Communications in Mathematics and Applications
Vol. 13, No. 3, pp. 1027–1035, 2022
ISSN 0975-8607 (online); 0976-5905 (print)
Published by RGN Publications
DOI: 10.26713/cma.v13i3.1781

http://www.rgnpublications.com

Research Article

Assessing the Efficiency of Bootstrap Method: Application SAS Programming Algorithm for Exponential Growth Streptococcus Sobrinus

Rabiatul Adawiyah Abdul Rohim[1], Kamariah Abdul Jalil*[1],
Puspa Liza Ghazali[2], Wan Muhamad Amir W. Ahmad[3],
Noor Huda Ismail[3] and Siti Nadhirah Mohamad Fauzi[2]

[1] *Faculty of Medicine, Universiti Sultan Zainal Abidin, Campus Kota, 20400, Kuala Terengganu, Terengganu, Malaysia*
[2] *Faculty of Business and Management, University Sultan Zainal Abidin, 21300 Kuala Nerus, Terengganu, Malaysia*
[3] *School of Dental Sciences, Health Campus, Universiti Sains Malaysia, 16150, Kubang Kerian, Kelantan, Malaysia*
***Corresponding author:** kamariahjalil@unisza.edu.my

Received: January 5, 2022 **Accepted:** May 1, 2022

Abstract. *Introduction*: The bootstrap method is commonly used to estimate values linked with estimator and test statistical sampling distributions. As a result, it is required to have a good manner of calculating with some improvement of the proposed strategy to maintain the proposed model's efficiency and accuracy. *Objective*: This study aimed to compare parameter estimates using SAS programming and determine the optimal computation for representing data using the bootstrapping and non-bootstrapping methods. *Methods*: In this study, the bootstrap method was applied using SAS programming. *Results*: In comparison to a model without bootstrapping, exponential growth demonstrates a very good model utilizing bootstrapping approach because the outcome provided the smallest standard error. This implies that the best model is the exponential model with bootstrapping approach. *Conclusions*: Finally, these prospective methodologies have resulted in the effective study and provide the best outcomes for decision making, particularly for decision-makers, among studies in their sector.

Keywords. Exponential growth, Exponential model, Bootstrap method, SAS programming, Parameter estimates

Mathematics Subject Classification (2020). 00A05, 00A06, 00A71

1. Introduction

Probiotic bacteria have been studied extensively and are characterised as live microorganisms that have a beneficial effect on the host. For years, rebalancing the compound microflora has been used to treat gastrointestinal disorders. The oral cavity, like the gastrointestinal tract, is heavily colonised by bacteria, and many different bacterial species make up the microbiota in the mouth, which provides ideal conditions for bacteria with a constant temperature, moist surface with a relatively constant PH, and consistent nutrient supply. Streptococcus Sobrinus, a probiotic bacteria, is a viable therapy option for oral diseases with a microbiological cause [11].

There are billions of microorganisms in the human body, and the mouth cavity is one of the greatest sources of microbes. The human mouth is home to between 700 and 1000 microbial species. Oral disease, such as dental caries, periodontal disease, and oral cancer, are directly linked to oral bacteria' prevalence and progression [7, 9, 10, 12]. Meanwhile, oral bacteria can enter the bloodstream via a damaged oral mucosa, causing a rise in systemic antibody levels, raising the risk of a number of cardivascular diseases [7]. According to Peres *et al*. [10], the social importance of preventing and treating oral disease is projected to be \$356.80 billion in direct expenditures and \$ 187.61 billion in indirect costs in 2015. Dental illness is unquestionably a public health issue and one of the most common diseases worldwide, especially dental caries, which is a biofilm-related disease [14].

According to (Rohim *et al*. 2020 [11]), the integrated exponential regression method is more efficient than the exponential regression technique. The bootstrap method can be used to analyse data from small samples, especially when data is scarce. Many researchers that examine and stress specific approaches have attempted to construct a new or superior methodology in order to achieve better outcomes [11]. This study lays out a step-by-step guide to exponential growth modelling using the oral bacteria dataset. For dependable performance with a limited sample size, the parametric bootstrap approach is advised [3, 6]. The bootstrap method is a statistical approach that belongs to the resampling category. This method is quite useful and may be applied to a variety of situations, including the estimation of practically any statistic [1]. This technique is pretty easy, although it is repeated numerous times depending on the researcher's needs. The bootstrap method relies significantly on computer calculations. We can use the bootstrap approach to get the estimating value of a parameter that represents the entire population. The value of a population's parameter cannot be measured directly without the bootstrap method. So we utilise the statistical sampling approach to sample a population, measure the sample's statistics, and then use these statistics to infer something about the population's corresponding parameter [1].

2. Literature Review

Bootstrap

Proposed by Efron (1979) [4], the bootstrap is a simulation technique to evaluate the standard error of the estimation of a parameter. The idea is to perform various resampling with replacement of the dataset. Let $\widehat{\theta} = t(x)$ the estimator of θ calculated from a sample $(x_1, \dots, x_n)$.

A sample bootstrap $x^*_{(1)} = (x^*_1, \dots, x^*_n)$ is a resample with the replacement of size n from $(x_1, \dots, x_n)$, the index $i = 1, \dots, B$ refers to the number of wanted replicas B. This way the bootstrap estimator of the variance of $\widehat{\theta}$ is given by

$$\operatorname{var}(\widehat{\theta}) = \frac{1}{B-1} \sum_{i=1}^{B} \left(t(x^*_{(i)}) - \widehat{\theta}^*\right)^2,$$

where $t(x^*_{(i)})$ is the estimator of based on the ith replica bootstrap and $\widehat{\theta}^* = B^{-1} \sum\limits_{i=1}^{B} t(x^*_{(i)})$. For a sufficiently large number of replicas, we can calculate a confidence interval (with significance level 2α) for $\widehat{\theta}$ through normal approximation, which is given by

$$IC_B[\widehat{\theta}, (1-2\alpha)] = \left[\widehat{\theta}^* - Z_\alpha[\operatorname{var}(\widehat{\theta})]\right]^{\frac{1}{2}} ; \widehat{\theta}^* + z_\alpha\left[\operatorname{var}(\widehat{\theta})\right]^{\frac{1}{2}}$$

where z_α corresponds to the α-the quantile of standard normal distribution. We can estimate the empiric density $\widehat{F}$ as a normal approximation of real F. As such, we can estimate the confidence interval using quantiles, so the interval is given by $IC_B[\widehat{\theta}, (1-2\alpha)] = [\widehat{F}(\alpha); \widehat{F}(1-\alpha)]$.

Bacteria

Streptococcus sobrinus are known to be associated with dental caries in humans. Mutans Streptococci (Streptococcus Sobrinus) are strongly associated with the development of dental caries in human. These bacteria are the most common pathogens isolated from human dental plaque, and their prevalence has been reported in epidemiological studies [8]. Various methods have been used for detecting and identifying mutans Streptococci, including culturing, direct enzyme tests, enzyme-linked immunosorbent assays, and DNA probes [5]. In previous cross-sectional and longitudinal studies, we reported that preschool children with primary dentition harboring both S. Mutans and S. Sobrinus had a significantly higher incidence of dental caries than those with S. Mutans alone. The mutans group of oral Streptococci consist of seven species: Streptococcus Cricetus, S. Rattus, S. Mutans, S. Sobrinus, S. Downei, S. Macacae and S. Ferus [13]. Among the group, S. Mutans and S. Sobrinus are most frequently isolated from human dental plaque and closely associated with human dental caries [11]. Prevalence of S. Sobrinus and S. Mutans in the human oral cavity have been reported by epidemiological studies in which the isolation frequency of S. Mutans from dental plaque is much higher than that of S. Sobrinus [8].

3. Methodology

Data which characterize the proliferative capacity of microorganisms are used, composed of two variables which are the days of the culture (x), bacterial growth (y), and the rate of exponential growth of a bacterial culture ($\ln y$). The data was gathered at the Microbiology Laboratory using several mediums in order to find the best concentration rate. The bacteria's growth rate was determined by monitoring the alga plate at the Microbiology Laboratory, PPSG, USM. The rate of reproduction of bacteria was studied at a concentration of (5×105 CFU/Ml). The bacteria samples were taken on several days (within 7 days). Then, the data were recorded in SAS Programme to evaluate microorganism growth before analyzing by using the new proposed statistical methodology. The chosen microorganisms were stored in the appropriate manner (Table 1 and Figure 1). After that, all bacteria were carefully disposed of in the clinical waste bin.

Table 1. Sample of microorganisms

Days of the culture (x)	1	2	3	5	9	11	13	15	20	23	25	29	33
Stereptococcus Sobrinus (y) (1×10^5)	82	84	86	86	87	87	87	4.50	4.52	4.54	4.53	4.55	4.60
The rate of exponential growth ($\ln y$)	4.41	4.43	4.45	4.45	4.47	4.47	4.47	4.50	4.52	4.54	4.53	4.55	4.60

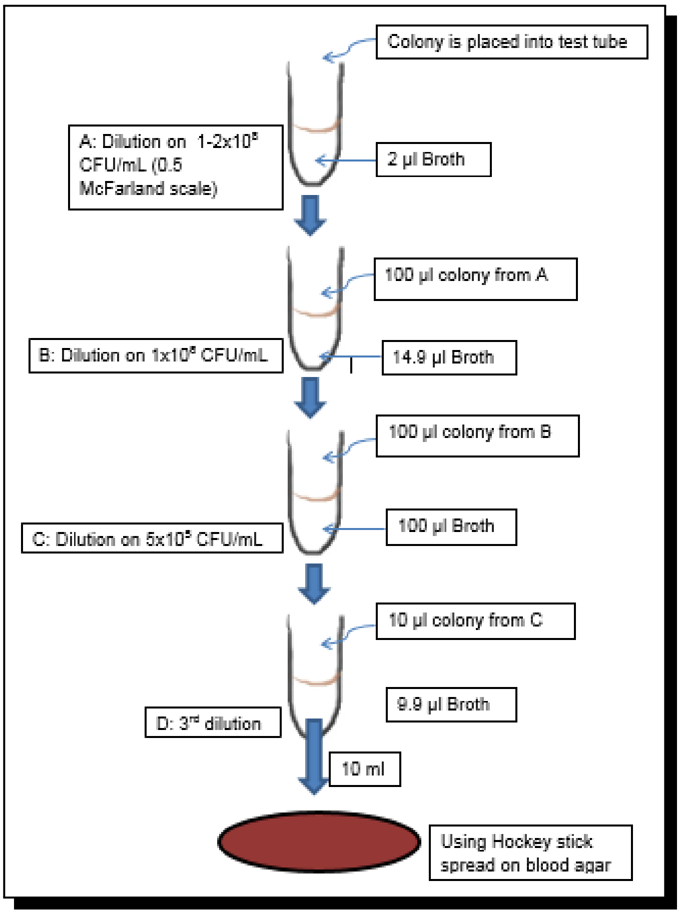

Figure 1. Step for broth microdilution method

The Algorithm of Exponential Calculation

The algorithm showed below is a way of inserting data in the SAS algorithm and the way of calculating the bootstrapping and without bootstrapping method.

- Data in SAS format. The name of the dataset is given as `cell_streptococcus`. The data consist of two variables which names as `x`, `y` and `ln y`

```
data cell_streptococcus;
input x y ln y;
datalines;
1.00  82.00 4.41
2.00  84.00 4.43
3.00  86.00 4.45
5.00  86.00 4.45
9.00  87.00 4.47
11.00 87.00 4.47
13.00 87.00 4.47
15.00 90.00 4.50
20.00 92.00 4.52
23.00 94.00 4.54
25.00 93.00 4.53
29.00 95.00 4.55
33.00 99.00 4.60
;
run;
```

- Adding a bootstrapping algorithm to the methodology building. `cell_streptococcus` data were bootstrap two times with the case resampling. The following procedure in detail was given in SAS syntax as follows. The new data which generate by the SAS procedure will be named as booted. The produce data in the study will be print through the proc print procedure.
- We also add the syntax of `''ods rtf file='abc.rtf' style=journal''` in the SAS language in order to get the output in Microsoft Word format.

```
\%MACRO bootstrap(data=cell_streptococcus,booted=booted,boots=2,seed=1234);
DATA \&booted;
pickobs = INT(RANUNI(&seed)*n)+1;
SET &data POINT = pickobs NOBS = n;
REPLICATE=int(i/n)+1;
i+1;
IF i > n*&boots THEN STOP;
RUN;
%MEND bootstrap;
ods rtf file='abc.rtf' style=journal;
%bootstrap(data=cell_streptococcus,boots=2);
run;
proc print data=booted;
run;
```

```
Title "Exponential Equation without bootstrap";
ods graphic/imagename="Exponential Equation without bootstrap";
proc nlin data=cell_streptococcus plots=fit;
parameters A=1 b=0;
model y=A*exp(b*x);
ods output EstSummary=summExp;
run;
Title "Exponential Equation with bootstrap";
ods graphic/imagename="Exponential Equation with bootstrap";
proc nlin data=booted plots=fit;
parameters A=1 b=0;
model y=A*exp(b*x);
ods output EstSummary=summExp;
run;
ods rtf
close;
```

4. Results and Discussions

Model with Bootstrapping Method

The result of the first model involving bootstrapping method is given in Table 2 and Table 3. Table 2 shows that the model predicts the dependent variable significantly well. The p-value is less than 0.05, and this indicates that, overall, the model statistically significantly predicts the outcome variable, this is a good fit for the data.

Table 2. ANOVA

Source	DF	Sum of squares	Mean square	F value	Approx Pr > F
Model	2	197028	98513.8	153369	<.0001
Error	24	15.4160	0.6423		
Total	26	197043			

Table 3. Parameter estimates

Parameter	Estimate	Approx Std. Error	Approximate 95% confidence limits	
A	82.9121	0.2208	82.4563	83.3679
b	0.00512	0.000197	0.00471	0.00553

From Table 3, we can write an exponential model as

$$y = 82.9121e^{0.00512x} \tag{4.1}$$

Model (4.1) can be transformed into a linear form by taking, it can be written as follows

$$\ln(y) = \ln(82.9121) + 0.00512x,$$

$$\ln(y) = 4.4178 + 0.00512x.$$

Given in Table 3 is the information on parameter estimates. From Table 3, the standard error indicates how much standard error value occurs in the model. In this case, 0.2208 and 0.000197 can explain, which is very small. The good model involving bootstrapping method is the one that has minimum standard error value.

Model without Bootstrapping Method

The result for the second model not involving bootstrapping method is given in Table 4.

Table 4. Parameter estimates

Parameter	Estimate	Approx Std. Error	Approximate 95% confidence limits	
A	83.0253	0.5096	81.9036	84.1469
b	0.00499	0.000327	0.00427	0.00571

From Table 4, we can write an exponential model as

$$y = 83.0253e^{0.00499x}. \tag{4.2}$$

Model (4.2) can be transformed into a linear form by taking, it can be written as follows

$$\ln(y) = \ln(83.0253) + 0.00499x,$$

$$\ln(y) = 4.4191 + 0.00499x.$$

Given in Table 4 is the information on parameter estimates. From Table 4, the standard error indicates how much standard error value occurs in the model. In this case, 0.5096 and 0.000327 can explain, which is larger than standard error involving the bootstrapping method. The good model involving bootstrapping method is the one that has minimum standard error value than model not involving the bootstrapping method.

Table 5. Comparison of parameter estimate

Parameter estimates	Bootstrapping method	Without bootstrapping method
Model	$y = 82.9121e^{0.00512x}$ or in a linear form $\ln(y) = 4.4178 + 0.00512x$	$\ln(y) = 4.4191 + 0.00499x$

Discussion

The objective of this research was to compare the parameter estimate using SAS programming and find the best calculation which can represent the data through the bootstrapping and without bootstrapping method. Two different methods were used: a model with bootstrapping method and a model without bootstrapping method. Exponential growth shows a very good model involving bootstrapping method because the result produced the smallest standard error reading compare with model without bootstrapping method. This indicates that the model of an exponential with bootstrapping method is the best model. The other without bootstrapping methods fit the model poor. Bootstrapping is a very versatile model for calculating a distribution of possible result for the unpaid claims, is comparatively easy to utilize and explain to others and can be readily "generalized" to be more flexible and blend with other related examples [2]. Therefore, to keep an efficiency and accuracy of the proposed model, it is necessary to have a good way of calculation with some improvement of the proposed strategy.

The model of an exponential with bootstrapping method is the best model. It is because advantage of a bootstrap model is that it can be specifically tailored to the statistical features found in the data under analysis [11]. This is particularly important as the results of any simulation model are only as good as the model used in the simulation process. Like all models and method, the quality of the bootstrap model depends on the quality of the assumptions.

5. Conclusion

In conclusion, the exponential model using bootstrapping approach is the best model.When only a small sample size is available, the bootstrap approach might be used. To keep the efficiency and accuracy of the proposed model, it is necessary to have a good way of calculation with some improvements of the proposed strategy.

Acknowledgments

The authors would like to thank all the participants who agreed to participate in this study and thanks to the anonymous referees for their valuable suggestions which led to the improvement of the manuscript.

Competing Interests

The authors declare that they have no competing interests.

Authors' Contributions

All the authors contributed significantly in writing this article. The authors read and approved the final manuscript.

References

[1] W. M. A. W. Ahmad, F. M. M. Ghazali, N. F. M. Noor and N. A. Aleng, Modeling for exponential growth and decay methodology in biometry using SAS syntax, *Bulletin of Applied Mathematics and Mathematics Education* **1**(1) (2021), 1 – 8, DOI: 10.12928/bamme.v1i1.3853.

[2] W. M. A. Ahmad, R. A. A. Rohim and N. H. Ismail, Forecasting parameter estimates: A modeling approach using exponential and linear regression, *Journal of Engineering, Technology & Applied Science Research* **8**(4) (2018), 3162 – 3167, DOI: 10.48084/etasr.2150.

[3] D. L. Cassel, *BootstrapMania: Resampling the SAS®*, SAS Global Forum 2010: Statistics and Data Analysis, Paper 268-2010, 1 – 11, URL: https://support.sas.com/resources/papers/proceedings10/268-2010.pdf.

[4] B. Efron, Bootstrap methods: Another look at Jacknife, *The Annals of Statistics* **7** (1979), 1 – 26, DOI: 10.1214/aos/1176344552.

[5] T. Igarashi, A. Yamamoto and N. Goto, PCR for detection and identification of Streptococcus sobrinus, *Journal of Medical Microbiology* **49** (2000), 1069 – 1074, DOI: 10.1099/0022-1317-49-12-1069.

[6] B. C. Jung, M. Jhun and J. W. Lee, Bootstrap tests for overdispersion in a zero-inflated poisson regression model, *Biometrics* **61** (2005), 626 – 629, DOI: 10.1111/j.1541-0420.2005.00368.x.

[7] M. Károly, N. Gábor, N. Ádám and B. Andrea, Characteristics, diagnosis and treatment of the most common bacterial diseases of the oral cavity, *Orvosi Hetilap* **160**(19) (2019), 739 – 746, DOI: 10.1556/650.2019.31377, (in Hungarian).

[8] W. J. Loesche, Role of streptococcus mutans in human dental decay, *Microbiological Reviews* **50** (1986), 353 – 380, DOI: 10.1128%2Fmr.50.4.353-380.1986.

[9] S. A. Mosaddad, E. Tahmasebi, A. Yazdanian, M. B. Rezvani, A. Seifalian, M. Yazdanian and H. Tebyanian, Oral microbial biofilms: An update, *European Journal of Clinical Microbiology & Infectious Diseases* **38** (2019), 2005 – 2019, DOI: 10.1007/s10096-019-03641-9.

[10] M. A. Peres, L. M. D. Macpherson, R. J. Weyant, B. Daly, R. Venturelli, M. R. Mathur, S. Listl, R. K. Celeste, C. C. Guarnizo-Herreño, C. Kearns, H. Benzian, P. Allison and R. G. Watt, Oral diseases: A global public health challenge, *The Lancet* **394**(10194) (2019), 249 – 260, DOI: 10.1016/S0140-6736(19)31146-8.

[11] R. A. A. Rohim, W. M. A. Ahmad, N. H. Ismail, M. A. Yaaqob, M. K. Alam and F. M. M. Ghazali, Study of oral lactobacillus towards developing a comprehensive structured for integrated exponential regression model, *Bangladesh Journal of Medical Science* **19**(3) (2020), 552 – 557, DOI: 10.3329/bjms.v19i3.45874.

[12] J. Štšepetova, J. Truu, R. Runnel, R. Nõmmela, M. Saag, J. Olak, H. Nõlvak, J.-K. Preem, K. Oopkaup, K. Krjutškov, E. Honkala, S. Honkala, K. Mäkinen, P.-L. Mäkinen, T. Vahlberg, J. Vermeiren, D. Bosscher, P. de Cock and R. Mändar, Impact of polyols on oral microbiome of Estonian schoolchildren, *BMC Oral Health* **19** (2019), Article number: 60, DOI: 10.1186/s12903-019-0747-z.

[13] R. A. Whiley and D. Beighton, Current classification of the oral streptococci, *Oral Microbiology and Immunology* **13** (1998), 195 – 216, DOI: 10.1111/j.1399-302x.1998.tb00698.x.

[14] K. Yadav and S. Prakash, Dental caries: A microbiological approach, *Journal of Clinical Infectious Diseases & Practice* **2**(1) (2017), 1 – 15, DOI: 10.4172/2476-213X.1000118.