ISSN Print: 2394-7500
ISSN Online: 2394-5869
Impact Factor: 5.2
IJAR 2016; 2(10): 513-517
www.allresearchjournal.com
Received: 14-08-2016
Accepted: 15-09-2016

Biswajit Das
Research Scholar, Assam Down Town University, Panikhaiti, Guwahati, Assam, India

Dhritikesh Chakrabarty
Department of Statistics, Handique Girls' College, Guwahati, Assam, India
& Research Guide, Assam Down Town University, Panikhaiti, Guwahati, Assam, India

Correspondence
Biswajit Das
Research Scholar, Assam Down Town University, Panikhaiti, Guwahati, Assam, India

Newton's backward interpolation: Representation of numerical data by a polynomial curve

Biswajit Das and Dhritikesh Chakrabarty

Abstract
In order to reduce the numerical computations associated to the repeated application of the existing interpolation formula in computing a large number of interpolated values, a formula has been derived from Newton's backward interpolation formula for representing the numerical data on a pair of variables by a polynomial curve. Application of the formula to numerical data has been shown in the case of representing the data on the total population of India corresponding as a function of time. The formula is suitable in the situation where the values of the argument (i.e. independent variable) are at equal interval.

Keywords: Interpolation, newton's backward interpolation formula, polynomial curve, representation of numerical data

1. Introduction
Interpolation, which is the process of computing intermediate values of a function from the set of given values of the function {Hummel (1947), Erdos & Turan (1938) *et al*}, plays significant role in numerical research almost in all branches of science, humanities, commerce and in technical branches. A number of interpolation formulas namely Newton's Forward Interpolation formula, Newton's Backward Interpolation formula, Lagrange's Interpolation formula, Newton's Divided Difference Interpolation formula, Newton's Central Difference Interpolation formula, Stirlings formula, Bessel's formula and some others are available in the literature of numerical analysis {Bathe & Wilson (1976) [1], Jan (1930), Hummel (1947) *et al*}.
In case of the interpolation by the existing formulae, the value of the dependent variable corresponding to each value of the independent variable is to be computed afresh from the used formula putting the value of the independent variable in it. That is if it is wanted to interpolate the values of the dependent variable corresponding to a number of values of the independent variable by a suitable existing interpolation formula, it is required to apply the formula for each value separately and thus the numerical computation of the value of the dependent variable based on the given data are to be performed in each of the cases. In order to get rid of these repeated numerical computations from the given data, one can think of an approach which consists of the representation of the given numerical data by a suitable polynomial and then to compute the value of the dependent variable from the polynomial corresponding to any given value of the independent variable. However, a method/formula is necessary for representing a given set of numerical data on a pair of variables by a suitable polynomial. One such formula has been developed in this study. The formula has been derived from Newton's backward interpolation formula. The formula obtained has been applied to represent the numerical data, on the total population of India since 1971, by a polynomial curve.

2. Newton's Backward Interpolation Formula
Then Newton's backward interpolation formula is

$$f(x) = f(x_n) + v\nabla f(x_{n-1}) + \frac{v\,(v+1)}{2!}\nabla^2 f(x_{n-2}) + \frac{v\,(v+1)(v+2)}{3!}\nabla^3 f(x_{n-3}) + \frac{v\,(v+1)(v+2))(v+3)}{4!}\nabla^4 f(x_{n-4}) + \ldots\ldots + \frac{v\,(v+1)(v+2)\ldots\ldots(v+n-1)}{n!}\nabla^n f(x_0)$$

where $v = \frac{x-x_n}{h}$

$$= f(x_n) + (x-x_n)\frac{\nabla f(x_{n-1})}{h} + (x-x_n)(x-x_{n-1})\frac{\nabla^2 f(x_{n-2})}{2!h^2} + (x-x_n)(x-x_{n-1})$$
$$(x-x_{n-2})\frac{\nabla^3 f(x_{n-3})}{3!h^3} + (x-x_n)(x-x_{n-1})(x-x_{n-2})(x-x_{n-3})\frac{\nabla^4 f(x_{n-4})}{4!h^4}$$
$$+\ldots\ldots\ldots\ldots + (x-x_n)(x-x_{n-1})(x-x_{n-2})(x-x_{n-3})\ldots\ldots\ldots(x-x_1)\ \frac{\nabla^n f(x_0)}{n!h^n} \quad (2.1)$$

This formula can be expressed as

$f(x) = C_n + C_{n-1}(x-x_n) + C_{n-2}(x-x_n)(x-x_{n-1}) + C_{n-3}(x-x_n)(x-x_{n-1})(x-x_{n-2}) +$
$C_{n-4}(x-x_n)(x-x_{n-1})(x-x_{n-2})(x-x_{n-3}) + \ldots\ldots\ldots\ldots +$
$C_0(x-x_n)(x-x_{n-1})(x-x_{n-2})(x-x_{n-3})\ldots\ldots\ldots(x-x_1)$
$\rightarrow$ (2.2)

where $C_n = f(x_n)$

$C_{n-1} = \frac{\nabla f(x_{n-1})}{h}$

$C_{n-2} = \frac{\nabla^2 f(x_{n-2})}{2!h^2}$

$C_{n-3} = \frac{\nabla^3 f(x_{n-3})}{3!h^3}$

$C_{n-4} = \frac{\nabla^4 f(x_{n-4})}{4!h^4}$

…………………....

……………………..

$C_0 = \frac{\nabla^n f(x_0)}{n!h^n}$

3. Representation of Numerical Data by Polynomial Curve:

By algebraic expansion, one can obtain that

$(x-x_n) = (x-x_n)$

Also,

$(x-x_n)(x-x_{n-1}) = x^2 - x.x_{n-1} - x.x_n + x_n x_{n-1}$
$= x^2 - (x_{n-1} + x_n)\,x + x_n x_{n-1}$
$= x^2 - \left(\sum_{i=n-1}^{n} x_i\right) x + x_n x_{n-1}$

Again,

$(x-x_n)(x-x_{n-1})(x-x_{n-2})$

$= (x^2 - x.x_{n-1} - x.x_n + x_n x_{n-1})\,(x-x_{n-2})$

$= x^3 - x^2.x_{n-1} - x^2.x_n + x\,x_n x_{n-1} - x^2.x_{n-2} + x.x_{n-2}x_{n-1} + x.x_n x_{n-2} - x_{n-2}x_{n-1}x_n$

$= x^3 - (x_{n-2} + x_{n-1} + x_n)\,x^2 + (x_{n-2}x_{n-1} + x_{n-2}x_n + x_{n-1}x_n)\,x - x_{n-2}x_{n-1}x_n$

$= x^3 - \left(\sum_{i=n-2}^{n} x_i\right) x^2 + \left(\sum_{i=n-2}^{n-1}\sum_{j=n-1}^{n} x_i x_j\right) x - x_{n-2}x_{n-1}x_n$

Similarly,

$(x-x_n)(x-x_{n-1})(x-x_{n-2})(x-x_{n-3})$

$= \{x^3 - (x_{n-2} + x_{n-1} + x_n)\,x^2 + (x_{n-2}x_{n-1} + x_{n-2}x_n + x_{n-1}x_n)\,x - x_{n-2}x_{n-1}x_n\}\,(x-x_{n-3})$

$= x^4 - (x_{n-2} + x_{n-1} + x_n)\,x^3 + (x_{n-2}x_{n-1} + x_{n-2}x_n + x_{n-1}x_n)\,x^2 - x_{n-2}x_{n-1}x_n\,x - x_{n-3}\,x^3$

$+ (x_{n-2} + x_{n-1} + x_n)\,x^2 x_{n-3} - (x_{n-2}x_{n-1} + x_{n-2}x_n + x_{n-1}x_n)x_{n-3}\,x + x_{n-3}x_{n-2}x_{n-1}x_n$

$= x^4 - (x_{n-3} + x_{n-2} + x_{n-1} + x_n)\,x^3 + (x_{n-3}x_{n-2} + x_{n-3}x_{n-1} + x_{n-2}x_{n-1} + x_{n-2}x_n + x_{n-3}x_n + x_{n-1}x_n)\,x^2 - (x_{n-3}x_{n-2}x_{n-1} + x_{n-3}x_{n-2}x_n + x_{n-3}x_{n-1}x_n + x_{n-2}x_{n-1}x_n)\,x + x_{n-3}x_{n-2}x_{n-1}x_n$

$= x^4 - \left(\sum_{i=n-3}^{n} x_i\right) x^3 + \left(\sum_{i=n-3}^{n-1}\sum_{j=n-2}^{n} x_i x_j\right) x^2 - \left(\sum_{i=n-3}^{n-2}\sum_{j=n-2}^{n-1}\sum_{k=n-1}^{n} x_i x_j x_k\right) x +$

$x_{n-3}x_{n-2}x_{n-1}x_n$

……………………………………………………………………………………………

……………………………………………………………………………………………

In general,

$(x - x_n)(x - x_{n-1})(x - x_{n-2})(x - x_{n-3}) \dots\dots\dots .(x - x_1)$

$= x^n - (\sum_{i=n-(n-1)}^{n} x_i)\, x^{n-1} + (\sum_{i=n-(n-1)}^{n-1} \sum_{j=n-(n-2)}^{n} x_i x_j)\, x^{n-2} -$

$(\sum_{i=n-(n-1)}^{n-(n-2)} \sum_{j=n-(n-2)}^{n-(n-1)} \sum_{k=n-(n-1)}^{n} x_i x_j x_k)\, x + \dots\dots\dots + (-1)^n (x_n x_{n-1} x_{n-2} \dots\dots\dots\dots x_1)$

$= x^n - (\sum_{i=1}^{n} x_i)\, x^{n-1} + (\sum_{i=1}^{n-1} \sum_{j=2}^{n} x_i x_j)\, x^{n-2} - (\sum_{i=1}^{2} \sum_{j=2}^{1} \sum_{k=1}^{n} x_i x_j x_k)\, x + \dots\dots\dots\dots\dots\dots$

$+ (-1)^n (x_n x_{n-1} x_{n-2} \dots\dots\dots\dots x_1)$

Now, equation (2.2), can be expressed as

$f(x) = C_n + C_{n-1}(x - x_n) + C_{n-2}\, [x^2 - (\sum_{i=n-1}^{n} x_i)\, x + x_n x_{n-1}\,] + C_{n-3}\, [\, x^3 - (\sum_{i=n-2}^{n} x_i)\, x^2 +$
$(\sum_{i=n-2}^{n-1} \sum_{j=n-1}^{n} x_i x_j)\, x - x_{n-2} x_{n-1} x_n] + C_{n-4}\, [x^4 - (\sum_{i=n-3}^{n} x_i)\, x^3 +$
$(\sum_{i=n-3}^{n-1} \sum_{j=n-2}^{n} x_i x_j)\, x^2 - (\sum_{i=n-3}^{n-2} \sum_{j=n-2}^{n-1} \sum_{k=n-1}^{n} x_i x_j x_k)x + x_{n-3} x_{n-2} x_{n-1} x_n\,]$

………………………………………………………………………………………..

$C_0\, [\, x^n - (\sum_{i=1}^{n} x_i)\, x^{n-1} + (\sum_{i=1}^{n-1} \sum_{j=2}^{n} x_i x_j)\, x^{n-2} - (\sum_{i=1}^{2} \sum_{j=2}^{1} \sum_{k=1}^{n} x_i x_j x_k)\, x +$
$+ \dots\dots\dots\dots\dots + (-1)^n (x_n x_{n-1} x_{n-2} x_{n-3} \dots\dots\dots\dots x_1)] \rightarrow$ **(3.1)**

This is of the form

$f(x) = A_0 + A_1 x + A_2\, x^2 + A_3\, x^3 + \dots\dots\dots\dots\dots\dots\dots\dots + A_n\, x^n \rightarrow$ **(3.2)**

Where

$A_0 = C_n - C_{n-1}\, x_n + C_{n-2}\, x_n x_{n-1} - C_{n-3}\, x_{n-2} x_{n-1} x_n + C_{n-4}\, x_{n-3} x_{n-2} x_{n-1} x_n\, -$
$\dots\dots\dots\dots\dots\dots\dots\dots + (-1)^n\, C_0\, (x_n x_{n-1} x_{n-2} x_{n-3} \dots\dots\dots\dots x_1)$

$A_1 = C_{n-1} - C_{n-2}\, (\sum_{i=n-1}^{n} x_i) + C_{n-3}\, (\sum_{i=n-2}^{n-1} \sum_{j=n-1}^{n} x_i x_j) -$
$C_{n-4}\, (\sum_{i=n-3}^{n-2} \sum_{j=n-2}^{n-1} \sum_{k=n-1}^{n} x_i x_j x_k) + \dots\dots\dots\dots - C_0 (\sum_{i=1}^{2} \sum_{j=2}^{1} \sum_{k=1}^{n} x_i x_j x_k)$

$A_2 = C_{n-2} - C_{n-3}\, (\sum_{i=n-2}^{n} x_i) + C_{n-4}\, (\sum_{i=n-3}^{n-1} \sum_{j=n-2}^{n} x_i x_j) - \dots\dots\dots\dots$
$+ C_0\, (\sum_{i=2}^{3} \sum_{j=3}^{2} \sum_{k=2}^{1} \sum_{l=1}^{n} x_i x_j x_k x_l)$

$A_3 = C_{n-3} - C_{n-4}\, (\sum_{i=n-3}^{n} x_i) + \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots +$
$C_0\, (\sum_{i=3}^{4} \sum_{j=4}^{3} \sum_{k=3}^{2} \sum_{l=2}^{1} \sum_{m=1}^{n} x_i x_j x_k x_l x_m)$

………………………………………………………………………………………

………………………………………………………………………………………

$A_n = C_n$

Equation (3.2), with the coefficients

$A_0, A_1, A_2, A_3, \dots\dots\dots\dots\dots\dots\dots\dots, A_n,$

,as defined above, is the required formula for representing a given set of numerical data on a pair of variables by a suitable polynomial we have aimed at.

Note: The formula is valid for representing a given set of numerical data on a pair of variables by a suitable polynomial under the following two conditions:

(i) values of the argument are at equal interval

(ii) the value of x corresponding to which the value of y is to be interpolated is in the last half of the series

4. An Example of the Formula:

The following table shows the data on total population of India corresponding to the years:

Year	1971	1981	1991	2001	2011
Total Population	548159652	683329097	846302688	1027015247	1210193422

Taking 1971 as origin and changing scale by 1/10, one can obtain the following table for independent variable *x* (representing time) and *f*(*x*) (representing total population of India):

Year	1971	1981	1991	2001	2011
x_i	0	1	2	3	4
$f(x_i)$	548159652	683329097	846302688	1027015247	1210193422

Here $x_0 = 0$, $x_1 = 1$, $x_2 = 2$, $x_3 = 3$, $x_4 = 4$

$f(x_0) = 548159652$, $f(x_1) = 683329097$, $f(x_2) = 846302688$, $f(x_3) = 1027015247$

$f(x_4) = 1210193422$

Difference Table

x	$f(x)$	$\nabla f(x)$	$\nabla^2 f(x)$	$\nabla^3 f(x)$	$\nabla^4 f(x)$
0	548159652				
		135169445			
1	683329097		27804146		
		162973591		– 10065178	
2	846302688		17738968		– 5208174
		180712559		–15273352	
3	1027015247		2465616		
		183178175			
4	1210193422				

Now, $C_4 = f(x_4) = 1210193422$

$C_3 = \frac{\nabla f(x_3)}{h} = 183178175$

$C_2 = \frac{\nabla^2 f(x_2)}{2!h^2} = \frac{2465616}{2.1.1} = \frac{2465616}{2} = 1232808$

$C_1 = \frac{\nabla^3 f(x_1)}{3!h^3} = \frac{-15273352}{3.2.1.1} = \frac{-15273352}{6} = -2545558.66$

$C_0 = \frac{\nabla^4 f(x_0)}{4!h^4} = \frac{-5208174}{4.3.2.1.1} = \frac{-5208174}{24} = -217007.25$

∴ The polynomial is

$f(x) = A_0 + A_1 x + A_2 x^2 + A_3 x^3 + A_4 x^4$……………………………………………………………………………....………. (i)

Where $A_0 = C_4 - C_3 x_4 + C_2 x_4 x_3 - C_1 x_2 x_3 x_4 + C_0 x_1 x_2 x_3 x_4$

= 1210193422 – 183178175 × 4 + 1232808 × 4 × 3 + 2545558.66 × 2 × 3 × 4
– 217007. 25 × 1 × 2 × 3 × 4
= 1210193422 – 732712700 + 14793696 + 61093407.84 – 5208174
= 1286080525.84 – 737920874
= 548159651.84

$A_1 = C_3 - C_2 (x_3 + x_4) + C_1 (x_2 x_3 + x_2 x_4 + x_3 x_4) - C_0 (x_1 x_2 x_3 + x_1 x_2 x_4 + x_1 x_3 x_4 + x_2 x_3 x_4)$
= 183178175 – 1232808 × (3 + 4) – 2545558.66 × (6 + 8 + 12) + 217007.25 (6 + 8 + 12 + 24)
= 183178175 – 1232808 × 7 – 2545558.66 × 26 + 217007.25 × 50
= 183178175 – 8629656 – 66184525.16 + 10850362.5
= 194028537.5– 74814181.16
= 119214356.34
$A_2 = C_2 - C_1 (x_2 + x_3 + x_4) + C_0 (x_1 x_2 + x_1 x_3 + x_2 x_3 + x_2 x_4 + x_1 x_4 + x_3 x_4)$
= 1232808 + 2545558.66 (2 + 3 + 4) – 217007.25 (2 + 3 + 6 + 8 + 4 + 12)
=1232808 + 2545558.66 × 9 – 217007.25 × 35
= 1232808 + 22910027.94 – 7595253.75
= 24142835.94 – 7595253.75
= 16547582.19
$A_3 = C_1 - C_0 (x_1 + x_2 + x_3 + x_4)$
= – 2545558.66 + 217007.25 × (1 + 2 + 3 + 4)
= – 2545558.66 + 217007.25 × 10
= – 2545558.66 + 2170072.5
= – 375486.16
$A_4 = C_0 = -217007.25$

∴ (i) ⇒ $f(x) = 548159651.84 + 119214356.34\,x + 16547582.19\,x^2 - 375486.16\,x^3 - 217007.25\,x^4$ ……………………..(ii)
Which is the required polynomial
Now, putting $x = 0, 1, 2, 3, 4$ in (ii), we have
$f(0) = 548159651.84$
= 548159652
$f(1)$ = 548159651.84 + 119214356.34 × 1 + 16547582.19 × 1– 375486.16 × 1 −217007.25 × 1
= 548159651.84 + 119214356.34 + 16547582.19 – 375486.16 −217007.25
= 683921590.37 − 592493.41
= 683329096.96
= 683329097
$f(2)$ = 548159651.84 + 119214356.34 × 2 + 16547582.19 × 4 – 375486.16 × 8 −217007.25 × 16
= 548159651.84 + 238428712.68 + 66190328.76 – 3003889.28 − 3472116

= 852778693.28 − 6476005.28
= 846302688
$f(3)$ = 548159651.84 + 119214356.34 × 3 + 16547582.19 × 9 – 375486.16 × 27 −217007.25 × 81
= 548159651.84 + 357643069.02 + 148928239.71– 10138126.32 −17577587.25
= 1054730960.57 – 27715713.57
= 1027015247
$f(4)$ = 548159651.84 + 119214356.34 × 4 + 16547582.19 × 16 – 375486.16 × 64 – 217007.25 × 256
= 548159651.84 + 476857425.36 + 264761315.04 – 24031114.24 − 55553856
= 1289778392.24 − 79584970.24
= 1210193422

5. Conclusion

The formula described by equation (3.2) can be used to represent a given set of numerical data on a pair of variables, by a polynomial.

The degree of the polynomial is one less than the number of pairs of observations.

The polynomial that represents the given set of numerical data can be used for interpolation at any position of the independent variable lying within its two extreme values.

The approach of interpolation, described here, can be suitably applied in inverse interpolation also.

Newton's backward interpolation formula is valid for estimating the value of the dependent variable under the following two conditions:

(i) Values of the argument are at equal interval

(ii) The value of x corresponding to which the value of y is to be interpolated is in the last half of the series

Therefore, the formula derived here is valid for representing a set of numerical data on a pair of variables by a polynomial under these two conditions only. Consequently, there is necessity of searching for some formula for representing a set of numerical data on a pair of variables by a polynomial if the value of the independent variable corresponding to which the value of the dependent variable is to be estimated lies in the last half of the series of the given values, which are at equal interval, of the independent variable. Moreover, there is also necessity of searching for some formula for representing a set of numerical data on a pair of variables by a polynomial if the given values of the independent variable are not at equal interval.

Reference

1. Bathe KJ, Wilson EL. Numerical Methods in Finite Element Analysis, Prentice-Hall, Englewood Cliffs, NJ, 1976.
2. Gerald CF, Wheatley PO. Applied Numerical Analysis, fifth ed., Addison-Wesley Pub. Co., MA, 1994.
3. David R Kincard, Ward Chaney E. Numerical analysis, Brooks /Cole, Pacific Grove, CA, 1991.
4. Endre S, David Mayers. An Introduction to Numerical Analysis, Cambridge, UK, 2003.
5. John H Mathews, Kurtis D Fink. Numerical methods using MATLAB, 4ed, Pearson Education, USA, 2004.
6. James B. Scarborough Numerical Mathematical Analysis, 6 Ed, The John Hopkins Press, USA, 1996.
7. Kendall E. Atkinson An Introduction to Numerical Analysis, 2 Ed., New York. 1989.
8. Chapra SC, Canale RP. Numerical Methods for Engineers, third ed., McGraw-Hill, New York, 2002.
9. Conte SD, Carl de Boor. Elementary Numerical Analysis, 3 Ed, McGraw-Hill, New York, USA, 1980.
10. Robert J Schilling, Sandra L Harries. Applied Numerical Methods for Engineers, Brooks /Cole, Pacific Grove, CA, 2000.