11])/(x[10] - x[11]) * y[10]$$

$$L20 = (2019 - x[10])/(x[11] - x[10]) * y[11]$$

Approximated value=L19+L20

Approximated population total=48,533,587

4.2 Discussions

Several authors documented criteria that assess the quality of the a model or technique. These criteria are based on the difference between the estimated model and the presumed known theoretical model. In the present study, the criterrion used compares to new observations resulting from the same population as individuals of the sample, the variability of the errors of predictions are carried out by a linear polynomial on the other hand these predictions are equal to the improvement of the quality of approximation by taking into account the degree of the polynomial. It also informs about the validity limits of the degree of the approximation polynomial. The empirical analysis demonstrated graphically in chapter four from figure one to six showed the approximation of the function (i.e. the population trend) with a set of two polynomials. The pictorial representations showed that, figure six portrayed a more accurate polynomial depicted by the linear polynomial in black line, formed by 1999 and 2009. However, annual population totals were obtained from 1999 to 2009. The Lagrange polynomial was then used to interpolate through 2007,2008 and 2009 with their corresponding population totals as 36,625,399; 37,617,748 and 38,610,097 respectively. The linear polynomial obtained from 'best fitting' was used to extrapolate the 2019 population total in Kenya as 48,533,587.

CHAPTER FIVE

5 CONCLUSIONS AND RECOMMENDATIONS

5.1 Conlusions

In this work, the Lagrange polynomial has proven to be a good technique in approximating the population total from data obtained from the Kenya National Bureau of Statistics (KNBS).
The research revealed that, subsequent population totals can better be approximated using a sample closest to the target population being approximated. Therefore, the best approximating polynomial must be a linear form in order to obtain convergence with a diminishing variation in a given interval.

5.2 Recommendations

Since estimation techniques are faced with a trade off between bias and variance. Regression models are not exceptions in this problem hence exposed to the effect of bias or variance problem. Thus, the researcher is faced with the choice of compromising with one in order to minimise the other. However, this research is recommending the use of quantile regressions instead of the usual regression techniques to narrow down these problems and avoid the effect of ill-conditions faced in regression analysis, since population is a function of time and considered to be nonlinear. However, ill-conditions are as a result of weak regression coefficients and this can be removed with the use of QR decomposition in order to improve the precision of the estimates.

References

Al-Jorarha, J. and Bataineh, A. (2014). Estimation of finite population ratio when other auxiliary variables are available in the study. *Austrian Journal of Statistics*, 44(1):33.

Arleck, P. and Settle,R. (2004). The survey research handbook. *New York:McGram-Hill Irwin.*

Barnett, V. and Moore, K. (1997). Best linear unbiased estimates in ranked-set sampling with particular reference to imperfect ordering. *Journal of Applied Statistics*, 24(6):697–710.

Breidt, F. J. (2005). Model-assisted estimation for complex surveys using penalised splines. *Biometrika*, 92(4):831–846.

Briggs, C. (1986). Learning how to ask: A sociolinguistic appraisal of the role of the interview in social science research. *Cambridge: Cambridge university.*

Burden, R. L. and Faires, J. D. (2001). Numerical analysis. 2001. *Brooks/Cole, USA.*

Cassel, C. M., Sarndal, C. E., and Wretman, J. H. (1976). Some results on generalized difference estimation and generalized regression estimation for finite populations. *Biometrika*, 63(3):615.

Cellmer, S. (2014). Least fourth powers: optimisation method favouring outliers. *Survey Review*, 47(345):411–417.

Chambers, R. L. (1986). Outlier robust finite population estimation. *JASA*, 81(396):1063–1069.

Choudhury, S. and Singh, B. K. (2012). A class of chain ratio product type estimators with two auxiliary variables under double sampling scheme. *Journal of the Korean Statistical Society*, 41(2):247–256.

Cochran, W. G. (1946). Graduate training in statistics. *The American Mathematical Monthly*, 53(4):193.

Cochran, W. G. and Goulden, C. H. (1940). Methods of statistical analysis. *Journal of the Royal Statistical Society*, 103(2):250.

Dash, P. R. and Mishra, G. (2011). An improved class of estimators in two-phase sampling using two auxiliary variables. *Communications in Statistics - Theory and Methods*, 40(24):4347–4352.

Deng, L.Y. and Chhikara, R. S. (1990). On the ratio and regression estimation in finite population sampling. *The American Statistician*, 44(4):282.

Deville, J.C. and Sarndal, C.E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association*, 87(418):376.

Fan and Gijbels (1992). Variable bandwidth and local linear regression smoothers. *Annals of Statistics*, 20(4):2008–2036.

Fan and Gijbels (1996). Local polynomial modelling and its applications. *Chapman and Hall*.

Foddy, W. (1993). Constructing questions for interviews and questionnaires: Theory and practice. *Melbourne: Cambridge university*.

Fowler, F. (2002). Survey research methods. *Thousand Oaks, CA: Sage*.

Fraenkel, J. and Wallen,N.E. (2006). How to design and evaluate research in education. *New York:McGram-Hill*.

Gasser, T. and Muller, H.G. (1984). Estimating regression functions and their derivatives by the kernel method. *Scandinavian Journal of statistics*, 11(3):171–185.

Hansen, M. H., Hurwitz, W. N., and Madow, W. G. (1953). *Sample survey methods and theory*, volume 1. Wiley New York.

Horvitz, D. G., and Thompson, D. J. (1952). . A generalization of sampling without replacement from a finite . *Journal of American Statistics Association*, 47:663–685

Johnson, A. A., Breidt, F. J., and Opsomer, J. D. (2008). Estimating distribution functions from survey data using nonparametric regression. *Journal of Statistical Theory and Practice*, 2(3):419–431.

Kai, B., Li, and Zou, H. (2010). Local composite quantile regression smoothing:an efficient and safe alternative to local polynomial regression. *Journal of the Royal statistical society*, Series B-statistical methodology 72(part 1):49–69.

Keele, L. J. (2008). *Semiparametric regression for the social sciences.* John Wiley and Sons.

Khan, M., Ullah, S., Al-Hossain, A. Y., and Bashir, N. (2014). Improved ratio-type estimators using maximum and minimum values under simple random sampling scheme. *Hacettepe Journal of Mathematics and Statistics*, 44(29):1–1.

KUO, L. (1988). Linear bayes estimators of the potency curve in bioassay. *Biometrika*, 75(1):91–96.

Lairez, P. (2016). A deterministic algorithm to compute approximate roots of polynomial systems in polynomial average time. *Foundations of Computational Mathematics.*

Lavrakas, P. (1993). Telephone survey methods: Sampling ,selection, and supervision(2nd ed.). *Newbury Park, CA: Sage.*

Lu, J. and Yan, Z. (2014). A class of ratio estimators of a finite population mean using two auxiliary variables. *PLoS ONE*, 9(2):e89538.

Madow, W. G. and Madow, L. H. (1944). On the theory of systematic sampling, i. *The Annals of Mathematical Statistics*, 15(1):1–24.

Martanez, S., Rueda, M., Arcos, A., Martanez, H., and Muaoz, J. F. (2011). On determining the calibration equations to construct model-calibration estimators of the distribution function. *Revista Matematica Complutense*, 25(1):87–95.

Montanari, G. E. and Ranalli, M. G. (2003). On calibration methods for design based finite population inferences. *Bulletin of the International Statistical Institute, 54 th session*, 60.

Mukerjee, R., Rao, T. J., and Vijayan, K. (1987). Regression type estimators using multiple auxiliary information. *Australian Journal of Statistics*, 29(3):244–254.

Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability and Its Applications*, 9(1):141–142.

Olkin, I. (1958). Multivariate ratio estimation for finite populations. *Biometrika*, 45(1/2):154.

Osborne, J. W. and Overbay, A. (2004). The power of outliers (and why researchers should always check for them). *Practical assessment, research and evaluation*, 9(6):1–12.

Rao, P. S. R. S., Cassel, C.-M., Sarndal, C.-E., and Wretman, J. H. (1980). Foundations of inference in survey sampling. *Technometrics*, 22(2):279.

Rao, P. S. R. S. and Mudholkar, G. S. (1967). Generalized multivariate estimator for the mean of finite populations. *Journal of the American Statistical Association*, 62(319):1009–1012.

Ray, S. K. and Sahai, A. (1980). Efficient families of ratio and product-type estimators. *Biometrika*, 67(1):211.

Robson, D. S. (1957). Applications of multivariate polykays to the theory of unbiased ratio-type estimation. *Journal of the American Statistical Association*, 52(280):511–522.

Ruppert and Wand (1994). Multivariate locally weighted least squares regression. *The Annals of statistics*, 22(3):1346–1370.

Saini, M. and Kumar, A. (2016). Ratio estimators for the finite population mean under simple random sampling and rank set sampling. *International Journal of System Assurance Engineering and Management*.

Searls, D. T. (1964). The utilization of a known coefficient of variation in the estimation procedure. *Journal of the American Statistical Association*, 59(308):1225.

Shuy, R. (2002). In person versus telephone interviewing. *J.F Gubriem,and J.A.*, pp.537-556.

Singh, H. P. and Espejo, M. R. (2000). An improved class of chain regression estimators in two phase sampling. *Statistics and Risk Modeling*, 18(2).

Singh, H. P., Pal, S. K., and Mehta, V. (2016). A generalized class of ratio-cum-dual to ratio estimators of finite population mean using auxiliary information in sample surveys. *Mathematical Sciences Letters*, 5(2):203–211.

Singh, H. P. and Vishwakarma, G. K. (2008). Some families of estimators of variance of stratified random sample mean using auxiliary information. *Journal of Statistical Theory and Practice*, 2(1):21–43.

Singh, M. P. (1967). Ratio cum product method of estimation. *Metrika*, 12(1):34–42.

Singh, V. K., Singh, H. P., Singh, H. P., and Shukla, D. (1994). A general class of chain estimators for ratio and product of two means of a finite population. *Communications in Statistics - Theory and Methods*,23(5):1341–1355.

Singleton, R. and Sraits (2002). Survey interviewing. in j.f. gubriem, and j.a.,holstein(eds.), handbook of interview research: Context and method. *Thousand Oaks, CA: Sage*, (pp.59-82).

Solanki, R. S., Singh, H. P., and Pal, S. K. (2014). Improved ratio-type estimators of finite population variance using quartiles. *Hacettepe Journal of Mathematics and Statistics*, 45(44):1–1.

Srivastava, S. K. (1983). Predictive estimation of finite population mean using product estimator. *Metrika*, 30(1):93–99.

Sukhatme, B. V. (1962). Some ratio-type estimators in two-phase sampling. *Journal of the American Statistical Association*, 57(299):628–632.

Sukhatme, V. (1984). Future dimensions of world food and population. richard g. woods. *Economic Development and Cultural Change*, 32(4):892–897.

Swain, A., Mishra, S. S., and Pradhan, B. (2013). On some predictive ratio type estimators in two stage sampling. *Journal of Statistical Theory and Applications*, 12(1):83.

Watson, G. (1964). Smooth regression analysis. *The Indian Journal of statistics*, Series A 26(4):359–372.

Webster, R., Welham, S. J., Potts, J., and Oliver, M. (2006). Estimating the spatial scales of regionalized variables by nested sampling, hierarchical analysis of variance and residual maximum likelihood. *Computers and Geosciences*, 32(9):1320–1333.

Welsh, A. H. and Ronchetti, E. (1998). Bias-calibrated estimation from sample surveys containing outliers. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 60(2):413–428.

Wu, C. and Sitter, R. R. (2001). A model-calibration approach to using complete auxiliary information from survey data. *Journal of the American Statistical Association*, 96(453):185–193.

Appendices

Appendix1

MANUSCRIPT NUMBER	1240914
FULL TITLE	Approximation of finite population totals using Lagrange polynomial
PUBLISHER	Scientific Research Publishing
JOURNAL	Open Journal of Statistics
AUTHORS	1. Lamin Kabareh 2. Dr Thomas Mageto 3. Dr Benjamin Muema
ABSTRACT	Approximation of finite population totals in the presence of auxiliary information is considered. A polynomial based on Lagrange polynomial is proposed. Like the local polynomial regression, Horvitz Thompson and ratio estimators, this approximation technique is based on annual population total in order to fit in the best approximating polynomial within a given period of time (years) in this study. This proposed technique has shown to be unbiased under a linear polynomial. The use of real data indicated that the polynomial is efficient and can approximate properly even when the data is unevenly spaced.

Appendix2

MANUSCRIPT NUMBER	1240969
FULL TITLE	**ESTIMATION OF BOUNDED POPULATIONS AND CARRYING CAPACITY WITH THE LOGISTIC MODEL**
PUBLISHER	Scientific Research Publishing
JOURNAL	Open Journal of Statistics
AUTHORS	1. Lamin Kabareh 2. Dr Thomas Mageto
ABSTRACT	Estimation of bounded populations and carrying capacity in the presence of a sample frame is considered. Models based on Logistic model is proposed. Like the existing estimators, this estimation technique deals with initial condition and is based on yearly population totals in order to fit in a model within a given period of time in this study. The proposed Logistic model technique has shown to be efficient especially with large data. The empirical study indicated that the Logistic model is efficient and can estimate properly even in the presence of outliers.

Appendix3

MANUSCRIPT NUMBER	12259
FULL TITLE	COMPARISON OF THE PIECEWISE POLYNOMIAL APPROXIMATE TO THE NEWTON BACKWARD DIFFERENCE POLYNOMIAL APPROXIMATE OF FINITE POPULATION TOTALS
PUBLISHER	International Journal of Multidisciplinary Research Academy
JOURNAL	International Journal of engineering, science and mathematics
AUTHORS	Lamin Kabareh 2. Dr Thomas Mageto
ABSTRACT	Approximation of finite population totals in the presence of auxiliary information is considered. Polynomials based on Piecewise polynomial and Newton backward difference polynomial are proposed. Like the local polynomial regression, Horvitz Thompson and ratio estimators, these approximation techniques are based on annual population totals in order to fit in the best approximating polynomial within a given period of time (years) in this study. The proposed Piecewise polynomial technique has shown to be unbiased under a first order polynomial as we approach the target value as opposed to the Newton backward difference polynomial. The use of real data indicated that the Piecewise polynomial is efficient and can approximate properly and give a smooth curve at the knots in the presence of outliers.