ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 5, Issue 6 , June 2018

Numerical Data on a Pair of Variables: Representation by Logistic Curve

Biswajit Das, Dhritikesh Chakrabarty

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

ABSTRACT: A set of available numerical data on a pair of variables are required to be represented by different types of mathematical curves in different situations. In some situation, it is required to represent the data as mentioned by a logistic curve. In this study, attempt has been made on the representation of numerical data by this type of mathematical curve namely logistic curve. This paper describes (i) the method of representing numerical data on a pair of variables by logistic curve and (ii) application of the method in representing total population of India by a logistic curve as an example of showing the application of the method to numerical data.

KEYWORDS: Pair of variables, numerical data, representation by logistic curve, total population of India

I. INTRODUCTION

In the existing formulae for numerical interpolation {Hummel (1947), Erdos & Turan (1938), Bathe & Wilson (1976), Jan (1930), Hummel (1947) et al}, 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 then 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 is necessary for representing a given set of numerical data on a pair of variables by a suitable polynomial. Das & Chakrabarty (2016*a*, 2016*b*, 2016*c* & 2016*d*) derived four formulae for representing numerical data on a pair of variables by a polynomial curve. They derived the formulae from Lagranges Interpolation Formula, Newton's Divided Difference Interpolation Formula, Newton's Forward Interpolation Formula and Newton's Backward Interpolation respectively. In another study, Das & Chakrabarty (2016*e*) derived one method for representing numerical data on a pair of variables by a polynomial curve. The method derived is based on the inversion of a square matrix by Caley-Hamilton theorem on characteristic equation of matrix *[Cayley (1858 , 1889) & Hamilton (1864a , 1864b , 1862)]*. In a separate study Das & Chakrabarty (2016*f* , 2016*g*) composed two methods, based on the inversion of matrix by elementary transformation, for the same purpose. The studies, made so far, are on the representation of numerical data on a pair of variables by polynomial curve It is be possible to represent the numerical data on a pair of variables by non-polynomial curve besides the representation of the said data by polynomial curve. Some methods have already been developed for representing the said data by polynomial curve [Das & Chakrabarty (2016*a*, 2016*b*, 2016*c*, 2016*d* , 2016*e* , 2016*f* , 2017*a*, 2017*b*)], exponential curve [Das & Chakrabarty (2017*c*)] and modified exponential curve [Das & Chakrabarty (2017*d*)]. In this study, attempt has been made on the representation of numerical data by a special non-polynomial curve namely logistic curve. This paper describes (i) the method of representing numerical data on a pair of variables by logistic curve and (ii) application of the method in representing total population of India by a logistic curve as an example of showing the application of the method to numerical data.

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 5, Issue 6 , June 2018

II. REPRESENTATION OF NUMERICAL DATA BY LOGISTIC CURVE:

The Logistic curve is of the form

$$y = \frac{A}{B + C^x} \quad (1)$$

where A , B and C are parameters

Let

$$y_0\ ,\ y_1\ ,\ y_2$$

be the values of y corresponding to the values of x_0 , x_1 and x_2 of x respectively. Then the points

$$(x_0, y_0), (x_1, y_1), (x_2, y_2)$$

lie on the curve described by equation (2.1).

Therefore,

$$y_0 = \frac{A}{B + C^{x_0}} \quad (2)$$

$$y_1 = \frac{A}{B + C^{x_1}} \quad (3)$$

$$y_2 = \frac{A}{B + C^{x_2}} \quad (4)$$

$$\therefore\ \frac{1}{y_0} = \frac{B}{A} + \frac{1}{A} c^{x_0} \quad (5)$$

$$\frac{1}{y_1} = \frac{B}{A} + \frac{1}{A} c^{x_1} \quad (6)$$

$$\frac{1}{y_2} = \frac{B}{A} + \frac{1}{A} c^{x_2} \quad (7)$$

Equation (6) – Equation (5) $\Rightarrow \Delta\left(\frac{1}{y_0}\right) = \frac{1}{A}(c^{x_1} - c^{x_0})$ (8)

Equation (7) – Equation (6) $\Rightarrow \Delta\left(\frac{1}{y_1}\right) = \frac{1}{A}(c^{x_2} - c^{x_1})$ (9)

If x_0 , x_1 , x_2 are equally spaced then

$$x_1 - x_0 = x_2 - x_1 = h$$

i.e. $x_1 = x_0 + h \qquad x_2 = x_0 + 2h$

This means,

$$y_0 = \frac{A}{B + C^{x_0}} \quad (10)$$

$$y_1 = \frac{A}{B + C^{x_0 + h}} \quad (11)$$

$$y_2 = \frac{A}{B + C^{x_0 + 2h}} \quad (12)$$

$$\therefore\ \frac{1}{y_0} = \frac{B + C^{x_0}}{A} \quad (13)$$

$$\frac{1}{y_1} = \frac{B + C^{x_0 + h}}{A} \quad (14)$$

$$\frac{1}{y_2} = \frac{B + C^{x_0 + 2h}}{A} \quad (15)$$

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 5, Issue 6 , June 2018

Accordingly,

$$\Delta\left(\frac{1}{y_0}\right) = \frac{1}{A}\, c^{x_0}\, (c^h - 1) \qquad (16)$$

$$\&\ \Delta\left(\frac{1}{y_1}\right) = \frac{1}{A} c^{x_1} (c^h - 1) = \frac{1}{A} c^{x_0 + h} (c^h - 1) = \frac{c^h}{A} c^{x_0} (c^h - 1) \qquad (17)$$

which implies,

$$\frac{\Delta\left(\frac{1}{y_1}\right)}{\Delta\left(\frac{1}{y_0}\right)} = C^h$$

$$\text{i.e. } \log \left\{ \frac{\Delta\left(\frac{1}{y_1}\right)}{\Delta\left(\frac{1}{y_0}\right)} \right\} = h \log C$$

$$\text{i.e. } \log C = \frac{1}{h} \log \left\{ \frac{\Delta\left(\frac{1}{y_1}\right)}{\Delta\left(\frac{1}{y_0}\right)} \right\} \qquad (18)$$

$$\text{i.e. } C = \text{antilog} \left[\frac{1}{h} \log \left\{ \frac{\Delta\left(\frac{1}{y_1}\right)}{\Delta\left(\frac{1}{y_0}\right)} \right\} \right] \qquad (19)$$

From equation (16) & equation (17), expression for A is obtained as

$$A = \frac{c^{x_0} (c^h - 1)}{\Delta\left(\frac{1}{y_0}\right)} = \frac{c^{x_1} (c^h - 1)}{\Delta\left(\frac{1}{y_1}\right)} \qquad (20)$$

where C is given by equation (2.19).
Again, from equations (10), (11) & (12), expression for B is obtained as

$$\therefore B = \frac{A}{y_0} - c^{x_0} = \frac{A}{y_1} - c^{x_1} = \frac{A}{y_2} - c^{x_2} \qquad (21)$$

where A & C are given by equations (20) & (19) respectively.

III. EXAMPLE OF DATA REPRESENTATION BY LOGISTIC CURVE:

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

Year	1951	1961	1971	1981	1991	2011
Total Population	361088090	439234771	548159652	683329097	846302688	1210193422

Let us first represent the total populations corresponding to the years 1951, 1961 & 1971 by the logistic curve described by equation (1).
In order to do it let us take the year 1951 as origin (i.e 0) and choose the scale 1/10 such that the value of x corresponding to the year 1961 becomes 1 and 1971 becomes 2. Thus we have the following table:

Table –1

Year	x	Total Population $f(x) = y$	$\frac{1}{y_i}$	$\Delta\left(\frac{1}{y_i}\right)$	$\frac{\Delta\left(\frac{1}{y_1}\right)}{\Delta\left(\frac{1}{y_0}\right)}$
1951	0	361088090	0.000000002769407321	– 0.000000000492720533	0.9181690619
1961	1	439234771	0.000000002276686788	– 0.0000000004524007496	
1971	2	548159652	0.0000000018242860384		

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 5, Issue 6 , June 2018

Now,

$C = 0.9181690619$

Accordingly,

$$A = \frac{c^{x_0}(c^h - 1)}{\Delta(\frac{1}{y_0})} = \frac{0.9181690619 - 1}{-0.000000000492720533} = \frac{1}{0.000000000492720533} - \frac{0.9181690619}{0.000000000492720533}$$

$$= 2029548056.2000447421 - 1863468234.8421635596$$
$$= 166079821.3578811825$$

$$\& \ B = \frac{A}{y_0} - c^{x_0} = \frac{166079821.3578811825}{361088090} - 1 = -0.5400573269$$

Therefore, the logistic curve that can represent the given data is

$$y = \frac{166079821.3578811825}{-0.5400573269 + (0.9181690619)^x}$$

This curve yields,

$$y_0 = \frac{A}{B + C^{x_0}} = \frac{166079821.3578811825}{-0.5400573269 + 1} = 361088090$$

$$y_1 = \frac{A}{B + C^{x_1}} = \frac{166079821.3578811825}{-0.5400573269 + 0.9181690619} = 439234771$$

$$y_2 = \frac{A}{B + C^{x_2}} = \frac{166079821.3578811825}{-0.5400573269 + (0.9181690619)^2} = \frac{166079821.3578811825}{-0.5400573269 + 0.8430344262} = 548159652$$

The data on total population corresponding to three consecutive years (at a gap of 10 years) can be represented by the logistic curve. The curves obtained have been shown in the following table.

Table – 2

Years	x	Logistic Curve	Quadratic curve	Modified exponential curve
1951 1961 1971	0 1 2	$y = \frac{166079821.3578811825}{-0.5400573269 + (0.9181690619)^x}$	$y = 361088090 + 62757581\,x + 15389100\,x^2$	$y = 162671556.48 + 198416533.52 \times (1.3938516595)^x$
1961 1971 1981	0 1 2	$y = \frac{447256612.3705644717}{0.0182632202 + (0.7976607733)^x}$	$y = 439234771 + 106280000\,x + 26445564\,x^2$	$y = -12844741.85 + 452079512.85 \times (1.2409418651)^x$
1971 1981 1991	0 1 2	$y = \frac{607034035.4099546391}{0.1074037156 + (0.7809442825)^x}$	$y = 548159652 + 107365299\,x + 27804146\,x^2$	$y = -108964607.87 + 657124259.87 \times (1.2056984549)^x$
1981 1991 2001	0 1 2	$y = \frac{930490540.3659358186}{0.36170191551 + (0.7377753206)^x}$	$y = 683329097 + 161199695\,x + 17738968\,x^2$	$y = -813961579.52 + 1497290676.52 \times (1.1088456594)^x$

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 5, Issue 6 , June 2018

1991 2001 2011	0 1 2	$y = \frac{1400316634.9560702305}{0.6546286037 + (0.7088532327)^x}$	$y = 846302688 + 178246943\,x + 2465616\,x^2$	$y = -12398675095.11 + 13244977783.11 \times (1.0136438552)^x$

Similarly, the data on total population corresponding to two consecutive years (at a gap of 30 years) can be represented by the logistic curve. The curves obtained have been shown in the following table:

Table – 3
(Mathematical curve representing total population at a gap of 30 years)

Years	x	Logistic Curve	Quadratic curve	Modified exponential curve
1951 1981 2011	0 1 2	$y = \frac{392165083.5070385244}{0.0860648533 + (0.4878388268)^x}$	$y = 361088090 + 117617689\,x + 204623318\,x^2$	$y = -146377372.04 + 507465462.04 \times (1.6350008643)^x$

IV. CONCLUSION

The method of representing numerical data on a pair of variables by logistic curve described by equation (2.1) has been discussed in section II. Its application to the total population of India has been shown in section III.

It is to be mentioned here that the method discussed here is applicable in representing a set of numerical data on a pair of variables if the given values of the independent variable are equidistant. In the case where the given values of the independent variable are not equidistant, the method fails to represent the given data by logistic curve.

The logistic curve contains three parameters. Accordingly, when three pairs of numerical data are available, they can be represented by logistic curve. However, when three pairs of numerical data are available then they can also be represented by quadratic curve and by modified exponential curve since each of these two curves also contains three parameters. Thus, one question arises- which of the three curves will suit the entire data best among the three ones.

The quadratic curve is of the form

$$y = \alpha x^2 + \beta x + \gamma \qquad (22)$$

where α , β , γ are the parameters of the curve.

In order to represent the given three pairs on numerical data by this curve, we have

$$y_0 = \alpha\, x_0^2 + \beta x_0 + \gamma \qquad (23)$$

$$y_1 = \alpha\,(x_0 + h)^2 + \beta\,(x_0 + h) + \gamma \qquad (24)$$

$$y_2 = \alpha\,(x_0 + 2h)^2 + \beta\,(x_0 + 2h) + \gamma \qquad (25)$$

Equation (24) – Equation (23) $\Rightarrow$

$$\Delta\, y_0 = \beta h + 2\alpha x_0 h + h^2 \qquad (26)$$

Equation (25) – Equation (24) $\Rightarrow$

$$\Delta\, y_1 = \beta h + 2\alpha x_0 h + 3h^2\alpha \qquad (27)$$

Equation (27) – Equation (26) $\Rightarrow$ $\Delta^2 y_0 = 2h^2\alpha$

Therefore, $$\alpha = \frac{\Delta^2 y_0}{2h^2} \qquad (28)$$

Consequently from equations (26) & (27) , β is obtained as

$$\beta = \frac{\Delta\, y_0}{h} - \frac{\Delta^2 y_0}{2h^2}\left(\frac{x_0}{h} + \frac{1}{2}\right) \qquad (29)$$

$$\beta = \frac{\Delta\, y_1}{h} - \frac{\Delta^2 y_0}{2h^2}\left(\frac{x_0}{h} + \frac{3}{2}\right) \qquad (30)$$

Thus, the quadratic curve described by equation (22) satisfies the property that

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 5, Issue 6 , June 2018

$$\Delta^2 y_i = \text{constant} \tag{31}$$

On the other hand, the modified exponential curve described by the equation

$$y = a + bc^x$$

where a , b & c are parameters; satisfies the property that

$$\frac{\Delta y_{i+1}}{\Delta y_i} = \text{constant} \tag{32}$$

Again, the logistic curve described by the equation (1) satisfies the property that

$$\frac{\Delta\left(\frac{1}{y_1}\right)}{\Delta\left(\frac{1}{y_0}\right)} = \text{constant}$$

Thus, comparing the values of

$$\Delta^2 y_i \quad , \quad \frac{\Delta y_{i+1}}{\Delta y_i} \quad \& \quad \frac{\Delta\left(\frac{1}{y_1}\right)}{\Delta\left(\frac{1}{y_0}\right)}$$

one can determine which of the three curves will suit the entire data best among the three ones.

In Table – 4 & Table – 5, it is observed the constancy property of $\frac{\Delta y_{i+1}}{\Delta y_i}$ is most dominant than that of $\Delta^2 y_i$ & $\frac{\Delta\left(\frac{1}{y_1}\right)}{\Delta\left(\frac{1}{y_0}\right)}$

Thus, logistic curve is most suitable than quadratic curve & modified exponential curve to represent the available data.

Table – 4

Year	X	Total population (y_i)	Δy_i	$\Delta^2 y_i$	$\frac{\Delta y_{i+1}}{\Delta y_i}$
1951	0	361088090	78146681		
1961	1	439234771	108924881	30778200	1.39385165
1971	2	548159652	135169445	26244564	1.24094186
1981	3	683329097	162973591	27804146	1.20569845
1991	4	846302688	180712559	17738968	1.10884565
2001	5	1027015247	183178175	2465616	1.01364385
2011	6	1210193422			
Total	21	5115322967	849105332		

Table – 5

Year	x	Total population (y_i)	$\frac{1}{y_i}$	$\Delta\left(\frac{1}{y_i}\right)$	$\frac{\Delta\left(\frac{1}{y_{i+1}}\right)}{\Delta\left(\frac{1}{y_i}\right)}$
1951	0	361088090	0.000000002769407321	– 0.000000000492720533	
1961	1	439234771	0.000000002276686788	– 0.0000000004524007496	0.9181690619
1971	2	548159652	0.0000000018242860384	– 0.0000000003608623318	0.7976607733
1981	3	683329097	0.0000000014634237066		0.7809442825

1991	4	846302688	0.0000000011816103318	– 0.0000000002818133748	0.7377753206
2001	5	1027015247	0.00000000097369537883	– 0.00000000020791495297	0.7088532327
2011	6	1210193422	0.00000000082631419227	– 0.00000000014738118656	
Total	21	5115322967			

REFERENCES

[1]. Bathe K. J. & Wilson E. L. (1976): "*Numerical Methods in Finite Element Analysis*", Prentice-Hall, Englewood Cliffs, NJ.

[2]. Biswajit Das & Dhritikesh Chakrabarty (2016*a*): "Lagranges Interpolation Formula: Representation of Numerical Data by a Polynomial Curve", *International Journal of Mathematics Trend and Technology* (ISSN (online): 2231-5373, ISSN (print): 2349-5758), 34 part-1 (2), 23-31.

[3]. Biswajit Das & Dhritikesh Chakrabarty (2016*b*): "Newton's Divided Difference Interpolation Formula: Representation of Numerical Data by a Polynomial Curve", *International Journal of Mathematics Trend and Technology* (ISSN (online): 2231-5373, ISSN (print): 2349-5758), 35 part-1 (3), 26-32.

[4]. Biswajit Das & Dhritikesh Chakrabarty (2016*c*): "Newton's Forward Interpolation Formula: Representation of Numerical Data by a Polynomial Curve", *International Journal of Applied Research* (ISSN (online): 2456 - 1452, ISSN (print): 2456 - 1452), 1 (2), 36-41.

[5]. Biswajit Das & Dhritikesh Chakrabarty (2016*d*): "Newton's Backward Interpolation Formula: Representation of Numerical Data by a Polynomial Curve", *International Journal of Statistics and Applied Mathematics* (ISSN (online): 2394-5869, ISSN (print): 2394-7500), 2 (10), 513-517.

[6]. Biswajit Das & Dhritikesh Chakrabarty (2016*e*): "Matrix Inversion: Representation of Numerical Data by a Polynomial curve", *Aryabhatta Journal of Mathematics & Informatics* (ISSN (print): 0975-7139, ISSN (online): 2394-9309), 8(2), 267-276.

[7]. Biswajit Das & Dhritikesh Chakrabarty (2016*f*): "Inversion of Matrix by Elementary Transformation: Representation of Numerical Data by a Polynomial Curve", *Journal of Mathematis and System Sciences* (ISSN (print): 0975-5454), 12(1 – 2), 27-32.

[8]. Biswajit Das & Dhritikesh Chakrabarty (2017*a*): "Inversion of Matrix by Elementary Column Transformation: Representation of Numerical Data by a Polynomial Curve", *International Journal of Mathematics Trends and Technology* (ISSN (online): 2231-5373, ISSN (print): 2349-5758), 42(1), 45-49

[9]. Biswajit Das & Dhritikesh Chakrabarty (2017*b*): "Backward Divided Difference: Representation of Numerical Data by a Polynomial curve" *International Journal of Statistics and Applied Mathematics*(ISSN (online): 2394-5869, ISSN (print): 2394-7500), 2(2), 01-06

[10]. Biswajit Das & Dhritikesh Chakrabarty (2017*c*): "Representation of Numerical Data by Exponential Curve", *Aryabhatta Journal of Mathematics & Informatics* (ISSN (print): 0975-7139, ISSN (online): 2394-930 9), 9(1), 157-162.

[11]. Biswajit Das & Dhritikesh Chakrabarty (2017*d*): "Representation of Numerical Data by Modified Exponential Curve", *Journal of Mathematis and System Sciences (JMASS)*, ISSN:0975-5454, 13(1-2), 1-5.

[12]. Cayley A. *(1858): "A Memoir on the Theory of Matrices". Phil.Trans.* ***148***.

[13]. *Cayley A. (1889):* "*The Collected Mathematical Papers of Arthur Cayley (Classic Reprint)", Forgotten books.* ASIN B008HUED9O.

[14]. Erdos P. & Turan P. (1938): "On Interpolation II: On the Distribution of the Fundamental Points of Lagrange and Hermite Interpolation", "*The Annals of Mathematics, 2nd Ser*", Jstor, 39(4), 703—724.

[15]. Hummel P. M. (1947): "A Note on Interpolation (in Mathematical Notes)", "*American Mathematical Monthly*", Jstor, 54(4), 218—219.

[16]. *Hamilton W. R. (1864a): "On a New and General Method of Inverting a Linear and Quaternion Function of a Quaternion"*, Proceedings of the Royal Irish Academy, Royal Irish Academy, ***viii****: 182–183.* (communicated on June 9, 1862)

[17]. *Hamilton W. R. (1864b): "On the Existence of a Symbolic and Biquadratic Equation, which is satisfied by the Symbol of Linear Operation in Quaternions"*, Proceedings of the Royal Irish Academy, Royal Irish Academy, ***viii****: 190–101.* (communicated on June 23, 1862)

[18]. *Hamilton W. R. (1862):* "On the Existence of a Symbolic and Biquadratic Equation which is satisfied by the Symbol of Linear or Distributive Operation on a Quaternion",The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science, *series* ***iv***,Taylor & Francis ***24****, 127–128,* ISSN 1478-6435, Retrieved 2015-02-14.

[19]. Jan K. Wisniewski (1930): "Note on Interpolation (in Notes)", "*Journal of the American Statistical Association"*, Jstor 25(170), 203—205.