



MODIFIED NONLINEAR MODEL FOR EXPONENTIAL GROWTH METHOD AND ITS APPLICATION IN BIOSTATISTICS USING SAS

Wan Muhamad Amir W Ahmad*, Nor Azlida Aleng***, Yosza Dasril**, Nurfadhina Abdul Halim***, Ruhaya Hasan*, Zalila Ali****, Adam Baharum**** & Syerrina Zakaria***

* School of Dental Sciences, Health Campus, Universiti Sains Malaysia (USM), Kubang Kerian, Kelantan, Malaysia

** Department of Industrial Electronics, Faculty of Electronics & Computer Engineering, Universiti Teknikal Malaysia Melaka (UTeM), Melaka, Malaysia

*** School of Informatics and Applied Mathematics, Universiti Malaysia Terengganu (UMT), Terengganu Malaysia

**** School of Mathematics Sciences, Universiti Sains Malaysia (USM), Malaysia.

Cite This Article: Wan Muhamad Amir W Ahmad, Nor Azlida Aleng, Yosza Dasril, Nurfadhina Abdul Halim, Ruhaya Hasan, Zalila Ali, Adam Baharum & Syerrina Zakaria, "Modified Nonlinear Model for Exponential Growth Method and Its Application in Biostatistics Using SAS", *International Journal of Multidisciplinary Research and Modern Education*, Volume 3, Issue 1, Page Number 89-94, 2017.

Copy Right: © IJMRME, R&D Modern Research Publication, 2017 (All Rights Reserved). This is an Open Access Article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract:

This paper supplied an alternative method for exponential growth modeling as a technique for regression analysis through SAS algorithm. This alternative method is a combination technique (using bootstrap and fuzzy regression for nonlinear model) for the small data set and gives the researcher an option to launch the analysis even there is not enough data set. This method current method improves the previous methodology with embedded bootstrapping and fuzzy technique to the step of nonlinear regression model. The aim of this principle is to propose an alternative method of doing analysis with better improved results. In our case, we applied this principle to the agriculture data and the gained results were compared by looking at the average width of predicted interval.

Key Words: Bootstrap, Fuzzy, Weighted Robust Regression, Exponential Growth & Biostatistics

Introduction to Algorithm Using SAS Language:

This paper provides a road map of the practical approach of exponential growth modeling using aquaculture dataset. Multiple linear regressions are very famous technique and are extensively used in many fields especially in agricultural research. The parametric bootstrap method is recommended for the small sample size for a reliable performance (Jung et.al, 2005; Cassel, 2010). Bootstrap method is a statistical technique that falls under the broad heading of resampling. This method is very useful and can be used various especially in the estimation of nearly any statistics (Cassel, 2010). This procedure involves a relatively simple procedure, but repeated so many times depending on the need of the researcher. Bootstrap technique is heavily dependent upon computer calculation. Using the bootstrap method we are able to determine the estimating value of a parameter that presenting the whole of a population. Without using bootstrap method, the value of the parameter of a population is impossible to measure directly. So, we use statistical sampling method and we sample a population, measure a statistic of this sample, and then use these statistics to say something about the corresponding parameter of the population (Cassel, 2010). For the case of nonlinear regression, we have to transform the equation from non-linear to a linear form. Multiple linear regressions $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k$ are an extension of simple linear regression. We used this technical to get a better result. The random error term is added to make the model probabilistic rather than deterministic. The value of the coefficient β_i determines the contribution of the independent variables x_i , and β_0 is the y-intercept (Diem Ngo & La Puente, 2012). A fuzzy regression model corresponding to $Y = Z_0 + Z_1 x_1 + Z_2 x_2 + \dots + Z_k x_k$ previously, explanation variables x_i 's are assumed to be precise. However, according to the equation above, response variable Y is not crisp but is instead fuzzy in nature. That means the parameters are also fuzzy in nature. Our objective is to estimate these parameters. In further discussion, Z_i 's are assumes symmetric fuzzy numbers which can be presented by interval. For example, Z_i can be express as fuzzy set given by $Z_i = \langle a_{ic}, a_{iw} \rangle$ where a_{ic} is centre and a_{iw} is radius or vagueness associated. Fuzzy set above reflects the confidence in the regression coefficients around a_{ic} in terms of symmetric triangular memberships function. Application of this method should be given more attention when the underlying phenomenon is fuzzy which

means that the response variable is fuzzy. So, the relationship is also considered to be fuzzy. This $Z_i = \langle a_{1c}, a_{1w} \rangle$ can be written as $Z_i = [a_{1L}, a_{1R}]$ with $a_{1L} = a_{1c} - a_{1w}$ and $a_{1R} = a_{1c} + a_{1w}$. In fuzzy regression methodology, parameters are estimated by minimizing total vagueness in the model. $y_j = Z_0 + Z_1x_{1j} + Z_2x_{2j} + \dots + Z_kx_{kj}$. Using $Z_i = \langle a_{1c}, a_{1w} \rangle$ we can write $y_j = \langle a_{0c}, a_{0w} \rangle + \langle a_{1c}, a_{1w} \rangle x_{1j} + \dots + \langle a_{nc}, a_{nw} \rangle x_{nj} = \langle a_{jc}, a_{jw} \rangle$. Thus this can be written as $y_{jc} = a_{0c} + a_{1c}x_{1j} + \dots + a_{nc}x_{nj}$ then it can be written straightly as $y_{jw} = a_{0w} + a_{1w}|x_{1j}| + \dots + a_{nw}|x_{nj}|$. As y_{jw} represent radius and so cannot be negative, therefore on the right-hand side of equation $y_{jw} = a_{0w} + a_{1w}|x_{1j}| + \dots + a_{nw}|x_{nj}|$, absolute values of x_{ij} are taken. Suppose there m data point, each comprising $a(n+1)$ -row vector. Then parameters Z_i are estimated by minimizing the quantity, which is total vagueness of the model-data set combination, subject to the constraint that each data point must fall within estimated value of response variable. This can be visualized as the following linear programming problem, minimized $\sum_{j=1}^m (a_{0w} + a_{1w}|x_{1j}| + \dots + a_{nw}|x_{nj}|)$ and Subject to $\left\{ \left(a_{0c} + \sum_{i=1}^n a_{ic}x_{ij} \right) + \left(a_{0w} + \sum_{i=1}^n a_{iw}x_{ij} \right) \right\} \geq Y_j$ and $\left\{ \left(a_{0c} + \sum_{i=1}^n a_{ic}x_{ij} \right) - \left(a_{0w} + \sum_{i=1}^n a_{iw}x_{ij} \right) \right\} \leq Y_j$ and $a_{iw} \geq 0$. Simple procedure is commonly used to solve the linear programming problem. (Kacprzyk and Fedrizzi, 1992). Data of this study is a sample which composed of two variables. Namely variables are as in Table 1.

Table1: Description of Data

Num.	Code	Explanation of user variables
1.	Y	Bacteria Growth Reading
2.	X	Dose

Table 2: Original Data

Dose(X)	Bacteria Growth Reading (Y)
2.000	1520.000
3.000	19110.00
4.000	16010.00
5.000	13610.00
8.000	5410.000
9.000	5010.000
12.00	2610.000
13.00	1510.000
15.00	910.0000

Table3: Taking LN for Data

Dose(X)	LN (Bacteria Growth Reading)(Y)
2.000	7.33
3.000	9.86
4.000	9.68
5.000	9.52
8.000	8.60
9.000	8.52
12.00	7.87
13.00	7.32
15.00	6.81

Exponential growth and decay regression was used in the analysis of relationship between variables. The algorithm is given as follows:

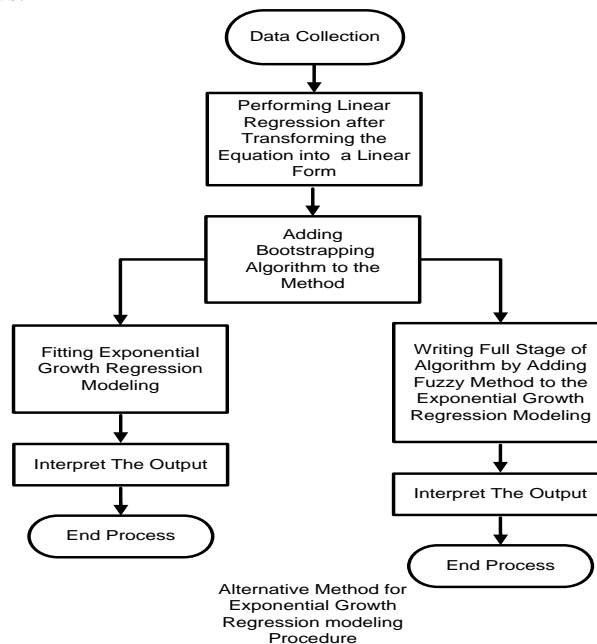


Figure 1: Flow Chart of an Alternative Exponential Modeling

Figure 1 showed the flow chart of an alternative method for exponential modeling procedure.

Step One: By transforming the $Y = Ae^{bX}$ into a linear form we obtained the following equation.

$Y = Ae^{bX}$ taking ln to the below, we obtained $\ln Y = \ln(Ae^{bX}) = \ln(A) + \ln(e^{bX}) = \ln(A) + bx$. Transforming data Y into lnY and let the X data in the original condition.

Step Two: /*Adding Bootstrapping Algorithm to the Method */
 title "Performing bootstrap with case resampling";
proc surveyselect data=bacteria out=boot1 method=urs samprate=1 outhits rep=2;
run;

Step Three: /*Running the Original Data using Bootstrap Method*/

```
data bacteria;
input x y;
datalines;
2.00 7.33
3.00 9.86
4.00 9.68
5.00 9.52
8.00 8.60
9.00 8.52
12.00 7.87
15.00 6.81
;
run;
title "Performing bootstrap with case resampling";
proc surveyselect data=bacteria out=boot1 method=urs samprate=1 outhits rep=2;
run;
proc print data= boot1;
run;
```

Table 4: Results: Data enlargement after performing bootstrapping method

Dose	LN (Bacteria Growth Reading) (Y)
2	7.33
2	7.33
3	9.86
4	9.68
4	9.68
13	7.32
13	7.32
15	6.81
2	7.33
2	7.33
4	9.68
4	9.68
9	8.52
9	8.52
9	8.52
13	7.32

Step Four: /*Running the Bootstrap Data Using Regression Method*/

```
ods rtf file='robdunc0.rtf' style=journal;
Title "Simple Linear Regression Using Proc Reg";
Proc Reg Data=boot1 ;
Model y = x;
Run;
ods rtf close;
```

Results: Results after performing bootstrapping method

Table 5: Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	8.90477	0.47669	18.68	<.0001
x	1	-0.09487	0.05846	-1.62	0.1269

$$\text{LN (Bacteria Reading)} = 8.90477 - 0.09487 \quad (i)$$

Standard Errors (0.47669) (0.05846)

Now we calculated width of predicted least square method,
 Referring to the (i) we can calculated width as follows:

$$\begin{aligned} \text{Width} &= (7367.030 + 0.4767) e^{(-0.09487+0.0584)X} - (7367.030 + 0.4767) e^{(-0.09487+0.0584)X} \\ &= 7367.50669 e^{-0.03647X} - (7366.55331 e^{-0.1533X}) \end{aligned}$$

Step Five: /*Performing Fuzzy Least Squares (FLS) to Exponential Growth and Decay Regression after Transforming the Equation into a Linear Form and Bootstrapping Method */

```

data bacteria;
input x y;
datalines;
2      7.33
2      7.33
3      9.86
4      9.68
4      9.68
13     7.32
13     7.32
15     6.81
2      7.33
2      7.33
4      9.68
4      9.68
9      8.52
9      8.52
9      8.52
13     7.32
;
ods rtf file='robdunc0.rtf' style=journal;
/* Method of fuzzy least squares (FLS)to the above data */
proc nlp;
min Y;
decvar ar br ac bc;
bounds ar>=0, br>=0, ac= 8.90477, bc= -0.09487;
lincon ac+2*bc-ar-2*br<=7.33;
lincon ac+2*bc-ar-2*br<=7.33;
lincon ac+3*bc-ar-3*br<=9.86;
lincon ac+4*bc-ar-4*br<=9.68;
lincon ac+4*bc-ar-4*br<=9.68;
lincon ac+13*bc-ar-13*br<=7.32;
lincon ac+13*bc-ar-13*br<=7.32;
lincon ac+15*bc-ar-15*br<=6.81;
lincon ac+2*bc-ar-2*br<=7.33;
lincon ac+2*bc-ar-2*br<=7.33;
lincon ac+4*bc-ar-4*br<=9.68;
lincon ac+4*bc-ar-4*br<=9.68;
lincon ac+9*bc-ar-9*br<=8.52;
lincon ac+9*bc-ar-9*br<=8.52;
lincon ac+9*bc-ar-9*br<=8.52;
lincon ac+13*bc-ar-13*br<=7.32;
lincon ac+2*bc+ar+2*br>=7.33;
lincon ac+2*bc+ar+2*br>=7.33;
lincon ac+3*bc+ar+3*br>=9.86;
lincon ac+4*bc+ar+4*br>=9.68;
lincon ac+4*bc+ar+4*br>=9.68;
lincon ac+13*bc+ar+13*br>=7.32;
lincon ac+13*bc+ar+13*br>=7.32;
lincon ac+15*bc+ar+15*br>=6.81;
lincon ac+2*bc+ar+2*br>=7.33;
    
```

```
lincon ac+2*bc+ar+2*br>=7.33;
lincon ac+4*bc+ar+4*br>=9.68;
lincon ac+4*bc+ar+4*br>=9.68;
lincon ac+9*bc+ar+9*br>=8.52;
lincon ac+9*bc+ar+9*br>=8.52;
lincon ac+9*bc+ar+9*br>=8.52;
lincon ac+13*bc+ar+13*br>=7.32;
Y=16*ar+108*br;
run;
ods rtf close;
```

Table 6: Fitting Exponential Growth and Decay Regression value using Fuzzy Least Square

Optimization Results				
Parameter Estimates				
N	Parameter	Estimate	Gradient Objective Function	Active Bound Constraint
1	ar	1.385030	16.000000	
2	br	1.97325E-17	108.000000	Lower BC
3	ac	8.904770	0	Equal BC
4	bc	-0.094870	0	Equal BC

Substituting the values of parameter estimates in model (Table 6) we obtained

$$\text{LN (Bacteria Reading)} = 8.90477 - 0.09487X \quad (\text{ii})$$

Standard Errors (1.385) (0.000)

Referring to the (ii) now we calculated width of predicted fuzzy least square method,

$$\text{Width} = (7367.030 + 1.385030) \times e^{(-0.09487+0.00)X} - (7367.030 - 1.385030) \times e^{(-0.09487-0.00)X}$$

$$= 7368.41503e^{-0.09487X} - (7365.64497e^{-0.09487X})$$

Table 7: Width of Predicted Interval by Least Square and Fuzzy Least Square

Dose	LN (Bacteria Growth Reading) (Y)	Width of Predicted Interval	
		Least Square in (LN)	Fuzzy Least Square (LN)
2	7.33	7.26	0.83
2	7.33	7.26	0.83
3	9.86	7.58	0.73
4	9.68	7.77	0.64
4	9.68	7.77	0.64
13	7.32	8.18	-0.21
13	7.32	8.18	-0.21
15	6.81	8.17	-0.40
2	7.33	7.26	0.83
2	7.33	7.26	0.83
4	9.68	7.77	0.64
4	9.68	7.77	0.64
9	8.52	8.15	0.17
9	8.52	8.15	0.17
9	8.52	8.15	0.17
13	7.32	8.18	-0.21
Average Width		7.81	0.38

Conclusion:

We computed the predicted interval using method least square and method of fuzzy least square. Fuzzy least square show the average much shorter compared to the method of least square. This indicated that fuzzy least square is more efficient than method least square.

Summary and Discussion:

This paper gives the explanation for an alternative programming method of bootstrap approach to nonlinear regression procedure using SAS software. The aim for the algorithm building is to provide the researcher with the alternative programming of a data analysis. This method can be applied for the small sample size data especially where the data is very difficult to collect. In our case, smaller width of predicted interval c will tell us how accurate our estimate parameter is likely to be.

References:

1. Diem Ngo, T.H., La Puente, C.A. (2012). The Steps to Follow in a Multiple Regression Analysis. SAS Global Forum 2012: Statistics and Data Analysis. Paper 333-2012, Pp 1-12.

2. Cassel, D.L., 2010. Bootstrap Mania: Re sampling the SAS. SAS Global Forum 2010: Statistics and Data Analysis. Paper 268-2010: pp: 1-11.
3. Jung, B.C., Jhun, M., Lee, J.W., 2005. Bootstrap Tests for Over dispersion in a Zero-Inflated Poisson Regression Model. Biometrics 61, 626-629.
4. Kacprzyk J. and Fedrizzi M. (1992) Fuzzy Regression Analysis, Omnitech Press, Warsaw.