

## Modeling for exponential growth and decay methodology in biometry using SAS syntax

Wan Muhammad Amir bin Wan Ahmad<sup>1\*</sup>, Farah Muna Mohamad Ghazali<sup>1</sup>, Nor Farid Mohd Noor<sup>1</sup>, Nor Azlida Aleng<sup>2</sup>

<sup>1</sup>School of Dental Sciences, Health Campus, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia

<sup>2</sup>Faculty of Ocean Engineering Technology and Informatics, Universiti Malaysia Terengganu, Kuala Terengganu, Terengganu, Malaysia

\*Corresponding E-mail: [wmamir@usm.my](mailto:wmamir@usm.my)

### ARTICLE INFO

### ABSTRACT

#### Article History

Received 16 March 2021

Revised 20 March 2021

Accepted 20 March 2021

#### Keywords

Bootstrap

Fuzzy

Weighted

Robust regression

#### How to cite this article:

Wan Ahmad, W. M. A., Ghazali, F. M. M., Mohd Noor, N. F., & Aleng, N. A. (2021). Modeling for exponential growth and decay methodology in biometry using SAS syntax. *Bulletin of Applied Mathematics and Mathematics Education*, 1(1), 1-8.

This paper provided an alternative method for exponential growth modeling as a regression analysis technique through the SAS algorithm. This alternative method is a combination technique (using nonlinear model bootstrap and fuzzy regression) for the small data set and gives the researcher an option to start the analysis, even if there is not enough data set. This method enhances the previous methodology with embedded bootstrapping and fuzzy technique to a nonlinear regression model. This principle aims to propose an alternative method of analysis with better results. In our case, we applied this principle to farm data and compared the results obtained by looking at the average width of the predicted interval.

This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



### Introduction to Algorithm Using SAS Language

This paper provides a road map of the practical approach using the biometry dataset to exponential growth modeling. Multiple linear regressions are a 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 reliable performance (Jung et al., 2005; Cassel, 2010). The bootstrap method is a statistical technique that falls under the broad heading of resampling. This method is very useful and can be used variously especially in the estimation of nearly any statistics (Cassel, 2010). This procedure involves a relatively simple procedure but is repeated so many times depending on the need of the researcher. The bootstrap technique is heavily dependent upon computer calculation. Using the bootstrap method, we can determine the estimating value of a parameter that presenting the whole of a population. Without using the bootstrap method, the value of the parameter of a population is impossible to measure directly. So, we use the 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 a non-linear to a linear form. Multiple linear regressions  $Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_kx_k$  are an extension of simple linear regression. We used this technique to get a better result. The random error term is added to make the model probabilistic rather than deterministic. The value of the coefficient  $\beta_i$  determines the contribution of the independent variables  $x_i$  and  $\beta_0$  is the  $y$ -intercept (Diem Ngo & La Puente, 2012). A fuzzy regression model corresponding to  $Y = Z_0 + Z_1x_1 + Z_2x_2 + \dots + Z_kx_k$  previously, explanatory variables  $x_i$ 's is assumed to be precise. However, according to the equation above, response variable  $Y$  is not crisp but is instead fuzzy. That means the parameters are also fuzzy. Our objective is to estimate these parameters. In further discussion,  $Z_i$ 's are assumed symmetric fuzzy numbers which can be presented by the interval. For example,  $Z_i$  can be express as a fuzzy set given by  $Z_i = \langle a_{ic}, a_{iw} \rangle$  where  $a_{ic}$  is the center and  $a_{iw}$  is radius or vagueness associated. The 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  $Z_i = \langle a_{ic}, a_{iw} \rangle$  can be written as  $Z_i = [a_{iL}, a_{iR}]$  with  $a_{iL} = a_{ic} - a_{iw}$  and  $a_{iR} = a_{ic} + a_{iw}$ . 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  $Z_i = \langle a_{ic}, a_{iw} \rangle$  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 the equation  $y_{jw} = a_{0w} + a_{1w}|x_{1j}| + \dots + a_{nw}|x_{nj}|$ , absolute values of  $x_{ij}$  being taken. Suppose there are  $m$  data point, each comprising  $a(n+1)$ -row vector. Then parameters  $Z_i$  are estimated by minimizing the quantity, which is the total vagueness of the model-data set combination, subject to the constraint that each data point must fall within the estimated value of the 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  $a_{iw} \geq 0$ . The simple procedure is commonly used to solve the linear programming problem (Kacprzyk & Fedrizzi, 1992). Data of this study is a sample which composed of two variables (See Table 1), while the data are in Table 2 and Table 3.

**Table 1.** Description of data

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

**Table 2.** Original data

Dose(X)	Bacteria 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

**Table 3.** Taking LN for data

Dose(X)	LN (Bacteria 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 the relationship between variables. The algorithm is given as follows.

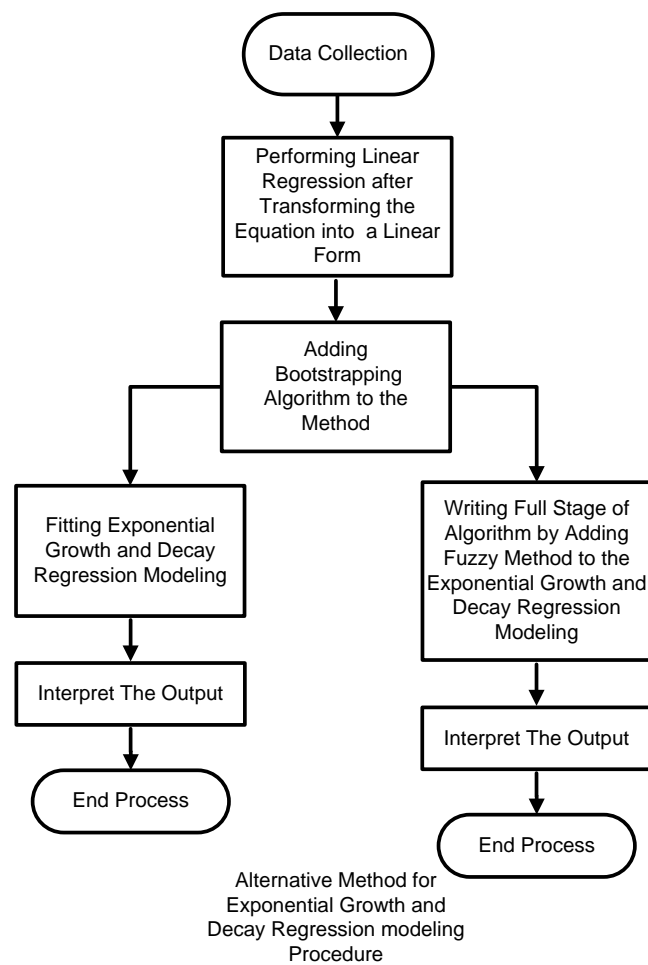
**Figure 1.** Flow chart of an alternative exponential modeling

Figure 1 showed the flow chart of an alternative method for the 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;
```

The data enlargement after performing the bootstrapping method is presented in Table 4.

**Table 4.** Data enlargement after performing the bootstrapping method.

<i>Dose</i>	<i>LN (Bacteria Reading) (Y)</i>
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='robduc0.rtf' style=journal;
Title "Simple Linear Regression Using Proc Reg";
Proc Reg Data=boot1 ;
Model y = x;
Run;
ods rtf close;
```

Results after performing the bootstrapping method is presented in Table 5.

**Table 5.** Parameter estimates

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

$$\text{LN (Bacteria Reading)} = 8.90477 - 0.09487x \quad (1)$$

Standard Errors      (0.47669)    (0.05846)

Now we calculated the width of the predicted least square method,

Referring to the (1) we can calculate width as follows:

$$\begin{aligned} \text{Width} &= (7367.030 + 0.4767)e^{(-0.09487 + 0.0584)x} - (7367.030 + 0.4767)e^{(-0.09487 + 0.0584)x} \\ &= 7367.50669 e^{-0.03641x} - (7366.55331 e^{-0.1533x}) \end{aligned}$$

**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='robduc0.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 shows the fitting exponential growth and decay regression value using fuzzy least square parameter estimates.

**Table 6.** Fitting exponential growth and decay regression value using fuzzy least square 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 the model (Table 6) we obtained

$$\begin{aligned}
 \text{LN (Bacteria Reading)} &= 8.90477 - 0.09487 & (2) \\
 \text{Standard Errors} & \quad (1.385) \quad (0.000)
 \end{aligned}$$

Referring to the (2) now we calculated the width of the predicted fuzzy least square method,

$$\begin{aligned} \text{Width} &= (7367.030 + 1.385030) \times e^{(-0.09487 + 0.00)x} - (7367.030 - 1.385030) \times e^{(-0.09487 - 0.00)x} \\ &= 7368.41503e^{-0.09487 x} - (7365.64497e^{-0.09487 x}) \end{aligned}$$

Table 7 presents the width of predicted interval by least square and fuzzy least square.

**Table 7.** Width of predicted interval by least square and fuzzy least square

Dose	LN (Bacteria 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.

## References

- 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, 1-12.
- Cassel, D. L. (2010). Bootstrap Mania: Resampling the SAS. *SAS Global Forum 2010: Statistics and Data Analysis*. Paper 268-2010, 1-11.
- Jung, B. C., Jhun, M., & Lee, J. W. (2005). Bootstrap Tests for Overdispersion in a Zero-Inflated Poisson Regression Model. *Biometrics*, 61, 626-629.
- Kacprzyk J., & Fedrizzi, M. (1992). *Fuzzy Regression Analysis*. Warsaw: Omnitech Press.

This page is intentionally left blank.