Confidence Bands for Survivor Function of Two Parameters Exponential Distribution under Double Type-II Censoring with Bootstrap Percentile

Akhmad Fauzy

Department of Statistics, Universitas Islam Indonesia, Yogyakarta, 55584
e-mail: afauzy@fmipa.uii.ac.id

Abstract

This paper describes existing methods and develops new methods for constructing confidence bands for survivor function of two parameters exponential distribution under double type-II censoring. Our results are built on extensions of previous work by [11] and Balakrishnan [1]. They use maximum likelihood estimator to construct interval estimation under double type-II censoring. The confidence bands are developed for survivor function using the confidence region about survivor function. We will use another method, known as the bootstrap percentile from [4]. This method gives shorter confidence bands compared to the traditional method.
Key words: air pollutant, bootstrap percentile, confidence bands, double type-II, survivor function

1. Introduction

The survivor function or reliability function is a property of any random variable that maps a set of events, usually associated with mortality or failure of some system, onto time. It captures the probability that the system will survive beyond a specified time. The term reliability function is common in engineering while the term survivor function is used in a broader range of applications, including human mortality.

The exponential distribution is often proposed as a model in life testing and reliability because of its simplicity and the ease with which estimates can be calculated. [2] deals with inference procedures for the one-parameter exponential model. Inference for the two-parameter exponential model has been studied by [9], [10], [12] and many others, based on complete samples and type-II censored data.

In reliability studies, due to time limitations and/or other restrictions on data collection, several lifetimes of units put on test may not be observed. In addition, sometimes the lowest and/or highest few observations in a sample could be due to some negligence or some other extraordinary reasons. In such cases, it is convenient to remove those outlying observations. Type-II censored samples are considered here, whereby, in an ordered sample of size n, a known number of observations is missing at either end (single censoring) or both ends (double censoring). Doubly censored samples have been considered, by authors, like [1] and [11]. They used maximum likelihood estimator to construct interval estimation for survivor function of two parameters exponential distribution under double type-II censoring. Using the intervals estimation for survivor functions at every lifetime develops confidence bands for survivor function. This band allows us to see the region in which the survivor function lies.

Bootstrap method is a computer-based method for assigning measures of accuracy to statistical estimates, especially to calculate the confidence interval. The aim of using bootstrap method is to gain the best estimation from minimal data [5]. [6] used bootstrap method to construct interval estimation for two parameters exponential distribution under double type-II censoring. In [7] bootstrap method was utilised to construct the interval estimation for survivor function for two parameters exponential distribution under double type-II censoring. In this paper the focus is to make comparison of the confidence bands for survivor function obtained by the conventional method and bootstrap percentile method.

2. Methodology and Theory

An example of real data is analysed to illustrate the procedure. The data is an air quality date extracted from the Malaysian Data Report 2000 obtained from the Department of Environment, Ministry of Science, Technology and Environment. The confidence band for the survivor function was first constructed by the traditional approach. From the bootstrap's repeated samples, the convergence condition is determined. This will be fallowed by developing the confidence band for the survivor function. The S-Plus software was used in the development of the programme.

The actual survival time of an individual, t, can be regarded as the value of a variable T, which can take any non-negative value. The survivor function, $S(t)$, is defined to be the probability that the survival time is greater than or equal to t, and so:

$$S(t) = P(T \geq t) = 1 - F(t). \tag{1}$$

The survivor function can therefore be used to represent the probability that an individual survives from the time origin to sometime beyond t [3].

The essential element in lifetime data analysis is the presence of a nonnegative response, t, with appreciable dispersion and often with censoring. Due to sampling methods or the occurrence of some competing risk of removal from the study, several lifetimes of individuals may be censored. By censored data we mean that, in a potential sample of size n, a known number of observations is missing at either end (single censoring) or both ends (double censoring). The type of censoring just described is often called type-II censoring [8].

Suppose some initial observations are censored in addition to some final observations being censored. Out of the n components put to test, suppose the experimenter fails to observe the first r and the last s, observations are then said to be double type II censoring.

$$t_{r+1:n} \leq t_{r+2:n} \leq \ldots \leq t_{n-s:n}. \tag{2}$$

Two Parameters Exponential Distribution

The two parameters exponential distribution has probability density function [10]:

$$f(t;\mu,\theta) = \frac{1}{\theta} \exp\left(-\frac{t-\mu}{\theta}\right) \; ; \; t \geq \mu, \; \mu \geq 0, \; \theta > 0. \tag{3}$$

Here μ is the warranty time and θ is the residual mean life. Once again, it is simple exercise to derive the maximum likelihood estimation of the θ as [1]:

$$\hat{\mu} = t_{r+1:n} + \hat{\theta} \ln\left(\frac{n-r}{n}\right), \tag{4}$$

$$\hat{\theta} = \frac{\sum_{i=r+1}^{n-s} t_{i:n} + s\, t_{n-s:n} - (n-r) t_{r+1:n}}{n-s-r}. \tag{5}$$

The following quantities are independent, with exact sampling distributions:

$$\frac{2n(t_{r+1:n} - \mu)}{\theta} \sim \chi_2^2 \quad \text{and} \quad \frac{2(n-s-r)\hat{\theta}}{\theta} \sim \chi_{(2(n-s-r)-2)}^2 , \tag{6}$$

where $\chi_{(2(n-s-r)-2)}^2$ is the chi-squared distribution with $(2(n-s-r)-2)$ degrees of freedom. It follows that the ratio of these quantities divided by the ratio of their degrees of freedom is a F variable:

$$\frac{2n(t_{1:n} - \mu)}{2\theta} \bigg/ \frac{2(n-s-r)\hat{\theta}}{\theta(2(n-s-r)-2)} = \frac{n(n-s-r-1)(t_{1:n} - \mu)}{(n-s-r)\hat{\theta}} \sim F_{(2,2(n-s-r)-2)}, \tag{7}$$

with $(2,2(n-s-r)-2)$ degrees of freedom.

A two-sided, equal-tail, $(1-\alpha)$ level confidence interval on μ is constructed from the probability statement:

$$\Pr\left(F_{(2,2(n-s-r)-2;\alpha/2)} \le \frac{n(n-s-r-1)(t_{1:n}-\mu)}{(n-s-r)\hat{\theta}} \le F_{(2,2(n-s-r)-2;1-\alpha/2)} \right) = 1-\alpha. \quad (8)$$

The $(1-\alpha)$ confidence intervals for μ is:

$$\left(\hat{\mu} - \frac{(n-s-r)\hat{\theta}\, F_{(2,2(n-s-r)-2;1-\alpha/2)}}{n(n-s-r-1)} \right) = \hat{\mu}_{\min} \le \mu \le \left(\hat{\mu} - \frac{(n-s-r)\hat{\theta}\, F_{(2,2(n-s-r)-2;\alpha/2)}}{n(n-s-r-1)} \right) = \hat{\mu}_{\max}.$$

$$(9)$$

A confidence interval on θ similarly constructed from the probability statement:

$$\Pr\left(\chi^2_{(2(n-s-r)-2;\alpha/2)} \le \frac{2(n-s-r)\hat{\theta}}{\theta} \le \chi^2_{(2(n-s-r)-2;1-\alpha/2)} \right) = 1-\alpha. \quad (10)$$

The $(1-\alpha)$ confidence intervals for θ is:

$$\frac{2(n-s-r)\hat{\theta}}{\chi^2_{(2(n-s-r)-2;1-\alpha/2)}} = \hat{\theta}_{\min} \le \theta \le \frac{2(n-s-r)\hat{\theta}}{\chi^2_{(2(n-s-r)-2;\alpha/2)}} = \hat{\theta}_{\max}. \quad (11)$$

Survivor function on two parameters exponential distribution is:

$$S(t) = \int_t^\infty f(t)\,dt = \int_t^\infty \theta^{-1} \exp(-(t-\mu)/\theta)\,dt = \exp(-(t-\mu)/\theta). \quad (12)$$

The $(1-\alpha)$ confidence for survivor function is:

$$\exp\left(-\frac{(t-\hat{\mu}_{\min})}{\hat{\theta}_{\min}} \right) \langle\ S(t)\ \langle\ \exp\left(-\frac{(t-\hat{\mu}_{\max})}{\hat{\theta}_{\max}} \right). \quad (13)$$

Bootstrap Percentile Method

Bootstrap method is a computer-based method for assigning measures of accuracy to statistical estimates, especially to calculate the confidence interval. Bootstrap itself comes from the phrase "*pull oneself up by one's Bootstrap*" which means to stand up by one's own feet and do with minimal resources. The minimal resource is a minimum data, data that are free from certain assumption or data with no assumption at all about the population distribution. The aim of using bootstrap method is to gain the best estimation from minimal observation. The Bootstrap's percentile procedures for the confidence bands for survivor function on two parameters exponential distribution under double type-II censoring are as follows:

1. give an equal opportunity $1/(n-s-r)$ to every observation of type-II censoring,

2. take $(n-s-r)$ sample with replication,

3. do step 2 until B times in order to get an "*independent bootstrap replications*", $\hat{\beta}^{*1}, \hat{\beta}^{*2}, \ldots, \hat{\beta}^{*B}$, and search for convergence condition. Calculate:

$$S(t)^{*b} = \exp\left(-\frac{(t^{*b}-\mu^{*b})}{\theta^{*b}} \right) \quad \text{with } \mu^{*b} = t^{*b}_{r+1:n} + \hat{\theta}^{*b} \ln\left(\frac{n-r}{n} \right) \quad \text{and}$$

$$\hat{\theta}^{*b} = \frac{\sum_{i=r+1}^{n-s} t^{*b}_{i:n} + s\, t^{*b}_{n-s:n} - (n-r)t^{*b}_{r+1:n}}{n-s-r} \quad (14)$$

4. define the confidence interval at the level $(1-\alpha)$ of the bootstrap percentile for survivor function of one and two parameters exponential distribution under double type-II censoring as:

$$\left[S(t)^{*b\,(\alpha/2)}, S(t)^{*b\,(1-\alpha/2)}\right],\tag{15}$$

5. confidence bands for survivor function are developed using the intervals estimation for survivor functions at every lifetime.

3. Results And Discussion

The data presented in air pollutant (special case: carbon monoxide) data report 2000 Malaysia from Department of Environment, Ministry of Science, Technology and Environment. The first 22 observations in a random sample of 28 lifetimes from carbon monoxide (ppm/part per million)) on 1st December 2000 are as follow:

- ,	- ,	- ,	0.1100,	0.4950,	0.5338,	0.6075,
0.6150,	0.7029,	0.7350,	0.7871,	0.8650,	0.8925,	0.8938,
0.9429,	0.9543,	0.9629,	1.0186,	1.0500,	1.0514,	1.0625,
1.2171,	2.3050,	2.8038,	2.9275,	- ,	- ,	- .

These data are of double type-II censoring. Using Lilliefors test, the data are exponentially distributed. As an illustration we will use these data to construct confidence bands for the survivor function.

Based on the two parameters exponential distribution under double type-II, the intervals estimation for survivor functions at every lifetime, are tabulated in Table 1.

Connecting the intervals estimation of every lifetime develops confidence bands for survivor function with traditional method and bootstrap percentile method.

Comparison of Confidence Bands

Figure 1 and 2 give the confidence bands for survivor function on two parameters exponential distribution under double type-II censoring using the traditional method and the bootstrap percentile method.
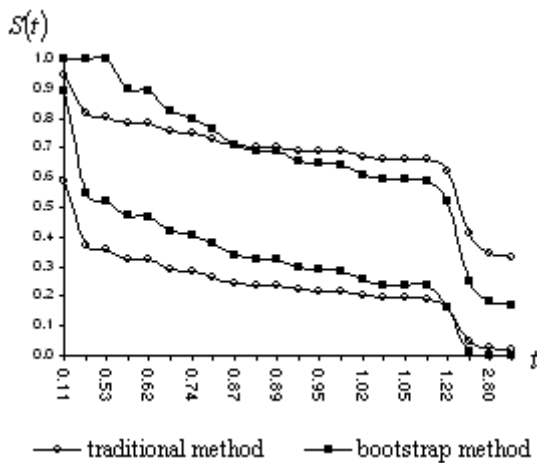

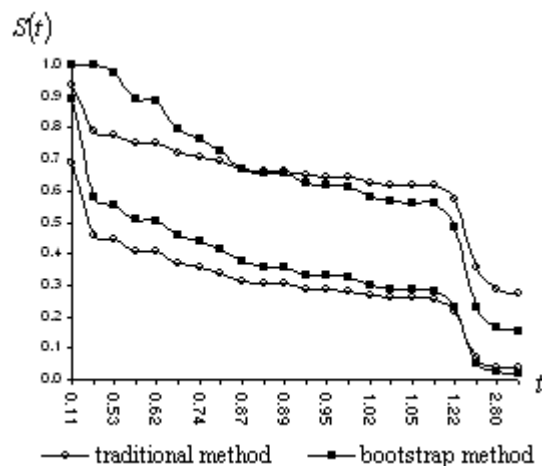Figure 1. 99% Confidence bands for survivor function

Figure 1. 95% Confidence bands for survivor function

From these figures 1 and 2 with 99% and 95% respectively, the width of the confidence regions for the survivor function with bootstrap percentile are narrower compared to the traditional method.

Table 1. The floor (F) and ceiling (C) for survivor functions to every lifetime
at the level of 99% and 95 % with traditional method and bootstrap percentile method

Lifetime	Traditional method				Bootstrap percentile method			
	99%		95%		99%		95%	
	F	C	F	C	F	C	F	C
0.1100	0.58631	0.94449	0.69007	0.93471	0.89286	1.00000	0.89286	1.00000
0.4950	0.37331	0.81770	0.46154	0.78915	0.54638	1.00000	0.57809	1.00000
0.5338	0.35671	0.80591	0.44321	0.77580	0.51999	1.00000	0.55332	0.97144
0.6075	0.32718	0.78398	0.41037	0.75107	0.47333	0.90161	0.50914	0.89286
0.6150	0.32431	0.78178	0.40716	0.74860	0.46883	0.89286	0.50484	0.88509
0.7029	0.29255	0.75647	0.37144	0.72022	0.41910	0.82269	0.45715	0.79586
0.7350	0.28174	0.74743	0.35919	0.71012	0.40228	0.79941	0.44088	0.76524
0.7871	0.26505	0.73300	0.34016	0.69404	0.37642	0.76298	0.41395	0.72364
0.8650	0.24191	0.71193	0.31357	0.67067	0.33921	0.70671	0.37316	0.67121
0.8925	0.23423	0.70464	0.30469	0.66261	0.32555	0.68877	0.35955	0.65420
0.8938	0.23388	0.70429	0.30428	0.66223	0.32501	0.68793	0.35901	0.65347
0.9429	0.22079	0.69146	0.28906	0.64809	0.29753	0.65688	0.33429	0.62463
0.9543	0.21786	0.68852	0.28564	0.64485	0.29205	0.65006	0.32938	0.61866
0.9629	0.21567	0.68631	0.28309	0.64242	0.28681	0.64488	0.32566	0.61362
1.0186	0.20204	0.67214	0.26708	0.62688	0.25797	0.61102	0.30017	0.58057
1.0500	0.19473	0.66429	0.25846	0.61828	0.23951	0.59429	0.28761	0.56416
1.0514	0.19442	0.66394	0.25809	0.61790	0.23904	0.59353	0.28712	0.56351
1.0625	0.19190	0.66119	0.25511	0.61489	0.23442	0.58821	0.28321	0.55775
1.2171	0.16008	0.62400	0.21706	0.57448	0.16171	0.51874	0.22914	0.48607
2.3050	0.04470	0.41524	0.06966	0.35608	0.01097	0.24709	0.05386	0.22642
2.8038	0.02491	0.34451	0.04137	0.28596	0.00324	0.18453	0.02610	0.16437
2.9275	0.02155	0.32891	0.03635	0.27082	0.00240	0.17164	0.02171	0.15208

4. Conclusion

Using the intervals estimation for survivor functions at every lifetime develops the confidence bands for survivor function. Bootstrap percentile method proved to be more potential in constructing confidence bands for survivor function on two parameters exponential distribution under double type-II censoring than the traditional method. Therefore, bootstrap method can be used as an alternative method in constructing confidence bands.

References

[1]. Balakrishnan. (1990). On the Maximum Likelihood Estimation of the Location and Scale Parameters of the Exponential Distribution Based on Multiply Type-II Censored Samples. *Journal of Applied Statistics.* 17:55-61.

[2]. Bartholomew, D. J. (1963). The Sampling Distribution of an Estimate Arising in Life Testing. *Technometrics* 5. 361-374.

[3]. Collett, D. (1996). *Modelling Data In Medical Research.* London: Chapman & Hall.

[4]. Efron, B. (1979). Bootstrap Method: Another Look at the Jackknife. *The Annals of Statistics*, 7(1), 1-26.

[5]. Efron, B. and Tibshirani, R. (1993). *An Introduction to the Bootstrap.* New York: Chapman & Hall.

[6]. Fauzy, A., Ibrahim, N. A., Daud, I. & Abu Bakar, M. R. (2003a). Interval Estimation for One and Two Parameters Exponential Distribution under Double Type-II Censoring with Bootstrap Percentile. *Weekly Seminar to Institute for Mathematical Research.* UPM.

[7].   Fauzy, A., Ibrahim, N. A., Daud, I. & Abu Bakar, M. R. (2003b). Interval Estimation for Survivor Function of Exponential Distribution under Double Type-II Censoring with Bootstrap Percentile. *International Conference on Research and Education in Mathematics*, Institute for Mathematical Research, UPM.

[8].   Fernandez, A. J. (2000). Estimation and Hypothesis Testing for Exponential Lifetime Models with Double Censoring and Prior Information. *Journal of Economic and Social Research*, 2(2), 1-17.

[9].   Hsieh, H. K. (1981). On Testing the Quality of Two Exponential Distributions. *Technometrics* 23. 263-269.

[10].   Lawless. (1982). *Statistical Models and Methods for Lifetime Data.* New York: John Wiley & Sons.

[11].   Raqab, M. Z. (1995). On the Maximum Likelihood Prediction of the Exponential Distribution Based on Double Type-II Censored Samples. *Pakistan Journal of Statistics*, 11, 1-10.

[12].   Singh, N. (1985). A Simple and Asymptotically Optimal Test for the Equality of K Exponential Distribution Based on Type II Censored Samples. *Comm. Statist. Theory Methods.* 14.1615-1625.