Mathematical Theory and Modeling
ISSN 2224-5804 (Paper) ISSN 2225-0522 (Online)
Vol.6, No.3, 2016

A generalized and efficient robust algorithm for handling censored values in one parameter exponential distribution.

Eric Boahen,

Department of Statistics Faculty of Mathematical Sciences University for Development Studies – Tamale.

Kaku Sagary Nokoe,

University of Energy and natural Resources Sunyani, Ghana

ABSTRACT

Often in survival analysis, response that is measured over time is not a continuous measure but is the occurrence of a particular event or the number of such event occurring in a particular interval. Events such as exacerbation or epileptic seizures during each month of follow-up are examples. In this situation it is very difficult to specify specific distribution for the data, but when distribution is not specified, maximum likelihood estimate cannot be used. To specify a functional form for the expectation and the marginal variance, generalized estimation model are used assuming that there is no repeated measurement. Asymptotically, censored distributions are not normal but to achieve consistency and sufficiency in estimate, the derivative of their distributions should be normal, chi-square, follow t-statistics or other well known functions. This is because the derivatives of these distributions are easier to understand and maximum likelihood estimates have minimum variance even if percentage of censored values increases.

***Keywords**:* hazard, censoring, generalize, normalize, Jacobian.

1.0 INTRODUCTION

In survival studies, patient' survival time are subject to paucity of patients who report for treatment. For this reason, clinical data are mostly censored. Estimating mean, variance and standard deviation are different from those in uncensored data. There is a margin of estimation bias because of censored observations. A censored observation is generated by several factors. In the process of the study, some patients may be still alive at the time of analysis. For this reason, these patients are right censored. In addition to this, patients can be lost because of insufficient follow-up. Censored observations occur because patients may loss to follow-up; patients may either drop out of the study completely or stop coming to the study center or may move away to

a different community for treatment. Death caused by competing risk rather than the cause of the disease being treated for is also a major contributing factor that generate censored observations in clinical studies.

The result of censored observation in clinical trial is estimation problem. An increase in proportion of censored observations in an entire clinical data equally increases the margin of bias in statistical estimates. As bias increases, precision decreases and this eventually affect validity of statistical estimates. Affected results in clinical estimates have an effect on statistical power leading to a questionable reliability of statistical inferences. In clinical trial, an already generated censored data are affected by estimation bias. This bias cannot be perfectly avoided but can be minimized. For this reason, statisticians have a duty to develop best models so as to reduce bias. Kaplan -Meier method is a nonparametric approach that does not depend on any assumption. Kaplan – Meier method in estimating parameters is the most widely used. This technique is popularly referred to as product limit. This method had been in use for many years before Kaplan- Meier gave a theoretical justification that , the method behaves as non-parametric maximum likelihood estimate that makes no distributional assumptions about the population. In the Kaplan –Meier (1958) approach, ordered observations are used instead of grouped data. This method has the advantage of yielding results that are not dependent on the choice of the event time interval. Even though this method behaves like maximum likelihood, it is not asymptotic under fairly general condition. Maximum likelihood has sufficient property in basic distributions and its asymptotic properties under general regularity conditions make its use very desirable. Maximum likelihood estimation affords a rather general method of estimation of parameters of survival distributions. This is applicable in censored observation if the distribution of the observations is specified. Censored data sometimes do not follow a specific distribution hence using parametric method in estimating parameters in censored data calls for specific assumptions. Sample statistics based on assumptions can be far from true reflection of the population parameters they are supposed to estimate, hence there is therefore the need for robust procedure in estimating parameters in censored data.

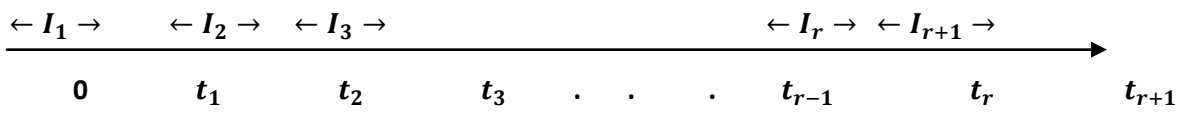
A robust procedure is a procedure where the accuracy of the procedure does not depend too heavily on the distribution being true. This means that even under the edifice of inaccurate assumption, robust procedure can still draw the sample statistic to the true population parameter in censored values.

2.0 NORMALIZED SPACING

It is not possible to admit all patients to the study sample at the same time because of paucity of patients. It is therefore reasonable to accept the patients as they enter into the study for treatment. For this reason, a common time to access the patients survival become a problem.

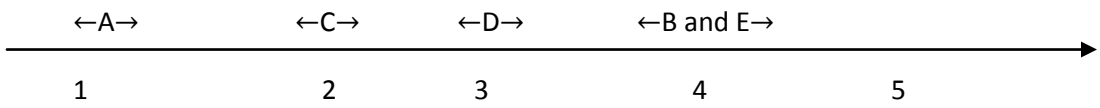
To make general observation, it is useful to put all observations at the same beginning time. At this point censored observations are assumed to be in any interval.

To make an analysis at time t. It is appropriate to put the observations in normalized spacing. Let$t_1, t_2, t_3, \ldots\ t_r,$ be the n patients time with $r\ failure\ and\ let\ t_1, \leq\ t_2, \leq\ t_3, \leq \cdots\ \leq t_r$ be the corresponding order statistics out of n patients. The event time are put into normalize space with interval $I_i = (t_{(i-1)}, t_i]$. This is an open ended interval of real numbers.



Suppose that the i^{th} interval contain m_i censored observations. Denote the time observed for these censored observation to be ${t_1}^i, {t_2}^i, {t_3}^i, \ldots, {t_{mi}}^i$ with order values ${t_1}^i \leq {t_2}^i \leq {t_3}^i \leq \ldots, \leq {t_{mi}}^i$. from table 2. It is seen that failure time are $1 < 2 < 3 < 4 < 5$

Figure 3.13.3



Let us define the total number of observations as N, the $N_1 = N - m_1$, $N_2 = N - m_1 - m_2$, $N_3 = N - m_1 - m_2, - m_3$,

$$N_i = N - \sum_{j=1}^{i} m_i \quad (2.1)$$

$$\mathrm{N} = r + \sum_{j=1}^{r+1} m_j. \quad (2.2)$$

Suppose the probability density function of an observation believe to be drawn at random from a population is $q(t)$ and its cumulative survival density function $Q(t) = P\{T > t\}$. In the i^{th} interval, there is a failure at the right hand side of the interval observations and point with probability $q(t_i)dt_i$ and m_i censored observations with probability $\prod_{j-1}^{m_i} Q(t_j^{(i)})$. The contribution of the i^{th} interval to the likelihood is $q(t_i)dt_i \prod_{j-1}^{m_i} Q(t_j^{(i)})$. From figure 1, we have r such intervals with failures at the right hand end point and two infinite with failure. The likelihood function is

$$\prod_{i=1}^{r} q(t_i)dt_i \prod_{j-1}^{m_i} Q(t_j^{(i)}) \prod_{j-1}^{m_{r+1}} Q(t_j^{(r+1)}) \quad (2.3)$$

To derive the probability density function, it is useful to obtain the constant of proportionality. If all the censored observations are in the last interval, the constant of proportionality is $\frac{n!}{(n-r)!}$, since at the first failure any of n can fail, at the second interval any of $n-1$ can fail, at the third interval any of $n-2$ can fail, and so on. The constant equals $n(n-1),\ \ldots\ , n-r+1$. Using the same idea above, N and m_1 are censored in the first interval, so there are $N-m_1=N_1$ ways that the first failure can occur; m_2 are censored in the interval and based on this second failure rate, there are $N-m_1-m_2-1=N_2-1$ ways it can occur. If there is to be a failure at 3 and m_3 observations are censored in the third interval, then there are $N-m_1-m_2-m_3-2=N_3-2$ ways this can happen. At this point the constant term is

$$N_1(N_2-1)(N_3-2)(N_4-3),\ \ldots(N_r-r+1)=K \tag{2.4}$$

Now the joint pdf of (t_i) when the data assumes exponential and if $q(t)=\lambda e^{-\lambda t}$, $Q\ (t)=e^{-\lambda t}$ is

$$[q(t_i)dt_i\prod_{j-1}^{m_i}Q\left(t_j^{(i)}\right)]=\lambda e^{-\lambda t_i}\prod_{j-1}^{m_i}e^{-\lambda t_j^i}=\lambda e^{-\lambda(t_i+\sum_{j=1}^{m_i}t_j^{(i)})} \tag{2.5}$$

Define S to be $\sum_{j=1}^{m_i}t_j^{(i)})$ which is the sum of censored observations in the i^{th} interval, then

$$f(t)=K\prod_{i=1}^{r}\lambda e^{-\lambda(t_i+S_i)}\lambda e^{-\lambda S_{r+1}}=K\lambda^r e^{-\lambda[\sum_{i=1}^{r}t_i+\sum_{i=1}^{r+1}S_i]}=K\lambda^r e^{-\lambda V} \tag{2.6}$$

Where V=$[\sum_{i=1}^{r}t_i+\sum_{i=1}^{r+1}S_i]$, (2.7)

is the total life observed. This is the same likelihood function in equation (2.3). Clearly, V is sufficient statistics for λ. Denote $\pi_{1=}$ total life observed in the i^{th} interval.

$$\pi_1=(N-m_1)t_1+S_1=N_1t_1+S_1 \tag{2.8}$$

$$\pi_2 = (N - m_1 - m_2 - 1)(t_2 - t_1) + S_2 - m_2 t_1 = (N_2 - 1)(t_2 - t_1) + S_2 - m_2 t_1 \quad (2.9)$$

$$\pi_1 = (N_i - i + 1)t_i - t_{(i-1)} + [S_i - m_i t_{(i-1)}] \quad (2.10)$$

For I = 1, . . , r where $t_0 = G \; or \; 0$

$$\pi_{r+1} = S_{r+1} - m_{r+1} t_r \quad (2.11)$$

The sum of $\pi_i = V$ ie $\sum_{i=1}^{r+1} \pi_i = V$ (2.12)

At this point we will concentrate and estimate the joint distribution of the π_i. This is obvious because each π_i is equal to the total up time observed in the i^{th} interval and V is the total up time observed in the i^{th} intervals.

$$\pi_i = N_1(N_2 - 1)(N_3 - 2)(N_4 - 3), \; \ldots (N_r - r + 1)\lambda^r e^{-\lambda \sum_{i=1}^{r+1} \pi_i} |\frac{\partial t}{\partial \pi}|. \quad (2.13)$$

To get the Jacobean, it is appropriate to get reciprocal $|\frac{\partial \pi}{\partial t}|$ and reciprocate that. The matrix of $|\frac{\partial \pi}{\partial t}|$ whose $(i, j)th$ element is

$$\begin{bmatrix} t_1 & t_2 \; \ldots & t_r \\ N_1 & 0 \; \ldots & 0 \\ -(N_2 - 1) - m_2 & N_2 - 1 \; \ldots & 0 \\ \vdots & \vdots & N_r - r + 1 \end{bmatrix}$$

$\frac{\partial \pi}{\partial t}$ is a triangular matrix so its determinant equals the product of the diagonal terms

$$N_1(N_2 - 1).. \; (N_r - r + 1) = N_i - i + 1. \quad (2.14)$$

Thus

$$|\frac{\partial \pi}{\partial t}| = N_1(N_2 - 1)(N_3 - 2)(N_4 - 3), \; \ldots (N_r - r + 1) = K, \quad (2.15)$$

where $K = {}^{N_i!}/_{(N_i - r)!}$

So $\frac{\partial t}{\partial \pi} = {}^{1}/_{K}$. (2.16)

$$f(\pi) = K\lambda^r e^{-\lambda \sum_{i=1}^{r+1} \pi_i} \, {}^{1}/_{K}. \quad (2.17)$$

$$f(\pi) = {}^{N_i!}/_{(N_i - r)!} \lambda^r e^{-\lambda \sum_{i=1}^{r+1} \pi_i} \, {}^{(N_i - r)!}/_{N_i!} \quad (2.18)$$

$$f(\pi)=\lambda^{r}e^{-\lambda\sum_{i=1}^{r+1}.\pi_i} \quad (2.19)$$

$$f(\pi)=\prod_{i=1}^{r+1}\lambda e^{-\lambda\pi_i} \quad (2.20)$$

It is obvious to see that π_i are independent identically distributed following an exponential distribution. $\sum_{i=1}^{r+1}.\pi_i = V$. in this case V follows a gamma distribution with parameters r and λ

For the gamma distribution

$$f(t)=\frac{(\lambda t)^{k-1}\lambda e^{-\lambda t}}{\Gamma(k)}, \quad (2.21)$$

The kernel of the likelihood for the gamma distribution with parameters r and λ is the same for that of V distribution

$$f(v)=\frac{(\lambda V)^{r-1}\lambda e^{-\lambda V}}{\Gamma(r)} \quad (2.22)$$

The aim is to estimate λ. To do this, first, it is useful to get the distribution of V

Let $\theta=\frac{1}{\lambda}$, $\theta=\frac{v}{r}$, then $E\sum_{i=1}^{r+1}.\pi_i) = V = r\theta$, so $E(\hat{\theta}) = \theta$. This is to say that $\hat{\theta}$ is an unbiased estimate of θ based on the sufficient statistics V provided λ is constant. By using Rao-Blackwell Theorem, $\hat{\theta}$ is the minimum variance unbiased estimator of θ.

From equation (3.13.19)

$$\log f(\pi) = r log\lambda - \lambda V + C \quad (2.23)$$

$$\frac{dL}{d\lambda}=\frac{r}{\lambda}-V. \text{ At } \frac{dL}{d\lambda}=0, \hat{\lambda}=\frac{r}{V} \quad (2.24)$$

$\hat{\lambda}=\frac{r}{v}$ is also a function of v alone. If $\hat{\lambda}$ is not constant, we may get into trouble, this is because an average of λ should be accounted for. In this situation, a robust procedure should be set up to address this issue.

3.0 ROBUST ALGORITHM

A robust procedure is a procedure where the accuracy of the procedure does not depend too heavily on the distribution assumptions being true. To ascertain whether $\hat{\lambda}$ is also minimum variance unbiased estimator, it is useful to estimate E($\hat{\lambda}$) under the procedure of logarithm of $\hat{\lambda}$. This is because, $\hat{\lambda}$ has a normal distribution with mean log($\hat{\lambda}$) and variance $\frac{\sigma^2}{r}$. If the data follow exponential distribution, then

$$X^2 = 2\lambda v = 2r\lambda\frac{v}{r} = 2r\,{}^{\lambda}/_{\hat{\lambda}}\text{ . in terms of } \hat{\lambda},\ \hat{\lambda} = \frac{2r\lambda}{X^2} \tag{2.25}$$

Where X^2 has a chi-square distribution with 2r degrees of freedom. This is because, most text tabulate the chi-square distribution but not the gamma distribution. In this situation, X^2 table is useful to find probability for gamma variates. If V follows a gamma distribution with parameters λ and r, then, it follows that

$$P[V \le v] = P[X^2 \le 2\lambda v]. \tag{2.26}$$

In application

$$\log(\hat{\lambda}) = log2r + \log\lambda - \log X^2. \tag{2.27}$$

Following from equation (2.27), it is useful to find the moment of $\log\hat{\lambda}$ because the mean and variance of $\log\hat{\lambda}$ largely depend on the moment of $\log X^2$. In the theory of moment statistics, if a variate X^2 with chi-square distribution has v degrees of freedom, then the expected expectation of the variate $E(X^2) = v$, the variance $v(X^2) = 2v$. $E(X^2 - v)^3 = 8v$ and $E(X^2 - v)^4 = 48v + 12v^2$. The bases of this is because, chi-square distribution is a special case of the gamma distribution where$\alpha = \frac{v}{2}, \beta = 2$. In relating chi-square to standardized normal distribution

$$\emptyset = \frac{(X^2 - v)}{\sqrt{2v}}. \tag{2.27}$$

It follows immediately that

$$X^2 = v + \emptyset\sqrt{2v} \text{ hence } X^2 = v(1 + \emptyset\sqrt{\frac{2}{v}}) \tag{2.28}$$

From standard normal theory, expectation of

$$E\left(\frac{(X^2 - v)}{\sqrt{2v}}\right) = 0 \text{ and } V\left(\frac{(X^2 - v)}{\sqrt{2v}}\right) = 1 \tag{2.29}$$

This implies that $E(\emptyset) = 0$ and $V(\emptyset) = 1$.

$$\text{E}(\emptyset^3) = \frac{8v}{(\sqrt{2v})^3} \Rightarrow \text{E}(\emptyset^3) = \frac{\sqrt[3]{2}}{\sqrt{v}} \tag{2.30}$$

$\text{E}(\emptyset^4) = \frac{48v + 12v^2}{4v^2} = \frac{12}{v} + 3$. In this case, as $v \to \infty$, $E(\emptyset) = 0$. This is an evidence that, any non normal data with chi-square distribution tend to normal as $v \to \infty$ under the influence of sufficient statistics. From equation (2.28)

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ . \ . \ . \tag{2.31}$$

$$\text{Log } X^2 = logv + \log(1 + \emptyset\sqrt{\frac{2}{v}}) = \log v + \emptyset\sqrt{\frac{2}{v}} - \frac{\emptyset^2}{2}\frac{2}{v} + \frac{\emptyset^3}{3}\frac{\sqrt[3]{2}}{\sqrt[3]{v}} + \ . \ . \ . \quad (2.32)$$

$$E(\text{Log } X^2) = \log v - \frac{1}{v} + \frac{\sqrt[3]{2}}{\sqrt[3]{v}}\left(\frac{\sqrt[3]{2}}{\sqrt{v}}\right) \ . \ . \ . \quad (2.33)$$

$$E(\text{Log } X^2) = \log v + \omega\left(\frac{1}{v}\right). \quad (2.34)$$

Where ω is $[-1 + \frac{2^{2/3}}{v}]$ (2.35)

$$E(\text{Log } X^2 - \log v)^2 = E[\frac{\emptyset^2}{2}\frac{2}{v} - 2\emptyset^3\frac{1}{v}\sqrt{\frac{2}{v}} + higher\ order\ terms\ in\ \emptyset] \quad (2.36)$$

$$E(\text{Log}\frac{X^2}{v})^2 = \frac{2}{v} - \frac{\sqrt[3]{2}}{\sqrt{v}}\frac{\sqrt[3]{2}}{\sqrt[3]{v}} + powers\ of\ reciprocals\ of\ v \quad (2.37)$$

$$V(\log X^2) = \frac{2}{v} + \omega\left(\frac{1}{v}\right) \quad (2.38)$$

Where ω is $[-1 + \frac{2^{2/3}}{v}]$

Let us come back to the estimation of $\hat{\lambda}$. $\hat{\lambda} = X^2$ has 2r degrees of freedom. This is to say that $v = 2r$. For this reason

$$E(\text{Log}\hat{\lambda}) = \log 2\text{r} + \log\lambda - \log 2r + \ \omega\left(\frac{1}{r}\right) \quad (2.39)$$

$$E(\text{Log}\hat{\lambda}) = \log\lambda + \omega\left(\frac{1}{r}\right) \quad (2.40)$$

$$V(\log\hat{\lambda}) = \frac{2}{2r} + \ \omega\left(\frac{1}{r}\right) \quad (2.41)$$

$$V(\log\hat{\lambda}) = \frac{1}{r} + \omega\left(\frac{1}{r}\right) \quad (2.42)$$

$$\sigma(\hat{\boldsymbol{\lambda}}) = \sqrt{\frac{\mathbf{1}}{\boldsymbol{r}} + \boldsymbol{\omega}\left(\frac{\mathbf{1}}{\boldsymbol{r}}\right)}$$

3.0 Confidence interval for $\hat{\lambda}$

Relating gamma distribution with parameter λ and r to chi-square

$$\text{P}[\text{V} \leq v] = P[X^2 \leq 2\lambda\text{V}] \quad (3.1)$$

Since $X^2 = 2\lambda V = 2r\frac{\lambda}{\hat{\lambda}}$ (3.2)

$\hat{\lambda} = {}^{2\lambda r}/_{X^2}$ (3.3)

Noting that $2\hat{\lambda}V$ has a chi-square distribution with 2r degrees of freedom, an exact (1- α)100 Confidence interval for the $\hat{\lambda}$ is given by $P[X^2_{1-\alpha} < X^2 < X^2_{\alpha}] = 1 - 2\alpha$. Replacing $2\lambda V$ for X^2

$$P[X^2_{1-\alpha} < 2\lambda V < X^2_{\alpha}] = P\left[\frac{X^2_{1-\alpha}}{2V} < \lambda < \frac{X^2_{\alpha}}{2V}\right] = 1 - 2\alpha \quad (3.4)$$

For this reason, $1 - 2\alpha$ for λ have the end points at $\frac{X^2_{1-\alpha}}{2V}$ lower point and $\frac{X^2_{\alpha}}{2V}$ upper point.

At v large, $v \to \infty$, $X^2_{\alpha,v}$ tend to normal function hence

$$X^2_{\alpha,v} = v[1 - \frac{2}{9v} + Z_{\alpha}\sqrt{\frac{2}{9v}}]^3 \quad (3.5)$$

APPLICATRION

A sample of 1011 consisting 807 censored observation and 204 event were to validate the model.

The hazard rate is given by 0.007616 as seen in equation. This means that the death rate among the patients who reported for cancer treatment is 0.007616 per month. This death rate has been influenced by 80% censored observations. This affects the reliability and unbiased estimates of the hazard rate. Looking at the data, the probability mass associated with censored values in between the observed failure times have been redistributed out each of the failure observations. The effect contributes to unequal interval hazard rate. As a result of this, there is an interval non constant variation. This shows in a large variance of 3,517,035 as shown in equation (4.1.10b). The estimate for this natural parameter (λ) may not be the one that makes inner product work nicely; there is therefore the need for an estimate of sufficient statistics. $\hat{\mu} = (0.007616)^{-1} = 131.730$ given by equation (4.11) is the sum of the times to death of patients who die on the study. At constant hazard rate, Halperin estimated the expected hazard as $E(\hat{\lambda}) = \frac{r\lambda}{r-1}$ and its variance as$Var(\hat{\lambda}) = \frac{\lambda^2}{r-1}$. These cannot work for non constant λ. It was against this discrepancy that Epstein and Sobel used the mean time to death $\hat{\mu}$ vis-a vis the exponential survival distribution to obtain $\hat{\mu} = \hat{\lambda}^{-1}$. This confirms our model in equation(4.1.10c).

We obtained $E(\hat{\lambda}) = E[log(\hat{\lambda})] = 2.12$ in equation (4.22) which a minimum variance of 7.33×10^{-03} (4.23). Equation (4.22) is the gradient of the log partition function which is the expectation of the sufficient statistics for the natural parameter λ.

Comparing our expectation model to one proposed by Welling (2006) (4.24) $E[\log\hat{\mu})] = \log(E|\mu|) - \frac{V(|\mu|)}{2E(|\mu|)^2} = 2.119684 - 0.049678 = 2.19.$

$\mathrm{E}\left[\log\left(\hat{\lambda}\right)\right] \approx \mathrm{E}[\log\hat{\mu})]$. This means that our model confirms that of Welling. Helsel (2012) indicated that for large amount of censoring greater or equal to 80%, mean, variance and standard deviation estimates are extremely biased. There is therefore the need to indicate appropriate distribution before maximum likelihood is used. The model proposed in equation (4.15) is a robust algorithm which can used to estimate $E\left(\hat{\lambda}\right)$ in exponential distribution from both non constant λ and non exponential distribution. Equation (4.22) allows the estimation of data which does not follow exponential distribution.

6.0 SIMULASTION STUDIES

In this section, we present simulation results from simulation of 1011 independent individuals of cancer patients used in the study. For each individual, the data consist of two parts $t_i, and\ \delta_i$, where t_i is the event time or censoring time, δ_i is the indicator variable with a value of 1 if t_i is uncensored or a value 0 if t_i is censored (status). We take the hazard function $E(log\lambda) = 2.11$ and considered censorship of 80%. The censoring times were generated from log normal distribution with a parameter selected to give 80% censorship in model.

7.0 DISCUSSION

From the simulation table, a sample size of 4000 patients at risk of cancer who report for treatment,760 of these patients are expected to die at a follow-up of 78 weeks which represent a proportion of 0.1905. Comparing this event proportion to the original data from korlebu with an event proportion of 0.20273, the simulated data differ from the original data by a margin of 0.01. This difference is observed because of interval hazard variations in the data. From the simulated data as shown in table, an increase in sample size maintains constant event proportions in the data.From figure 6, the shaded portion represents observations that were covered by follow-up period of 78 weeks. Observations sparsely spaced above the shaded region beyond follow-up period signify those patients who survived at the end of the clinical trial as seen from the original data. For this reason, the simulated data is a true reflection of the original data drawn from the study area.

References

Afifi A. A. and S.P. Azen (1972), statistical analysis: A computer oriented Approach, Academic Press, New York

Aitken, A. C. (1935), On least-squares and Linear Combinations. Proceedings of the Royal Society of Edinburgh

Kaplan, E. L and O. Meier, 1958 nonparametric estimation from incomplete observations *Journal of the American statistical association vol. 53, 457-481*

Armitage, P. 1959 The comparison of survival curves *Journal of the Royal Statistical Society, vol.A122, 279-300*

Berkson J. and R. P Gage 1952 Survival curve for cancer patients following treatment *Journal of the American Statistical Association vol. 47, 501-515*

Anderson, T. W. (1964). Sequential Analysis with Delayed Observations, Journal of the American Statistical Association, Vol.59, pp. 1006-1015

Armitage, P. (1960) Sequential Medical trials, Charles C. Thomas, Springfield, III

Bain, L. J. and C. E. Antle (1967), Estimation of Parameter in the Weibull Distribution, Technometrics, Vol. 9, pp.621-628

Barlow, R. E. and F. proschan (1965). Mathematical Theory of Reliability. John Willey and Sons. New York

APPENDIX

Table 1. table od model

MODEL	HAZARD	MEAN	VARIANCE	STANDARD DEVIATION	COEFFICIENT OF VARIATION	CONFIDENCE INTERVAL
EXPONENTIAL	0.007616	131.73	3517035	1875.37	14.23	-0.00693< $\lambda < 0.02215$
ROBUST	2.119684	0.471768	0.00733	0.005615	0.0026	

Table 4.2. *Plotting position*

t_i	value	$(1 - {}^{i}/_{n+1})$	Value $-\log(1 - {}^{i}/_{n+1})$
1	0	0.000	0.0000
2	2	0.975	0.0109
3	3	0.962	0.0168
4	4	0.949	0.0227
5	5	0.937	0.0282
6	6	0.925	0.0338
7	7	0.924	0.0343
8	8	0.898	0.0467
9	9	0.886	0.0526
10	10	0.873	0.0589
11	11	0.861	0.0649
12	12	0.848	0.0716
13	13	0.835.	0.0783
14	14	0.823	0.0846
15	15	0.810	0.0915
16	16	0.797	0.0985
17	17	0.785	0.1051
18	18	0.772	0.1124
19	19	0.759	0.1197
20	20	0.747	0.1266

21	21	0.734	0.1343
22	22	0.721	0.1421
23	23	0.708	0.1499
24	24	0.696	0.1574
25	25	0.683	0.1655
26	26	0.670	0.1739
27	27	0.658	0.1817
28	28	0.645	0.1904
29	29	0.633	0.1985
30	30	0.620	0.2076
31	31	0.607	0.2168
32	32	0.595	0.2254
33	33	0.582	0.2350
34	34	0.569	0.2448
35	35	0.557	0.2541
36	36	0.544	0.2644
37	37	0.532	0.2740
38	38	0.518	0.2856
39	39	0.506	0.2958
40	40	0.494	0.3062
41	41	0.481	0.3178
42	42	0.468	0.3297
43	43	0.455	0.3419
44	44	0.443	0.3535

45	45	0.430	0.3665
46	46	0.417	0.3798
47	47	0.405	0.3925
48	48	0.392	0.4067
49	49	0.379	0.4213
50	50	0.367	0.4353
51	51	0.354	0.4509
52	52	0.342	0.4659
53	53	0.329	0.4828
54	54	0.316	0.5003
55	55	0.304	0.5171
56	56	0.291	0.5361
57	57	0.278	0.5559
58	58	0.265	0.5767
59	59	0.253	0.5968
60	60	0.240	0.6197
61	61	0.227	0.6439
62	62	0.215	0.6675
63	63	0.202	0.6946
64	64	0.189	0.7235
65	65	0.177	0.7520
66	66	0.164	0.7851
67	67	0.152	0.8181
68	68	0.139	0.8569

69	69	0.126	0.8996
70	70	0.114	0.9430
71	71	0.101	0.9956
72	72	0.088	1.0555
73	73	0.076	1.1191
74	74	0.063	1.2006
75	75	0.050	1.3010
76	76	0.037	1.4317
77	77	0.025	1.6020
78	78	0.012	1.9208

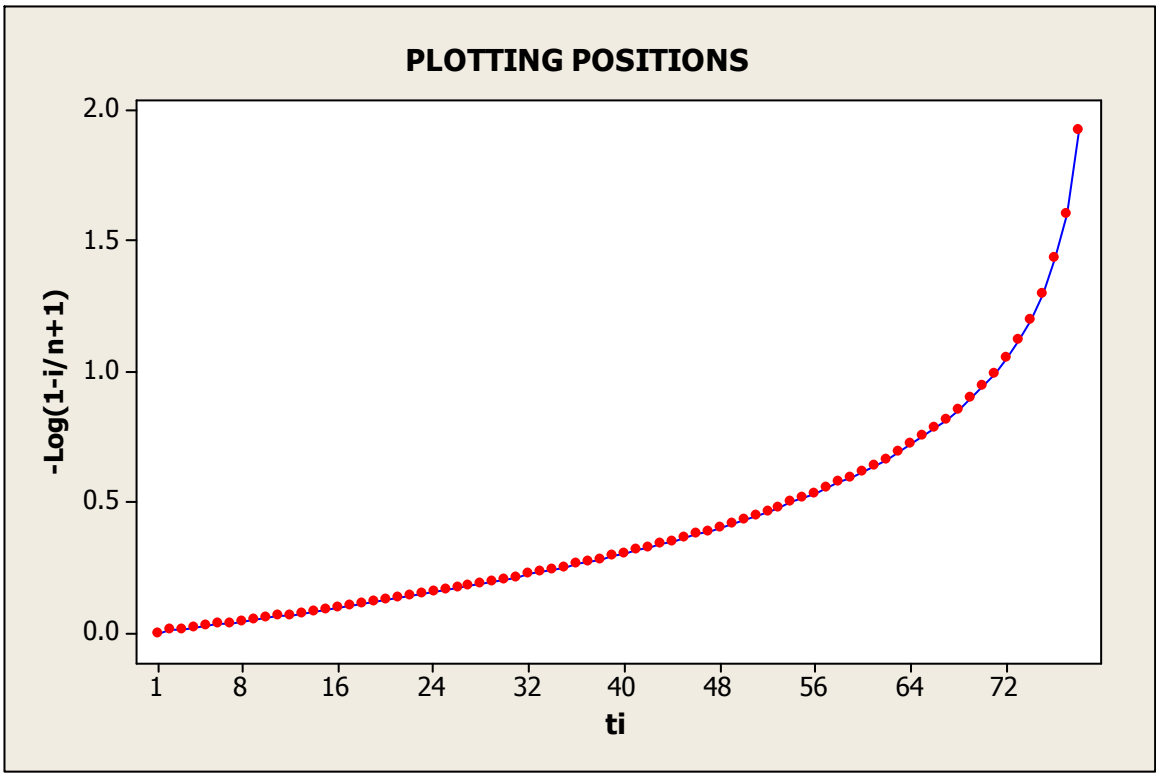

PLOTTING POSITIONS
-Log(1-i/n+1)
2.0
1.5
1.0
0.5
0.0
1
8
16
24
32
40
48
56
64
72
ti

Figure 3.13.1

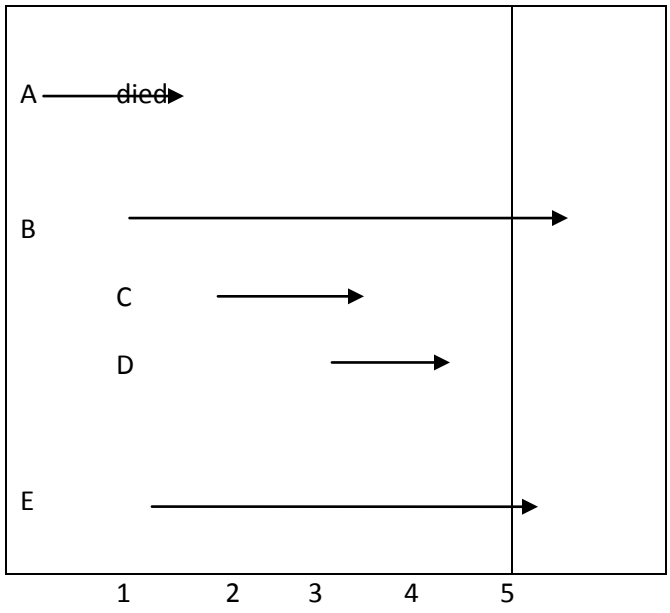
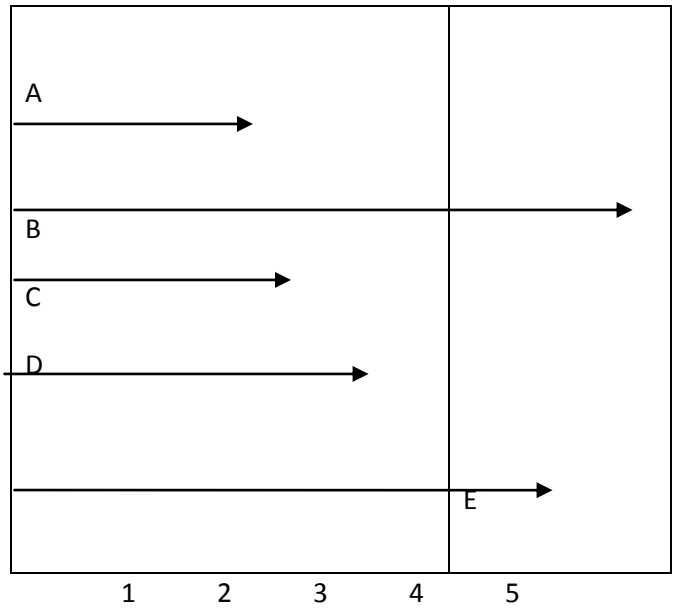

A
died
B
C
D
E
1
2
3
4
5

Table 3.13.2



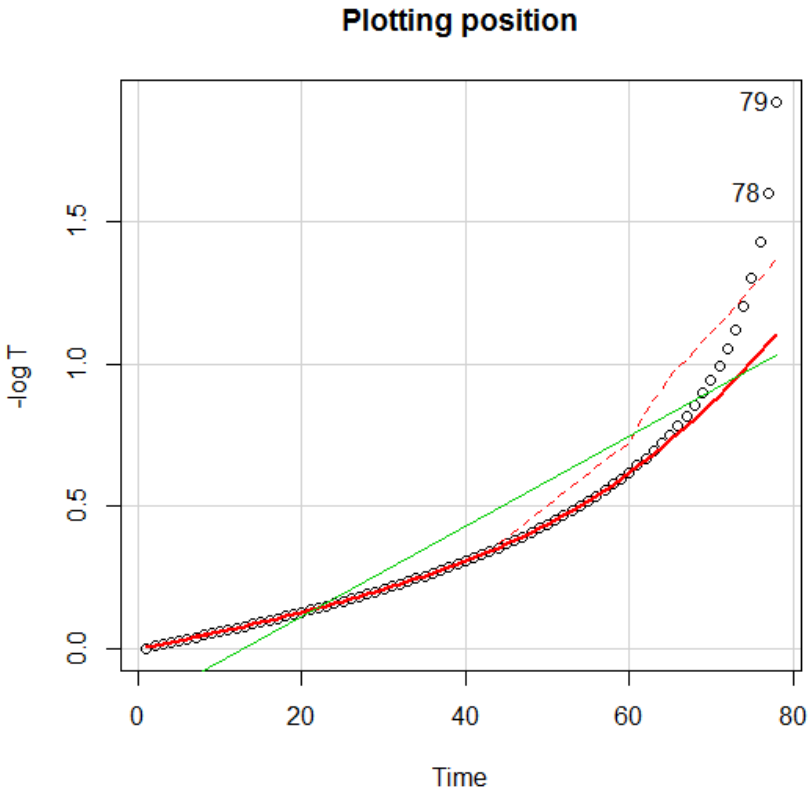
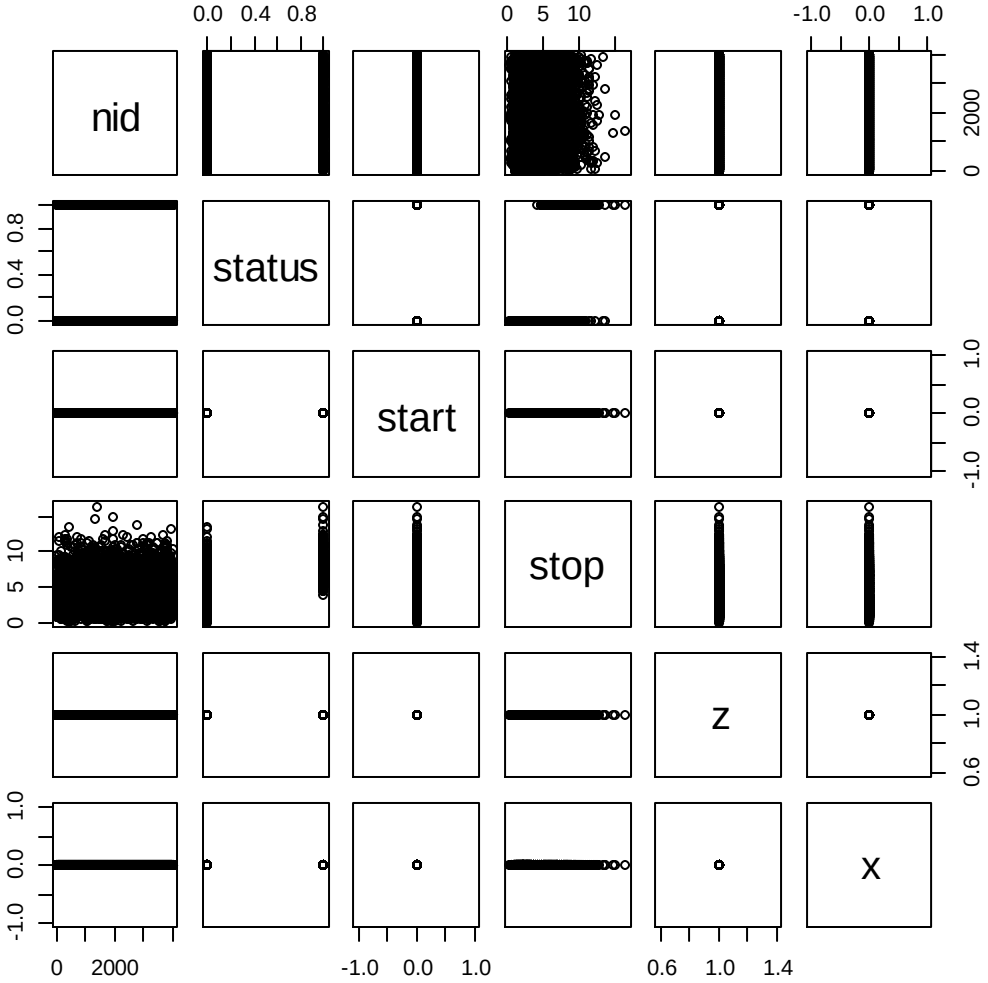


TABLE 4.5 : SIMULATED PARAMETERS

NO. OF PATIENTS	NO. OF EVENT	MEAN EP. OF EVENT	TOTAL TIME OF FOLLOW-UP	MEDIAN FOLLOW-UP TIME	DENSITY OF INCIDENCE
4000	760	0.1905	17982.13	4.015501	0.04237541
10000	1980	0.198	45364.27	4.069318	0.04364669
50000	9799	0.19598	225603.3	4.056644	0.04343465
80000	15655	0.1956875	361085	4.059307	0.04335545
100000	19181	0.19181	449228.4	4.039394	0.04269765
150000	28982	0.1932133	674962.8	4.040854	0.04293866
200000	38533	0.192665	899924.4	4.044289	0.04281804


nid
status
start
stop
z
x
0.0 0.4 0.8
0 5 10
-1.0 0.0 1.0
0 2000
0.0 0.4 0.8
-1.0 0.0 1.0
0 5 10
0.6 1.0 1.4
-1.0 0.0 1.0
0 2000
-1.0 0.0 1.0