

# Calculus of Finite Difference and Representation of Numerical Data by Mathematical Curve: A Brief Review

**Dr. Biswajit Das**

Department of Business Mathematics and Statistics, Chhaygaon College, Chhaygaon.

## **Abstract**

Some formulas have been derived / developed for mathematical representation of numerical data on a pair of variables by suitable mathematical equation / mathematical curve applying some commonly used operations in calculus of finite difference namely usual algebraic operation, forward difference operation, backward difference operation, divided difference operation, backward divided difference operation, difference & ratio operation and backward difference & ratio operation. This paper is a brief review on these recent developments of the formulas for representing numerical data on a pair of variables by mathematical curve.

**Keywords:** Calculus of Finite Difference, Pair of Variables , Numerical Data, Mathematical Representation

## **1. Introduction:**

Observations or data, collected from experiment or survey, normally suffer from various types of errors / causes which can be broadly divided into two types namely (1) Assignable error / cause that is avoidable / controllable & (2) Chance error / cause that is unavoidable / uncontrollable Even if all the assignable causes of errors are controlled or eliminated, observations still do not become free from error. Each of them still suffers from error which occurs due to unknown and unintentional cause that is nothing but the chance cause. Consequently the findings obtained by analyzing the observations which are free from the assignable errors are also subject to errors due to the presence of chance error in the observations. Determination of constant(s) associated to mathematical model(s), in different situations, based on the observations is also subject to error due to the same reason.

A number of mathematical models have been identified for describing the association of chance error(s) in determining constant(s) in some distinct situations where observations/data are of measurement type There are two broad aspects of statistical determination of parameters involved in the respective models describing the dependence of the dependent variable on the independent variable(s). One of them is based on the basic philosophy behind statistics which consists of determining the parameter(s) from numerical data compromising with some degree of error in findings. Several statistical methods have already been developed for determination (estimation in statistical literature) of such parameter(s) which are available in the standard literatures in statistics. However, existing statistical methods of estimation cannot normally yield error free estimate(s) of parameters The same fact happens in the case of some recently developed methods of estimation Recently, some studies have been made on attempting of

determining error free estimates of such parameter(s) along with attempting of application in the case of real data. These studies have been done on the basis of various measures of average namely arithmetic mean, geometric mean, harmonic mean, median, generalized mean and others.

Several statistical methods like the least squares method, orthogonal polynomial method etc have already been developed for representing numerical data on a pair of variables by polynomial curve. Some studies, extension in nature of the earlier ones, have recently been done on statistical representation of numerical data by some special mathematical curves namely linear curve, quadratic curve, exponential curve etc. In some recent studies some formulas / methods have been developed for representing numerical data on a pair of variables by polynomial curve in connection with the development of some more convenient formulas/methods of interpolation which is a technique of estimating approximately the value of the dependent variable corresponding to a value of the independent variable lying between its two extreme values on the basis of the given values of the independent and the dependent variables. The formulas developed are based on some commonly used operations in calculus of finite difference namely usual algebraic operation, forward difference operation, backward difference operation, divided difference operation, backward divided difference operation, difference & ratio operation and backward difference & ratio operation. On the other hand, the methods developed are based on matrix inversion by Cayley-Hamilton Theorem, matrix inversion by Gauss Jordan method and matrix inversion by elementary column transformation have been developed for mathematical representation of numerical data on a pair of variables by suitable mathematical equation / mathematical curve in connection with the development of some more meritorious formula / method of interpolation. This paper is a brief review on the recent developments of the formulas, based on these commonly used operations in calculus of finite difference, for representing numerical data on a pair of variables by mathematical curve.

**2. Representation of Numerical Data:**

Formulas of representation of numerical data on a pair of variables by suitable mathematical curve, as mentioned above, have been outlined below:

**2.1. Formula Based on Forward Difference Operation:**

Let us consider the situation where the argument  $x$  assume the values which are equally spaced i.e.  $x$  assume the values at equal interval and let the length of the interval be  $h$ .

Thus,  $x_{i+1} - x_i = h, (i = 0, 1, 2, \dots, n - 1)$

Let us define a variable  $u$  by  $u = \frac{x-x_0}{h}$

Then the formula for representing a given set of numerical data on a pair of variables by a suitable polynomial curve in the situation where the argument assumes the values which are equally is given by

$$y = f(x) = \alpha_0 x^0 + \alpha_1 x^1 + \alpha_2 x^2 + \dots + \alpha_n x^n \tag{1}$$

where

$$\alpha_0 = A_0 - A_1 x_0 + A_2 x_0 x_1 - A_3 x_0 x_1 x_2 + A_4 x_0 x_1 x_2 x_3 - \dots + A_n (-1)^n (x_0 x_1 x_2 x_3 \dots x_{n-1}),$$

$$\alpha_1 = A_1 - A_2 (\sum_{i=0}^1 x_i) + A_3 (\sum_{i=0}^1 \sum_{j=1}^2 x_i x_j) - A_4 (\sum_{i=0}^1 \sum_{j=1}^2 \sum_{k=2}^3 x_i x_j x_k) + \dots + (-1)^n A_n (x_0 x_1 x_2 x_3 \dots x_{n-2})$$

$$\alpha_2 = A_2 - A_3 \left( \sum_{i=0}^2 x_i \right) + A_4 \left( \sum_{i=0}^2 \sum_{j=1}^3 x_i x_j \right) - \dots + (-1)^n A_n (x_0 x_1 x_2 x_3 \dots x_{n-3} + x_0 x_1 x_2 x_3 \dots x_{n-2} + x_0 x_1 x_2 x_3 \dots x_{n-1}),$$

$$\alpha_i = A_i - A_{i+1} \left( \sum_{j=0}^i x_j \right) + \dots + (-1)^{n-i} A_n (x_0 x_1 x_2 x_3 \dots x_{n-i+1} \dots + x_0 x_1 x_2 x_3 \dots x_{n-1}),$$

$$\alpha_n = A_n$$

with

$$A_0 = y_0, A_1 = \frac{\Delta y_0}{h}, A_2 = \frac{\Delta^2 y_0}{2!h^2}, \dots, A_i = \frac{\Delta^{i-1} y_0}{(i-1)!h^{i-1}}, \dots, A_n = \frac{\Delta^n y_0}{n!h^n}$$

(Das & Chakrabarty, 2016c).

### 2.2. Formula Based on Backward Difference Operation:

As in the earlier case, here also let us consider the situation where the argument  $x$  assume the values which are equally spaced i.e.  $x$  assume the values at equal interval and let the length of the interval be  $h$ .

$$\text{Thus, } x_{i+1} - x_i = h, \quad (i = 0, 1, 2, \dots, n-1)$$

Let us define a variable  $v$  by  $v = \frac{x-x_n}{h}$

The formula for representing a given set of numerical data on a pair of variables by a suitable polynomial curve in the situation is given by

$$y = f(x) = \beta_0 x^0 + \beta_1 x^1 + \beta_2 x^2 + \dots + \beta_n x^n \tag{2}$$

where

$$\begin{aligned} \beta_0 &= B_n - B_{n-1} x_n + B_{n-2} x_n x_{n-1} - B_{n-3} x_{n-2} x_{n-1} x_n + \dots \\ &\quad - (-1)^n B_0 (x_n x_{n-1} x_{n-2} x_{n-3} \dots x_1) \\ \beta_1 &= B_{n-1} - B_{n-2} \left( \sum_{i=n-1}^n x_i \right) + B_{n-3} \left( \sum_{i=n-2}^{n-1} \sum_{j=n-1}^n x_i x_j \right) - B_{n-4} \\ &\quad \left( \sum_{i=n-3}^{n-2} \sum_{j=n-2}^{n-1} \sum_{k=n-1}^n x_i x_j x_k \right) + \dots - B_0 \left( \sum_{i=1}^2 \sum_{j=2}^1 \sum_{k=1}^n x_i x_j x_k \right) \\ \beta_2 &= B_{n-2} - B_{n-3} \left( \sum_{i=n-2}^n x_i \right) + B_{n-4} \left( \sum_{i=n-3}^{n-1} \sum_{j=n-2}^n x_i x_j \right) - \dots + \\ &\quad B_0 \left( \sum_{i=2}^3 \sum_{j=3}^2 \sum_{k=2}^1 \sum_{l=1}^n x_i x_j x_k x_l \right) \end{aligned}$$

$$\beta_n = B_n$$

With

$$B_n = f(x_n), B_{n-1} = \frac{\nabla f(x_{n-1})}{h}, B_{n-2} = \frac{\nabla^2 f(x_{n-2})}{2!h^2}, \dots, B_0 = \frac{\nabla^n f(x_0)}{n!h^n}$$

### 2.3. Formula Based on Divided Difference Operation:

The formula, Newton's divided difference interpolation formula [Chwaiger (1994), De Boor (2003), Gertrude (1954) et al], for representing a given set of numerical data on a pair of variables by a polynomial curve in this situation where the argument assume the values which are not necessarily equally spaced is

$$y = f(x) = \gamma_0 x^0 + \gamma_1 x^1 + \gamma_2 x^2 + \dots + \gamma_n x^n \tag{3}$$

where

$$\begin{aligned} \gamma_0 &= D_0 - D_1 x_0 + D_2 x_0 x_1 - D_3 x_0 x_1 x_2 + D_4 x_0 x_1 x_2 x_3 - \dots + D_n (-1)^n \prod_{i=0}^{n-1} x_i \\ \gamma_1 &= D_1 - D_2 (\sum_{i=0}^1 x_i) + D_3 (\sum_{i=0}^1 \sum_{j=1}^2 x_i x_j) - D_4 (\sum_{i=0}^1 \sum_{j=1}^2 \sum_{k=2}^3 x_i x_j x_k) + \\ &\dots + (-1)^n D_n (\prod_{i=0}^{n-2} x_i + \prod_{i=0}^{n-1} x_i) \\ \gamma_2 &= D_2 - D_3 (\sum_{i=0}^2 x_i) + D_4 (\sum_{i=0}^2 \sum_{j=1}^3 x_i x_j) - \dots - (-1)^n D_n (\prod_{i=0}^{n-3} x_i + \prod_{i=0}^{n-2} x_i + \prod_{i=0}^{n-1} x_i) \\ &\dots \\ \gamma_i &= D_i - D_{i+1} (\sum_{j=0}^i x_j) + \dots + (-1)^{n-i} D_n (\prod_{i=0}^{n-i} x_i + \dots + \prod_{i=0}^{n-1} x_i) \\ &\dots \\ \gamma_n &= D_n \end{aligned}$$

with

$$D_0 = f(x_0), D_1 = f(x_0, x_1), D_2 = f(x_0, x_1, x_2), \dots, D_n = f(x_0, x_1, x_2, x_3, \dots, x_n)$$

### 2.4. Formula Based on Algebraic Operation:

The formula, obtained from Lagrange’s interpolation formula [Whittaker & Robinson (1967), Echols (1893), Mills (1977) et al], for representing a given set of numerical data on a pair of variables by a mathematical curve in the situation where the argument assume the values which are not necessarily equally spaced is

$$y = f(x) = \rho_0 x^0 + \rho_1 x^1 + \rho_2 x^2 + \dots + \rho_n x^n \tag{4}$$

where

$$\begin{aligned} \rho_0 &= \sum_{i=0}^n L_i S_i(n), \rho_1 = \sum_{i=0}^n L_i S_i(n-r), \rho_2 = \sum_{i=0}^n L_i S_i(2), \dots, \\ \rho_{n-r} &= \sum_{i=0}^n L_i S_i(r), \dots, \rho_{n-2} = \sum_{i=0}^n L_i S_i(2), \rho_{n-1} = S_i(1), \rho_n = \sum_{i=0}^n L_i \end{aligned}$$

with

$$L_0 = \frac{f(x_0)}{\prod_{i=0}^n (x_0 - x_i)}, L_1 = \frac{f(x_1)}{\prod_{i=0}^n (x_1 - x_i)}, \dots, L_n = \frac{f(x_n)}{\prod_{i=0}^{n-1} (x_{n-1} - x_i)}$$

$$i \neq 0, \quad i \neq 1 \quad i \neq n - 1$$

and

$$S_r = \sum_{i=0}^n x_i, \quad S_r(p) = \sum_{i=0}^n \sum_{j=0}^n x_i x_j$$

$$i \neq r, \quad i \neq r, j \neq r, i \neq j$$

$$S_{r_1 r_2 \dots r_p} = \sum_{i_1=0}^n \sum_{i_2=0}^n \dots \sum_{i_p=0}^n x_{i_1} x_{i_2} \dots x_{i_p}$$

$$i_1 \neq r_1, i_2 \neq r_1 \dots i_p \neq r_p$$

$$i_1 \neq i_2 \neq \dots \neq i_p$$

### 2.5. Formula Based on Backward Divided Difference Operation:

The formula, based on backward divided difference operation, for representing a given set of numerical data on a pair of variables by a polynomial curve in the situation where the argument assume the values which are not necessarily equally spaced is

$$y = f(x) = \delta_0 x^0 + \delta_1 x^1 + \delta_2 x^2 + \dots + \delta_n x^n \tag{5}$$

where

$$\begin{aligned} \delta_n &= E_n - E_{n-1} x_n + E_{n-2} x_{n-1} x_n - E_{n-3} x_{n-2} x_{n-1} x_n + \dots \\ &+ (-1)^n E_0 (x_{n-1} x_{n-2} x_{n-3} \dots x_1 x_0) \end{aligned}$$

$$\delta_{n-1} = E_{n-1} - E_{n-2} (\sum_{i=n-1}^n x_i) + E_{n-3} (\sum_{i=n-1}^{n-1} \sum_{j=n-1}^{n-2} x_i x_j) - \dots$$

$$\begin{aligned}
 &+ (-1)^n C_0 (x_n x_{n-1} x_{n-2} x_{n-3} \dots x_1 x_2 + x_n x_{n-1} x_{n-2} x_{n-3} \dots x_1 x_0) \\
 \delta_{n-2} = &E_{n-2} - E_{n-3} \left( \sum_{i=n-2}^n x_i \right) + E_{n-4} \left( \sum_{i=n-2}^n \sum_{j=n-3}^{n-1} x_i x_j \right) - \dots + \\
 &(-1)^n E_0 (x_n x_{n-1} x_{n-2} x_{n-3} \dots x_3 + x_{n-1} x_{n-2} x_{n-3} \dots x_2 \\
 &+ x_n x_{n-1} x_{n-2} x_{n-3} \dots x_1) \\
 \dots \dots \dots \\
 \delta_{n-i} = &E_{n-i} - E_{n-(i+1)} \left( \sum_{i=n-1}^n x_i \right) + \dots + (-1)^{n-i} E_0 (x_n x_{n-1} x_{n-2} \dots x_{n-(i+1)} \\
 &+ \dots + x_n x_{n-1} x_{n-2} \dots x_1) \\
 \dots \dots \dots \\
 \delta_0 = &E_0 \\
 \text{with} \\
 E_n = &f(x_n), E_{n-1} = f(x_n, x_{n-1}), E_{n-2} = f(x_n, x_{n-1}, x_{n-2}), \dots, \\
 E_{n-i} = &(x_n, x_{n-1}, x_{n-2}, x_{n-3}, \dots, x_{n-i}) \\
 \dots, &E_0 = f(x_n, x_{n-1}, x_{n-2}, x_{n-3}, \dots, x_0)
 \end{aligned}$$

**3. Application to Numerical Data:**

The following table shows the number of persons in the population of India:

**Table – 3(i)**  
(Number of persons in the population of India)

Year	2000	2005	2010	2015	2020
Number of persons	1057922733	1154676322	1243481564	1328024498	1402617695

Taking 2000 as origin and changing scale by 1/5, one can obtain the following table for the values of the argument  $x$  (representing year) and the entry  $y = f(x)$  (representing the number of persons in the population of India):

**Table – 3(ii)**  
(Values of argument and entry of Number of persons in the population of India)

Year ( $i$ )	2000	2005	2010	2015	2020
Value of $x_i$	0	1	2	3	4
Value of $y_i = f(x_i)$ (in thousands)	105792273 3	115467632 2	124348156 4	1328024498	1402617695

Here,  $x_0 = 0, x_1 = 1, x_2 = 2, x_3 = 3, x_4 = 4$   
 $f(x_0) = 1057922733, f(x_1) = 1154676322, f(x_2) = 1243481564,$   
 $f(x_3) = 1328024498, f(x_4) = 1402617695$

**Representation by Forward Difference Formula:**

The values of the parameters of Equation (1) have been obtained as  
 $\alpha_0 = 548159.652, \alpha_1 = 119214.356, \alpha_2 = 16547.583, \alpha_3 = -375.487$   
 &  $\alpha_4 = -217.007$

Consequently, the equation of the polynomial that can represent the given numerical data becomes  $f(x) = 548159.652 + 119214.356 x + 16547.583 x^2 - 375.487 x^3 - 217.007 x^4$

**Representation by Backward Difference Formula:**

The values of the parameters of Equation (2) have been obtained as

$$\beta_0 = 548159.652, \beta_1 = 119214.35634, \beta_2 = 16547.58219, \beta_3 = -375.48616$$

$$\beta_4 = -217.00725$$

Consequently, the equation of the polynomial that can represent the given numerical data becomes  $f(x) = 548159.652 + 119214.35634 x + 16547.58219 x^2 - 375.48616 x^3 - 217.00725 x^4$

**Representation by Divided Difference Formula:**

The values of the parameters of Equation (3) have been obtained as

$$\gamma_0 = 548159.652, \gamma_1 = 119214.35618, \gamma_2 = 16547.58223, \gamma_3 = -375.48616$$

$$\& \gamma_4 = -217.00725$$

Consequently, the equation of the polynomial that can represent the given numerical data becomes  $f(x) = 548159.652 + 119214.35618 x + 16547.58223 x^2 - 375.48616 x^3 - 217.00725 x^4$

**Representation by Algebraic Formula:**

The values of the parameters of Equation (1) have been obtained as

$$\rho_0 = 548159.652, \rho_1 = -119214.356, \rho_2 = 16547.5823, \rho_3 = 375.4862$$

$$\& \rho_4 = -217.00725$$

Therefore, the polynomial that can represent the given numerical data becomes

$$f(x) = -217.00725 x^4 - 375.4862 x^3 + 16547.5823 x^2 + 119214.356 x + 548159.652$$

**Representation by Backward Divided Difference Formula:**

The values of the parameters of Equation (5) have been obtained as

$$\delta_4 = 1210193.422, \delta_3 = -178017.82218, \delta_2 = -8790.94773, \delta_1 = 3847.60216, \delta_0 = -217.00725$$

Consequently, the equation of the polynomial that can represent the given numerical data becomes

$$f(x) = 1210193.422 - 178017.82218 x - 8790.94773 x^2 + 3847.60216 x^3 - 217.00725 x^4$$

**4. Conclusion:**

Each of the formulas yields the same values of the function  $f(x)$  corresponding to the respective observed values. This implies that the formulas, mentioned above, are equivalent.

It is to be noted that the degree of the polynomial is one less than the number of pairs of observations. Finally, from the meaning of research [Chakrabarty (2018a, 2018g, 2019c)], it can be concluded that the development of the formulas for representing numerical data on a pair of variables by mathematical curve, as described above, can be regarded as research findings carrying fundamental importance and high significance in the theory of mathematical representation of data.

**References:**

1. Echols W. H., On Some Forms of Lagrange's Interpolation Formula, *The Annals of Mathematics*, Jstor, 8(1/6), 22—24.
2. Jan K. Wisniewski, Note on Interpolation (in Notes), “*Journal of the American Statistical Association*”, Jstor 25(170), 203—205, 1930.
3. Corliss J. J., Note on an Extension of Lagrange's Formula, *American Mathematical Monthly*, Jstor, 45(2), 106—107, 1938.
4. Hummel P. M., A Note on Interpolation (in Mathematical Notes), *American Mathematical Monthly*, Jstor, 54(4), 218—219, 1947.
5. Gertrude Blanch, On Modified Divided Differences, *Mathematical Tables and Other Aids to Computation*, Vol. 8, No. 45, 1-11, 1954.
6. Herbert E. Salzer, Multi-Point Generalization of Newton's Divided Difference Formula, *Proceedings of the American Mathematical Society*, Jstor, 13(2), 210—212, 1962.
7. Traub J. F., On Lagrange-Hermite Interpolation, *Journal of the Society for Industrial and Applied Mathematics*, Jstor, 12(4), 886—891, 1964.
8. Jordan C., *Calculus of Finite Differences*, 3rd ed, New York, Chelsea, 1965.
9. Quadling D. A., Lagrange's Interpolation Formula, *The Mathematical Gazette*, L(374), 372—375, 1966.
10. Whittaker E. T. and Robinson G., The Gregory-Newton Formula of Interpolation and An Alternative Form of the Gregory-Newton Formula, *The Calculus of Observations: A Treatise on Numerical Mathematics*, 4th ed, pp. 10-15, 1967.
11. Whittaker E. T. & Robinson G., Divided Differences & Theorems on Divided Differences, *The Calculus of Observations: A Treatise on Numerical Mathematics*, 4<sup>th</sup> ed., New York, 20-24, 1967.
12. Abramowitz M., *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th printing, New York: Dover, p. 880, 1972.
13. Bathe K. J. & Wilson E. L., *Numerical Methods in Finite Element Analysis*, Prentice-Hall, Englewood Cliffs, NJ., 1976.
14. Mills T. M., An introduction to analysis by Lagrange interpolation, *Austral. Math. Soc. Gaz.*, 4(1), MathSciNet, 10—18, 1977.
15. Dokken T & Lyche T., A divided difference formula for the error in Hermite interpolation, *BIT*, 19, 540—541, 1979.
16. Fred T., Recurrence Relations for Computing with Modified Divided Differences, *Mathematics of Computation*, Vol. 33, No. 148, 1265-1271, 1979.
17. Conte S.D. & Carl de Boor, *Elementary Numerical Analysis*, 3 Ed, McGraw-Hill, New York, USA., 1980.
18. Jeffreys H. & Jeffreys B.S., Divided Differences, *Methods of Mathematical Physics*, 3<sup>rd</sup> ed, 260-264, 1988.
19. Lee E.T.Y., A Remark on Divided Differences, *American Mathematical Monthly*, Vol. 96, No 7, 618-622, 1989.
20. Vertesi P., *SIAM Journal on Numerical Analysis*, Vol. 27, No. 5, pp. 1322-1331, Jstor, 1990.
21. David R. Kincaid & E. Ward Chaney, *Numerical analysis*, Brooks /Cole, Pacific Grove, CA, 1991.
22. Chwaiger J., On a Characterization of Polynomials by Divided Differences, *Aequationes Math*, 48, 317-323, 1994.

23. C. F. Gerald & P. O. Wheatley , *Applied Numerical Analysis*, fifth ed., Addison-Wesley Pub. Co., MA., 1994.
24. James B. Scarborough ,*Numerical Mathematical Analysis*, 6 Ed, The John Hopkins Press, USA, 1996.
25. Robert J. Schilling& Sandra L. Harries ,*Applied Numerical Methods for Engineers*, Brooks /Cole, Pacific Grove, CA, 2000.
26. Revers & Michael Bull, On Lagrange interpolation with equally spaced nodes, *Austral. Math.SocMathSciNet.*, 62(3), 357—368, 2000.
27. S.C. Chapra& R.P. Canale,*Numerical Methods for Engineers*, third ed., McGraw-Hill, NewYork, 2002.
28. Floater M, Error formulas for divided difference expansions and numerical differentiation, *J. Approx. Theory*, 122, 1–9, 2003.
29. De Boor C, A divided difference expansion of a divided difference, *J. Approx. Theory* ,122 , 10–12, 2003.
30. Endre S & David Mayers,*An Introduction to Numerical Analysis*, Cambridge, UK, 2003.
31. Wang X & Wang H, Some results on numerical divided difference formulas, [www.math.uga.edu/~mjlai/pub.html](http://www.math.uga.edu/~mjlai/pub.html), 2003.
32. John H. Mathews &KurtisD.Fink,Numerical methods using MATLAB, 4ed, Pearson Education, USA., 2004.
33. Wang X & Yang S, On divided differences of the remainder of polynomial interpolation, [www.math.uga.edu/~mjlai/pub.html](http://www.math.uga.edu/~mjlai/pub.html), 2004.