International Journal of Engineering, Science and Mathematics
Vol. 6 Issue 7, November 2017,
ISSN: 2320-0294 Impact Factor: 6.238
Journal Homepage: http://www.ijesm.co.in, **Email: ijesmj@gmail.com**
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

COMPARISON OF THE PIECEWISE POLYNOMIAL APPROXIMATE TO THE NEWTON BACKWARD DIFFERENCE POLYNOMIAL APPROXIMATE OF FINITE POPULATION TOTALS

Lamin Kabareh[1*]
[1]*Pan African University Institute for Basic Sciences, Technology and Innovation (PAUSTI)*
Thomas Mageto[2**]
[2]*Jomo Kenyatta University of Agriculture and Technology,P.O.Box*: 62000−00200,*Nairobi,Kenya*

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.

Keywords: Piecewise polynomial, Newton backward difference polynomial, approximation, finite population total,auxiliary information,Local polynomial regression, Horvitz Thompson and Ratio estimator, outliers, knots.

1.1 INTRODUCTION

Sample surveys are widely used as a cost effective apparatus of data collection and for making valid inference about population parameters. Government bureaus and organizations use such methods to obtain the current information. The foremost aim of a statistician in a sample survey is to obtain information about the population by deriving reliable estimates of unknown population parameters.

This study is using approximation techniques to approximate the finite population total called the piecewisepolynomial and Newton's backward difference polynomial that do not require any selection of bandwidth as in the case of local polynomial regression estimator or have to follow a specific form, instead we allow the data to speak for itself.The piecewise and Newton's backwarddifference polynomials are used for polynomial interpolation and extrapolation. For a given set of distinct points x_j and numbers y_j, these polynomials of lowest degrees that assume at each point x_j the corresponding value y_j (i.e. the functions coincide at each point). The backward difference interpolation method was first introduced for the evaluation of univariate polynomials at many equidistant points on the real line [1]as will be seen later on how it works. [2] has developed specialized difference formulas for interpolation for equidistant abscissae in which piecewise polynomials, not involving derivatives of the given function, have continuous derivatives at the mesh points.

[3] 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 $\pi_i = p(i \in s)$ be the first order inclusion probabilities. In 1940, 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 [4]. He developed the ratio estimator to estimate the population mean or the total of the study variable y. The ratio estimator of population $\overline{Y}$ is of the form

$$\overline{y_r} = \frac{\overline{y}}{\overline{x}}\overline{X}; \quad \overline{x} \neq 0$$

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 $\overline{y}_r$ is most efficient than the sample mean estimator $\overline{y}$, 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

$$\rho_{yx} > \frac{1}{2}\left(\frac{C_x}{C_y}\right) \dots \dots \dots \dots (1.0)$$

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. [5] further extended the work of [6] on systematic sampling. [7] 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 $\overline{Y}$ is of the form

$$\overline{y_q} = \frac{\overline{y}}{\overline{X}}\overline{x}; \quad \overline{X} \neq 0 \dots \dots \dots \dots \dots .. (1.1)$$

was proposed by [8]. It has been observed that the product estimator gives higher precision than the sample mean estimator $\overline{y}$under the condition that is if

$$\rho_{yx} =< -\frac{1}{2}\left(\frac{C_x}{C_y}\right) \dots \dots \dots \dots \dots . (1.2)$$

The expressions for bias and mean square errors of $\overline{y}_r$ and $\overline{y}_q$ have been derived by [9].

[10] made use of known value of $\overline{X}$ for defining the difference estimator

$\overline{y}_d = \overline{y} + \beta(\overline{X} - \overline{x})$.......(1.3)

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

$$\beta = \frac{S_{yx}}{S_x^2} \dots \dots \dots \dots .. (1.4)$$

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

$$b = \frac{s_{yx}}{s_x^2} \dots \dots \dots \dots . (1.5)$$

Using sample regression coefficient (i.e. b), [22] defined simple linear regression estimator as

$\bar{y}_{1r} = \bar{y} + b(\bar{X} + \bar{x})$........(1.6)

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, [11] suggested a one-parameter family of factor-type(F-T) ratio estimators defined as

$$\bar{y}_f = \bar{y}\left[\frac{(A+C)\bar{X} + fB\bar{x}}{(A+fB)\bar{X} + C\bar{x}}\right] \dots\dots\dots\dots (1.7)$$

where A=(d-1)(d-2), B=(d-1)(d-4), C=(d-2)(d-3)(d-4), d>0, $f = \frac{n}{N}$. The literature on survey sampling describes a great variety of techniques of using auxiliary information to obtained more efficient estimators. 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 [12] and Singh et al. [13].

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

$$\hat{R} - R = \frac{\bar{y} - R\bar{x}}{\bar{x}} = \frac{\bar{y} - R\bar{x}}{\bar{X}}$$

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

The concept of nonparametric models within a model assisted framework was first introduced by [14] 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. The estimator could also be written as

$$\hat{Y}_{gen} = \sum_{i\in s}\frac{y_i}{\pi_i} + \left(\sum_{j=1}^{N}\hat{\mu}\,(x_j) - \sum_{i\epsilon s}\frac{\hat{\mu}(x_i)}{\pi_i}\right) \dots\dots\dots\dots (1.8)$$

The first term in (1.8) is a design estimator which 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. [15] proposed 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 $\tilde{Y} = \sum_{i\epsilon s}\frac{y_i}{\pi_i}$

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 [16], let the (X_i, Y_i) be ordered pairs such that

$$Y_i = m(X_i) + \sigma(X_i)\varepsilon_i \dots\dots\dots\dots (1.9)$$

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

$$m(x) \approx m(x_o) + m^{'}(x_o)(x - x_o) + \cdots + \frac{m^n(x_o)}{n!}(x - x_o)^n \ldots \ldots \ldots \ldots (1.91)$$

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

$$\sum_{i=1}^{n}[Y_i - \sum_{j=0}^{q}\beta_j(X_i - x_0)^j]^2 K_h(X_i - x_0) \quad ....(1.92)$$

In (1.92), h controls the size of the neighborhood around x_0, and $K_h(.)$ controls the weights, where $K_h(.) \equiv \frac{K\left(\frac{\cdot}{h}\right)}{h}$, and K is a kernel function. Denote the solution to (1.92) as $\hat{\beta}$. Thenestimated $m^v(x_0) = v!\,\hat{\beta}_V$. [17] proposed to use nonparametric method to obtain $\mu(.)$.

In nonparametric regression, the user approaches the problem without assuming a model and attempts to fit a curve to the data points by employing a weighting scheme [18] and [19]. Most often, nonparametric regression (for example, the kernel regression, local linear regression) is employed when a theoretical reference curve is unavailable for a process and the data size is large [20] and[21]. However, this estimator experiences a twin problem of how to determine the optimal degrees of the local polynomial. A higher degree polynomial yields a smoother $\overline{\mu}(.)$ but worsens the boundary variance [22]. Such estimators are challenging to employ in cases of multiple covariates and when data is sparse. Another challenge is how to incorporate categorical covariates. A disadvantage of the parametric regression method is that if the assumed model is misspecified, the fitted curve is affected by high bias [23] ,[24] and [25].

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 [26], splines are piecewise regression functions we constrain to join at points called knots.
The Horvitz- Thompson (HT) estimator, which is originally discussed by [27] doesn't make use of the auxiliary information x_i but instead uses only the study variable y_i to obtain the population total.
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.
Let π_i be the probability of including i^{th} unit of the population in samples. 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.
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

$$\hat{Y}_{HT} = \sum_{i=1}^{N}\frac{y_i}{\pi_i} = \sum_{i=1}^{N} y_i\,\pi_i^{-1} \ldots \ldots \ldots \ldots \ldots . (1.93)$$

$\hat{Y}_{HT}$ is unbiased under design based approach [28]

Variance

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

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 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.

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 especially with an increase in the sample size which increases the likelihood of outliers in the sample.
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

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

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 [29]

2. APPROXIMATION OF FINITE POPULATION TOTALS

In this section, we are basically introducing approximators, that is the piecewise polynomial and Newton's backward difference polynomial approximates of the finite population totals.

2.1 Proposed Piecewise polynomial

Let [a,b] be denoted as time interval that is divided into subintervals of time as $[x_i, x_{i+1}]$, where i=0,1,….,n-1; $x_0 = a and x_n = b$. A piecewise polynomial is a function $p(x)$ denoted as the total at each given time $defined on [a,b] by$

$$p(x) = p_i(x), \ x_i \le x \le x_{i+1}, \ i = 0,1, \dots, n-1, \quad where, \ for$$
$$i = 0,1, \dots, n each function p_i(x) is a polynomial defined on [x_i, x_{i+1}].$$

The degree of $p(x)$ is the maximum degree of each polynomial $p_i(x)$, $for\ i = 0,1, \dots \dots n-1$. It is essential to realize that a piecewise polynomial defined on [a,b] is equal to some polynomial on each subinterval$[x_i, x_{i+1}]$ of $[a, b]$, a successive sample of size n is drawn for $i = 0,1, \dots, n-1$, but a different polynomial may be used for each subinterval. Typically, piecewise polynomials are used to fit smooth functions, and therefore are required to have a certain number of continuous derivatives. This requirement imposes additional constraints on the piecewise polynomial, and therefore the degree of the polynomials used on each subinterval must be chosen sufficiently high to ensure that these constraints can be satisfied.

2.2 Proposed Newtonbackward difference polynomial

Consider a finite population $V = \{V_1, \ V_2, V_3 \dots, V_N\}$ of N units. Let (y, x) be the $(total\ ,\ year\ )$ variables taking non-negative real values (y_i, x_i) respectively, on the unit $V_i(i = 1,2, \dots. N)$. From the population V, a successive sample of size n is drawn.Then, the backward difference polynomial is defined as follows:
Say for the data set (x_i, y_i) , $i = 0,1,2, \dots. n$
First order backward difference $\nabla y_i is defined as$:

$$\nabla y_i = y_i - y_{i-1} \dots \dots \dots \dots (1)$$

Second order backward difference $\nabla^2 y_i\ is\ defined\ as$:

$$\nabla^2 y_i = \nabla y_i - \nabla y_{i-1} \dots \dots \dots . (2)$$

In general, the k^{th} order backward difference is defined as :

$$\nabla^k y_i = \nabla^{k-1} y_i - \nabla^{k-1} y_{i-1} \dots\dots..(3)$$

In this case the reference point is x_m and therefore we can write the Newton backward difference polynomial as:

$$P_n(S) = y_n + s\nabla y_n + \frac{s(s+1)}{2!}\nabla^2 y_n + \dots + \frac{s(s+1)\dots\dots.(s+n-1)}{n!}\nabla^n y_n \dots\dots\dots\dots..(4)$$

Where $s = \frac{x-x_m}{h}$ and h= step size

2.2. Asymptotic properties of polynomial approximations

Polynomial Approximation of Functions:

Karl Weierstrass Theorem:

$f: [a,b] \to R \ \ continuous$

Then there exists a sequence of polynomials $P_n(x)$ such that

$||f - P_n||_\infty = max_{x\in[a,b]}|f(x) - P_n(x)| \to 0 \quad as \quad n \to \infty$

$\boldsymbol{Proof\ of\ Theorem}$:

$f: [a,b] = [0,1] \to R \ \ continuous.$

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

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

We now consider 3 functions: $f(x) = 1$, $f(x) = x$ and $f(x) = x^2$. We shall now show for uniform convergence below:

$\boldsymbol{f(x) = 1}$

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

$\boldsymbol{= (x + 1 - x)^n = 1,\ n \geq 0}$ **hence converged**

$\boldsymbol{Similarly}$;

$\boldsymbol{f(x) = x}$

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

$$\boldsymbol{= \sum_{k=1}^{n} \binom{n-1}{k-1} x^k (1-x)^{n-k}}$$

$\boldsymbol{Let\ r = k - 1}$

$$\boldsymbol{= x \sum_{r=0}^{n-1} \binom{n-1}{r} x^r (1-x)^{n-1-r}}$$

$$= x\{x + 1 - x\}^{n-1} = x, \quad n \geq 1 \quad \text{hence converged}$$

$Finally;$

$$f(x) = x^2$$

$$P_n(x) = B_n(x) = \sum_{k=0}^{n} \binom{n}{k} \left(\frac{k}{n}\right)^2 x^k (1-k)^{n-k}$$

$$= \sum_{k=1}^{n} \binom{n-1}{k-1} \frac{k-1+1}{n} x^k (1-x)^{n-k}$$

$$= \sum_{k=2}^{n} \binom{n-1}{k-2} \frac{x^k}{n} (1-x)^{n-k} + \frac{1}{n} \sum_{k=1}^{n} \binom{n-1}{k-1} x^k (1-x)^{n-k}$$

$$P_n(x) = B_n(f)(x) = \frac{(n-1)}{n} x^2 \sum_{k=2}^{n} \binom{n-2}{k-2} x^{k-2} (1-x)^{n-k} + \frac{1}{n} x$$

$$B_n(f)(x) = \frac{n-1}{n} x^2 + \frac{1}{n} x$$

$$|f(x) - B_n(f)(x)| = \frac{1}{n} |x(1-x)|, \quad for \quad n \geq 2$$

$$||f - B_n(f)||_\infty = (4n)^{-1} \to 0 \ as \ n \to \infty$$

$Comment$: *For one to get a good approximating polynomial that minimizes the error, one need choose an interpolating polynomial that is closest to the target point.*

3. MAIN RESULTS:

3.1 Empirical analysis

Table 1

i	YEAR (X_i)	POPULATION TOTAL($Y_i = P_i(X_i)$)
1	1969	10,942,705
2	1979	15,327,061
3	1989	21,448,774
4	1999	28,686,607
5	2009	38,610,097

The Kenya population census data since 1969 to 2009 is shown in Table 1 above.However, we aimed at selecting successive sample sizes of two and also three as well from the table. These sample sizes will be used to approximate the population total in 2019 census using the proposed techniques.

3.2 Piecewise Polynomial
3.2 .1 Linear Piecewise Polynomial

$$P_1(x) = a_1 + b_1x \qquad 1969 \le x \le 1979$$

$$P_2(x) = a_2 + b_2x \qquad 1979 \le x \le 1989$$

$$P_3(x) = a_3 + b_3x \qquad 1989 \le x \le 1999$$

$$P_4(x) = a_4 + b_4x \qquad 1999 \le x \le 2009$$

3.2.2Polynomials from samples:

$$a_1 + 1969b_1 = 10{,}942{,}705 \ldots\ldots\ldots\ldots\ldots(1)$$

$$a_1 + 1979b_1 = 15{,}327{,}061 \ldots\ldots\ldots\ldots\ldots\ldots(2)$$

$$a_2 + 1979b_2 = 15{,}327{,}061 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3)$$

$$a_2 + 1989b_2 = 21{,}448{,}774 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4)$$

$$a_3 + 1989b_3 = 21{,}448{,}774 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(5)$$

$$a_3 + 1999b_3 = 28{,}686{,}607 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(6)$$

$$a_4 + 1999b_4 = 28{,}686{,}607 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(7)$$

$$a_4 + 2009b_4 = 38{,}610{,}097 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(8)$$

3.2.3 $\boldsymbol{Derivatives\ at\ the\ Knots\ of\ the\ polynomials}$

$$\frac{d}{dx}(a_1 + b_1x) = \frac{d}{dx}(a_2 + b_2x) \qquad at\ x = 1979$$

$$b_1 = b_2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(9)$$

$$\frac{d}{dx}(a_2 + b_2x) = \frac{d}{dx}(a_3 + b_3x) \quad at \quad x = 1989$$

$$b_2 = b_3 \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots (10)$$

$$\frac{d}{dx}(a_3 + b_3x) = \frac{d}{dx}(a_4 + b_4x) \qquad at\ x = 19$$

$$b_3 = b_4 \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots (11)$$

$Table$ 2: Coefficients of linear spline

i	a_i	b_i
1	-852336991.4	438435.6
2	-1196159942	612171.3
3	-1418156210	723783.3
4	-1955019044	992349

Table 3: Approximates of population totals

$P_i(x)$	2009	2019
1	28480129	32864485
2	33693200	39813913
3	35924440	43162273
4	38,610,097	48,533,587

3.2.4 Quadratic Piecewise Polynomial

$$P_1(x) = a_1x^2 + b_1x + c_1 \qquad 1969 \le x \le 1979$$

$$P_2(x) = a_2x^2 + b_2x + c_2 \qquad 1979 \le x \le 1989$$

$$P_3(x) = a_3x^2 + b_3x + c_3 \qquad 1989 \le x \le 1999$$

$$P_4(x) = a_4x^2 + b_4x + c_4 \qquad 1999 \le x \le 2009$$

3.3.1 Polynomials from samples

$$a_1(1969)^2 + b_1(1969) + c_1 = 10{,}942{,}705 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.(12)$$

$$a_1(1979)^2 + b_1(1979) + c_1 = 15{,}327{,}061 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.(13)$$

$$a_2(1979)^2 + b_2(1979) + c_2 = 15{,}327{,}061 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.(14)$$

$$a_2(1989)^2 + b_2(1989) + c_2 = 21{,}448{,}774 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(15)$$

$$a_3(1989)^2 + b_3(1989) + c_2 = 21{,}448{,}774 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.(16)$$

$$a_3(1999)^2 + b_3(1999) + c_3 = 28{,}686{,}607 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots..(17)$$

$$a_4(1999)^2 + b_4(1999) + c_4 = 28{,}686607 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(18)$$

$$a_4(2009)^2 + b_4(2009) + c_4 = 38{,}610{,}097 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(19)$$

3.2.5 Derivatives at the knots

$$\frac{d}{dx}(a_1x^2 + b_1x + c_1) = \frac{d}{dx}(a_2x^2 + b_2x + c_2) \quad at \;\; x = 1979$$

$$3958a_1 + b_1 - 3958a_2 - b_2 = 0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(20)$$

$$\frac{d}{dx}(a_2x^2 + b_2x + c_2) = \frac{d}{dx}(a_3x^2 + b_3x + c_3) \quad at \;\; x = 1989$$

$$3978a_2 + b_2 - 3978a_3 - b_3 = 0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(21)$$

$$\frac{d}{dx}(a_3x^2 + b_3x + c_3) = \frac{d}{dx}(a_4x^2 + b_4x + c_4) \quad at \;\; x = 1999$$

$$3998a_3 + b_3 - 3998a_4 - b_4 = 0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(22)$$

Let $a_1 = 0 \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots . (23)$

Table 4: coefficients of the quadratic spline

i	a_i	b_i	c_i
1	0	438435.6	-847952635.4
2	17373.57	-68326154.46	67190265197
3	-6212.37	25498714.86	-26118607666
4	33068.94	-131547963.1	130849610097

Table 5: Approximates of population totals

$P_i(x)$	2009	2019
1	32,864,485	37,248,841
2	44,156,666	60,702,521
3	34,681,966	39,434,851
4	38,710,097	55,210,097

3.3 Newton backward difference polynomial

Table 6

x	y=f(x)	∇y	$\nabla^2 y$	$\nabla^3 y$	$\nabla^4 y$
1969	10,942,705				
1979	15,327,061	4,384,356			
1989	21,448,774	6,121,713	1,737,357		
1999	28,686,607	7,237,833	1,116,120	-621,237	
2009	38,610,097	9,923,490	2,685,657	1,569,537	2,190,774

$$Suppose\ we\ want\ to\ approximate\ the\ population\ in\ 2009.$$

Here $s = \frac{x-x_m}{h} wherex = 2009,\ \ x_m = 1999\ \ (referencevalue) andh = 10$

Therefore $s = 1$

3.3.1 First order Newton backward approximate for 2009 will be:

$$y_{2009} = 28{,}686{,}607 + (1)(7{,}237{,}833)$$

$$y_{2009} = 35{,}924{,}440$$

3.3.2 Second order Newton backward for 2009 will be:

$$y_{2009} = 28{,}686{,}607 + (1)(7{,}237{,}833) + \frac{(1)(1+1)}{2!}(1116120)$$

$$y_{2009} = 37{,}040{,}560$$

$$Suppose\ we\ want\ to\ approximate\ the\ population\ in\ 2019$$

$$x = 2019, \qquad x_m = 2009\ (reference\ value)\ \ and\ h = 10$$

3.3.3 First order Newton backward difference for 2019 will be:

$$y_{2019} = 38{,}610{,}097 + 9923490$$

$$y_{2019} = 48{,}533{,}587$$

3.3.4 Second order Newton backward difference for 2019 will be:

$$y_{2019} = 38{,}610{,}097 + 9{,}923{,}490 + 2{,}685{,}657$$

$y_{2019} = 51{,}219{,}244$and order four has approximated to be 54,979,555 people in 2019 census.

3.3.5 DISCUSSIONS:

The empirical study has shown that the linear piecewise polynomial in Table 3 gave little or no variations as we approach the target value. Which means a faster convergence can be obtained if our interpolating polynomial is linear and closest to the target value. The result shown on the table at $P_4(2009)$ has confirmed the Karl Weierstrass theorem of uniform convergence. One can rely on this single polynomial to talk about the future (extrapolate), which approximated the population total in 2019 to be 48,533,587.
On the other hand, the same real data study with a quadratic polynomial in Table 5 showed a slight variation with regard to approximating the population total in 2009 making an overestimation error of 100,000 and projected the population in 2019 to be 55,210,097.
In comparison with the Newton backward difference polynomial, the linear polynomial in 3.3.1 for 2009 is approximated to be 35,924,440 making an underestimation error of 2,685,657. Furthermore, the second order Newton backward difference for 2009 approximated the population as 37,040,560 making an underestimation error of 1,569,537.
Lastly, the first order Newton backward difference approximated the population in 2019 as 48,533,587 while the second order Newton backward difference approximated it to be 51,219,244 as shown in 3.3.3 and 3.3.4 respectively.

4.0. CONCLUSION:

In this work, the piecewise polynomial is more efficient in terms of prediction than the Newton backward polynomial. The piecewise polynomials are very effective with the presence of outliers in trying to get smooth curves at the knots which many polynomials lack. They can perform well as the degree of the polynomial increases with little variations. The linear piecewise polynomials, when combined to give a single polynomial do not satisfy the condition of differentiability at the knots which is a requirement for uniform convergence to take place. That is why, we rely on piecewise polynomials of degree greater than one for prediction purpose.

Disclosure of potential conflicts of interest: The author(s) declare(s) that there is no conflict of interest regarding the publication of this paper.

Acknowledgements: We are grateful to the authors for their numerous and valuable contributions to this work, most especially the first author.

References:

[1] W.Volk,1988. An efficient raster evaluation method for univariate polynomials, computing 40, 163-173.

[2] H. Rutishauser, 1959. Notes on smooth interpolation, ZAMP, vol.1, pp.60-63

[3] Deville, J.-C., Sarndal, C.-E., 1992. Calibration estimators in survey sampling. Journal of the American Statistical Association 87 (418), 376.

[4] Cochran, W. G., Goulden, C. H., 1940. Methods of statistical analysis. Journal of the Royal Statistical Society 103 (2), 250.

[5] Cochran, W. G., 1946. Graduate training in statistics. The American Mathematical Monthly 53 (4), 193.

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

[7] Singh, V. K., Singh, H. P., Singh, H. P., 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.

[8] 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.

[9] Wu, C., 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.

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

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

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

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

[14] Lairez, P., 2016. A deterministic algorithm to compute approximate roots of polynomial systems in polynomial average time. Foundations of Computational Mathematics.

[15] Robson, D. S., 1957. Applications of multivariate polykays to the theory of unbiased ratiotype estimation. Journal of the American Statistical Association 52 (280), 511–522.

[16] Montanari, G. E., Ranalli, M. G., 2003. On calibration methods for design based finite population inferences. Bulletin of the International Statistical Institute, 54 th session 60.

[17] Hansen, M. H., Hurwitz, W. N., Madow, W. G., 1953. Sample survey methods and theory. Vol. 1. Wiley New York.

[18] Robson, D. S., 1957. Applications of multivariate polykays to the theory of unbiased ratiotype estimation. Journal of the American Statistical Association 52 (280), 511–522.

[19] Uysal, I., & Guvenir, A. (1999). An overview of regression techniques for knowledge discovery. The Knowledge Engineering Review, 14(4), 319-340.

[20] DiNardo, J., & Tobias, J. L. (2001). Nonparametric density and regression estimation. The Journal of Economic Perspectives, 15(4), 11-28.

[21] Hens, N. (2005). Non- and semiparametric techniques for handling missing data. (Doctoral dissertation). Diepenbeek, Belgium: Transnationale Universiteit Limburgs.

[22] Hernández-Lobato, J. M. (2010). Balancing flexibility and robustness in machine learning: Semiparametric methods and sparse linear models. (Doctoral dissertation). Madrid, Spain: University of Autonoma Department of Computer Science.

[23] Einsporn, R., & Birch, J. B. (1993). Model robust regression: Using non-parametric regression to improve parametric regression analysis. (Technical Report, 93-5). Blacksburg, VA: Department of Statistics, Virginia Polytechnic Institute & State University.

[24] Mays, J. E. (2001b). Small-sample model-robust confidence intervals in regression.In

Proceedings fromAnnual Meeting of the American Statistical Association.Retrieved from http://www.amstat.org/sections/SRMS/Proceedings/y2001/Proceed/00530.pdf

[25] Pickle, S. M. (2006). Semiparametric techniques for response surface methodology (Doctoral dissertation). Virginia Polytechnic Institute and State University: Blacksburg, VA.

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

[27] Keele, L. J., 2008. Semiparametric regression for the social sciences. John Wiley and Sons.

[28] Singh, H. P., Pal, S. K., 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.

[29] Burden, R. L., Faires, J. D., 2001. Numerical analysis. 2001. Brooks/Cole, USA.