Journal Homepage: www.ijrpr.com ISSN: 3049-0103 (Online)

International Journal of Advance Research Publication and Reviews

Vol 02, Issue 07, pp 447-459, July 2025

A Bayesian Approach to Estimating Parameters of the Log-Normal Survival Distribution Under Censored Circumstances with Applications

Ogar Obol Tom[1*]; Etowa Fidelis Vincent[2]; Dr. Emem Okon Peter[3]

[1]*Department of Statistics and Research, College of Nursing Sciences, Calabar, Cross River State, Nigeria*
[2]*Department of Medical Microbiology, College of Nursing Sciences, Calabar, Cross River State, Nigeria*
[3]*Department of Plant Science and Biotechnology, College of Nursing Sciences, Calabar, Cross River State, Nigeria*
**ogardbishop@gmail.com*

ABSTRACT

We gave and implemented a Bayesian estimation procedure for the log-normal survival distribution without covariates, under censored circumstance. This survival distribution in question is one among many available for solving problems in the statistical field of survival analysis, and may be considered under uncensored or censored circumstances (with or without covariates). Perusal and scrutiny of a host of contemporary literature unraveled the presence of only a few studies which suggested the need for adopting the Bayesian alternative procedure, whilst considering the peculiar cases of a log-normal survival distribution; hence, exposing the gap warranting this study. To this end, we adopted the Bayesian estimation procedure for estimating the parameters of the log-normal survival distribution, and consequently deduced the mean and variance of the survival time. Our results showed that one could obtain the parameter estimates of μ and σ, via maximum likelihood estimation, as well as Bayesian estimation procedures under censored circumstance. Our results, more so, confirmed that μ and σ existed for both the MLE and Bayesian procedures under censored circumstance, particularly at $(u, v) = (0, 1)$ being a standard normal instance of our prior used for the simulation done in the study. However, our results confirmed the non-existence of μ_T and σ_T^2 for both the MLE and Bayesian procedures under censored circumstances.

Keywords: Log-normal Survival Distribution, Censored Circumstance, Bayesian Estimation, Normal Informative Prior.

1. INTRODUCTION

1.1 Background of the study

Survival analysis, as a branch of statistics, is one that estimates the expected duration of time until one, or more, occurrences (like: the death in biological organisms, or the failure in mechanical systems) take place [2; 6; 8]. This tool called "survival analysis" has now been applied, in contemporary times, to studies in various fields such as: clinical sciences, engineering, and industry, implying that its applications in most real-life scenarios is unquestionable [1; 5; 9].

Although engineers aim to estimate the time which is available until a machine fails, in the field of engineering, clinical scientists are often concerned with estimating the time available until ailments reoccur, or until an accident victims dies from injuries, or until an individual succumbs to a terminal illness, etc. [4; 5; 7] Sadly, in making such estimations, these non-survival analysts may depend on previous ideas absent the knowledge of survival analysis [9; 10], whereas in such an instance, survival analysis should have been a preferred technique for making non-intuitive estimations, with high precision [5; 11; 16].

During the early years of the 20th century, survival analysts developed survival distributions in survival analysis studies. These distributions are themselves statistical distributions that were only specifically redesigned to cater for the needs of survival analysis. These needs majorly comprise obtaining the survivorship functions and hazard functions [10; 12; 16].

However, since developing survival analysis in the early years of the 20th century, about six statistical distributions, namely: the exponential, Weibull, log-normal, gamma, generalized gamma, and log-logistic distributions, have been redesigned for use in survival analysis studies, with what is known as the "exponential survival distribution" being the most popular amongst the lot [3; 6]. Notwithstanding, the log-normal survival distribution has also been extensively utilized owing to its efficiency and reliability in modelling times to failure [3; 7; 13].

Log-normal survival distributions are a type of survival distributions that are redesigned from the log-normal statistical distribution as it were, but having minor interpretational variations [3; 12; 14]. In comparison with the log-normal statistical distribution, the log-normal survival distribution can be utilized to deduce the hazard and survivorship functions. Overtime, for all survival distributions, as well as the log-normal survival distribution in particular, all parameters have been estimated through the maximum likelihood estimation (MLE) procedure; no doubt, this has produced formidable results [4; 10; 13]. Regardless, in this article, we have attempted the adoption of a Bayesian estimation alternative, as tendencies are that this could produce stronger theoretical and applied results.

1.2 Statement of the problem

The method of maximum likelihood estimation – MLE is often used for estimating the parameters of survival distributions, in survival analysis. Until the recent adoption of the Bayesian technique in survival analysis studies, MLE had remained the only estimation approach. Regardless, from sufficient review of available literature on survival distributions, a majority of researchers who have investigated this matter have approached it usually disregarding the nature and sensitivity of the survival distributions whose parameters they intend to estimate. For instance, concerns about how the parameters of the distributions should be estimated under censored circumstances seem not be considered at all. [15] was one (among several) attempts to surmount this pitfall, using a reference Bayesian approach for the estimating the parameters of the generalized log-normal distribution in the presence of survival data. Notwithstanding, their study did not consider the cases of censored observations. Addressing the gap using the log-normal survival distribution with an informative prior is, thus, the interest of this study.

1.3 Aim and objectives of the study

This study adopted a Bayesian approach to estimating parameters of the log-normal survival distribution without covariates (under censored circumstance) with applications. In order to achieve the stated aim, the objectives of the study were to: (i) review the MLE procedure for estimating the parameters of the log-normal survival distribution, (ii) formulate the Bayesian estimation alternative to the MLE procedure, and (iii) establish any necessary theorem and axiom based on findings.

2. MATERIALS AND METHODS

2.1 General maximum likelihood estimation procedure

2.1.1 Estimation procedures for data with right-censored observations

Suppose that n persons were followed to their deaths or censored in a study. Let $t_1, t_2, \dots, t_r,\ \ t_{r+1}^+, t_{r+2}^+, \dots, t_n^+$ be the survival times observed from the n individuals, with r exact times and $(n-r)$ right-censored times. Assume that the survival times follow a distribution with the density function $f(t, \boldsymbol{b})$ and survivorship function $S(t, \boldsymbol{b})$, where $\boldsymbol{b} = (b_1, \dots, b_p)$ denotes unknown p parameters $b_1, \dots, b_p$ in the distribution. If the survival time is discrete (i.e., it is observed at discrete time only), $f(t, \boldsymbol{b})$ represents the probability of observing t and $S(t, \boldsymbol{b})$ represents the probability that the survival or event time is greater than t. In other words, $f(t, \boldsymbol{b})$ and $S(t, \boldsymbol{b})$ represent the information that can be obtained respectively from an observed uncensored survival time and observed right-censored survival time. Thus, the product $\prod_{i=1}^{n} f(t_i,\ \boldsymbol{b})$ represents the joint probability of observing the uncensored survival times, and $\prod_{i=r+1}^{n} S(t_i^+,\ \boldsymbol{b})$ represents the joint probability of those right-censored survival times. The product of these two probabilities, denoted by $L(\boldsymbol{b})$,

$$L(\boldsymbol{b}) = \prod_{i=1}^{n} f(t_i,\ \boldsymbol{b}) \prod_{i=r+1}^{n} S(t_i^+,\ \boldsymbol{b})$$

represents the joint probability of observing $t_1, t_2, \dots, t_r,\ t_{r+1}^+, t_{r+2}^+, \dots, t_n^+$. A similar interpretation applies to continuous survival $L(\boldsymbol{b})$ is called the likelihood function of $\boldsymbol{b}$, which can also be interpreted as a measure of the likelihood of observing a specific set of survival times $t_1, t_2, \dots, t_r,\ t_{r+1}^+, t_{r+2}^+, \dots, t_n^+$, given a specific set of parameters $\boldsymbol{b}$. The method of the MLE is to find an estimator of $\boldsymbol{b}$ that maximizes $L(\boldsymbol{b})$, or in other words, which is "most likely" to have produced the observed data $t_1, t_2, \dots, t_r,\ t_{r+1}^+, t_{r+2}^+, \dots, t_n^+$. Take the logarithm of $L(\boldsymbol{b})$ and denote it by $l(\boldsymbol{b})$,

$$l(\boldsymbol{b}) = log\, L\,(\boldsymbol{b}) = \sum_{\boldsymbol{i=1}}^{\boldsymbol{r}} log[f(t_i, \boldsymbol{b})] + \sum_{\boldsymbol{i=r+1}}^{\boldsymbol{r}} log[S(t_i^+, \boldsymbol{b})] \quad (1)$$

Then the MLE $\widehat{\boldsymbol{b}}$ is a $\boldsymbol{b}$ is the set of $\hat{b}_1, \hat{b}_2, \dots, \hat{b}_p$ that maximizes $l(\boldsymbol{b})$:

$$l(\widehat{\boldsymbol{b}}) = \underset{allb}{max}(l(\boldsymbol{b}))$$

It is clear that $\widehat{\boldsymbol{b}}$ is a solution of the following simultaneous equations, which are obtained by taking the derivative of $l(\boldsymbol{b})$ with respect to each b_j:

$$\frac{\partial l(\boldsymbol{b})}{\partial b_j} = 0 \qquad j = 1,2,\dots,p \quad (2)$$

To obtain the MLE $\widehat{\boldsymbol{b}}$, one can use a numerical method. A commonly used numerical method is the Newton-Raphson iterative procedure, which can be summarized as follows.

i. Let the initial values $b_1, \dots, b_p$ be zero; that is, let

$$\boldsymbol{b}^{(0)} = 0$$

ii. The changes for $\boldsymbol{b}$ at each subsequent step, denoted by $\Delta^{(j)}$, is obtained by taking the second derivative of the log-likelihood function:

$$\Delta^{(j)} = \left[-\frac{\partial^2 l(\boldsymbol{b}^{(j-1)})}{\partial \boldsymbol{b} \partial \boldsymbol{b}'}\right]^{-1} \frac{\partial l(\boldsymbol{b}^{(j-1)})}{\partial \boldsymbol{b}} \quad (3)$$

iii. Using $\Delta^{(j)}$, the value of $\boldsymbol{b}^{(j)}$ at j^{th} step is

$$\boldsymbol{b}^{(j)} = \boldsymbol{b}^{(j-1)} + \Delta^{(j)} \qquad j = 1,2,3,\dots$$

The iteration terminates at, say, the m^{th} step if $\|\Delta^{(m)}\| < \delta$, where δ is a given precision, usually a very small value, 10^{-4} or 10^{-5}. Then the MLE $\widehat{\boldsymbol{b}}$ is defined as

$$\widehat{\boldsymbol{b}} = \boldsymbol{b}^{(m-1)} \quad (4)$$

The estimated covariance matrix of the MLE $\widehat{\boldsymbol{b}}$ is given by

$$\overset{\Lambda}{var}(\widehat{\boldsymbol{b}}) = \overset{\Lambda}{cov}(\widehat{\boldsymbol{b}}) = \left[-\frac{\partial^2 l(\widehat{\boldsymbol{b}})}{\partial \boldsymbol{b} \partial \boldsymbol{b}'}\right]^{-1} \quad (5)$$

One of the good properties of a MLE is that if $\widehat{\boldsymbol{b}}$ is the MLE of $\boldsymbol{b}$, then $g(\widehat{\boldsymbol{b}})$ is the MLE of $g(\boldsymbol{b})$ if $g(\boldsymbol{b})$ is a finite function and need not be one-to-one.

The estimated $100(1-\alpha)\%$ confidence interval for any parameter b_i is

$$\left(\widehat{\boldsymbol{b}}_{\boldsymbol{i}} - \boldsymbol{Z}_{\boldsymbol{\alpha/2}}\sqrt{\boldsymbol{v}_{\boldsymbol{ii}}}\widehat{\boldsymbol{b}}_{\boldsymbol{i}} + \boldsymbol{Z}_{\boldsymbol{\alpha/2}}\sqrt{\boldsymbol{v}_{\boldsymbol{ii}}}\right) \quad (6)$$

where v_{ii} is the i^{th} diagonal element of $\hat{V}(\widehat{\boldsymbol{b}})$ and $Z_{\alpha/2}$ is the $100(1-\alpha/2)$ percentile point of the standard normal distribution $[P(Z > Z_{\alpha/2}) = \alpha/2]$. For a finite function $g(b_i)$ of b_i, the estimated $100(1-\alpha)\%$ confidence interval for $g(\boldsymbol{b_i})$ is its respective range R on the confidence interval in equation (6), that is,

$$R = \left\{g(\boldsymbol{b_i}) : \boldsymbol{b_i} \in \left(\widehat{\boldsymbol{b}}_{\boldsymbol{i}} - \boldsymbol{Z}_{\boldsymbol{\alpha/2}}\sqrt{\boldsymbol{v}_{\boldsymbol{ii}}}\widehat{\boldsymbol{b}}_{\boldsymbol{i}} + \boldsymbol{Z}_{\boldsymbol{\alpha/2}}\sqrt{\boldsymbol{v}_{\boldsymbol{ii}}}\right)\right\} \quad (7)$$

In case $g(b_i)$ is monotone in b_i, the estimated $100(1-\alpha)\%$ confidence interval for $g(b_i)$ is

$$R = \{g(\widehat{\boldsymbol{b}}_i - \boldsymbol{Z}_{\alpha/2}\sqrt{\boldsymbol{v}_{ii}}),\ g(\widehat{\boldsymbol{b}}_i + \boldsymbol{Z}_{\alpha/2}\sqrt{\boldsymbol{v}_{ii}})\} \quad (8)$$

2.1.2 Estimation procedures for data with right-, left-, and interval-censored observations

If the survival times $t_1, t_2, \dots, t_n$ observed for the n persons consist of uncensored left-, right-, and interval-censored observations, the estimation procedures are similar. Assume that the survival times follow a distribution with the density function $f(t, \boldsymbol{b})$ and the survivorship function $S(t, \boldsymbol{b})$, where $\boldsymbol{b}$ denotes all unknown parameters of the distribution. Then the log-likelihood function is

$$\left.\begin{aligned} l(\boldsymbol{b}) = log\, L\,(\boldsymbol{b}) = \sum log[f(t_i, \boldsymbol{b})] + \sum log[S(t_i, \boldsymbol{b})] \\ + \sum log[1 - S(t_i, \boldsymbol{b})] + \sum log[S(v_i, \boldsymbol{b}) - S(t_i, \boldsymbol{b})] \end{aligned}\right\} \quad (9)$$

where the first sum is over the uncensored observations, the second sum over the right-censored observations, the third sum over the left-censored observations, and the last sum over the interval-censored observations, with v_i as the lower end of a censoring interval. The other steps for obtaining the MLE $\widehat{\boldsymbol{b}}$ of $\boldsymbol{b}$ are similar to the steps shown in section (2.1.1) by substituting the log-likelihood function defined in equation (1) with the log-likelihood function in equation (9).

2.2 Log-normal distribution

If the survival time T follows the log-normal distribution with density function given by equation (10), then the mean and the variance are respectively $exp\left(\mu + \frac{1}{2}\sigma^2\right)$ and $[exp(\sigma^2) - 1]exp(2\mu + \sigma^2)$.

$$f(t) = \frac{1}{t\sigma\sqrt{2\pi}} exp\left[-\frac{1}{2\sigma^2}(logt - \mu)^2\right] \quad (10)$$

Estimation of the two parameters μ and σ^2 has been investigated either by using equation (10) directly or by using the fact that $Y = \log T$ follows the normal distribution with mean μ and variance σ^2.

2.2.1 Estimation of μ and σ^2 for data without censored observations

Estimations of μ and σ^2 for complete samples by maximum likelihood methods have been studied by many authors. But the simplest way to obtain estimates of μ and σ^2 with optimum properties is by considering the distribution of $Y = \log T$. Let $t_1, t_2, t_3, \dots, t_n$ be the survival times of n subsets. The MLE of μ is the sample mean of Y given by:

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} logt_i \quad (11)$$

The MLE of σ^2 is

$$\hat{\sigma}^2 = \frac{1}{n}\left[\sum_{i=1}^{n}(logt_i)^2 - \frac{(\sum_{i=1}^{n} logt_i)^2}{n}\right] \quad (12)$$

The estimate $\hat{\mu}$ is also unbiased but $\hat{\sigma}^2$ is not. The best unbiased estimates of μ and σ^2 are $\hat{\mu}$ and the sample variance $s^2 = \hat{\sigma}^2[n/(n-1)]$. If n is moderately large, the difference between s^2 and $\hat{\sigma}^2$ is negligible.

One of the properties of the MLE is that if $\hat{\theta}$ is the MLE of θ, $g(\hat{\theta})$ is the MLE of $g(\theta)$ if $g(\theta)$ is a finite function. Therefore, the MLEs of the mean and variance of T are, respectively, $exp\left(\hat{\mu} + \frac{1}{2}\hat{\sigma}^2\right)$ and $[exp(\hat{\sigma}^2) - 1]exp(2\hat{\mu} + \hat{\sigma}^2)$.

It is known that $\hat{\mu} = \hat{y}$ is normally distributed with mean μ and variance σ^2/n. Hence, if σ is known, a $100(1-\alpha)\%$ confidence interval for μ is $\hat{\mu} \pm Z_{\alpha/2}\,\sigma/\sqrt{n}$. If σ is unknown, we can use Student's t-distribution. A $100(1-\alpha)\%$ confidence interval for μ is $\hat{\mu} \pm t_{\alpha/2,(n-1)}\, s/\sqrt{n-1}$, where $t_{\alpha/2,(n-1)}$ is the $100\alpha/2$ percentage point of Student's t-distribution with $n-1$ degrees of freedom.

Confidence intervals for σ^2 can be obtained by using the fact that $n\hat{\sigma}/\sigma^2$ has a chi-square distribution with $n-1$ degrees of freedom. A $100(1-\alpha)\%$ confidence interval for σ^2 is

$$\frac{n\hat{\sigma}^2}{\chi^2_{(n-1),\alpha/2}} < \sigma^2 < \frac{n\hat{\sigma}^2}{\chi^2_{(n-1),1-\alpha/2}} \tag{13}$$

2.2.2 Estimation of μ and σ^2 for data with censored observations

We first consider samples with singly censored observations. The data consist of r exact survival times $t_{(1)} \leq t_{(2)} \leq \cdots \leq t_{(r)}$ and $n-r$ right-censored survival times that are at least $t_{(r)}$ denoted by $t^{+}_{(r)}$. Furthermore, we use the fact that $Y = logT$ has normal distribution with mean μ and variance σ^2. Estimates of μ and σ^2 can be obtained from the transformed data $y_i = logt_i$. Many authors have investigated the estimation of μ and σ^2.

The best linear estimates of μ and σ^2 proposed by Saharan and Greeberg are linear combinations of the logarithms of the r exact survival times:

$$\hat{\mu} = \sum_{i=1}^{r} a_i logt_{(i)} \tag{14}$$

$$\hat{\sigma} = \sum_{i=1}^{r} b_i logt_{(i)} \tag{15}$$

where the coefficients a_i and b_i are calculated and tabulated by Saharan and Greeberg for $n \leq 20$.

MLEs for the log-normal distribution can be used for $n > 20$. Let

$$\bar{y} = \frac{1}{r}\sum_{i=1}^{r} logt_{(i)} \tag{16}$$

and

$$s^2 = \frac{1}{r}\left[\sum \left(logt_{(i)}\right)^2 - \frac{\left(\sum logt_{(i)}\right)^2}{r}\right] \tag{17}$$

Then the MLEs of μ and σ^2 are

$$\hat{\mu} = \bar{y} - \hat{\lambda}\left(\bar{y} - logt_{(r)}\right) \tag{18}$$

and

$$\hat{\sigma}^2 = s^2 + \hat{\lambda}\left(\bar{y} - logt_{(r)}\right)^2 \tag{19}$$

where the value of $\hat{\lambda}$ has been tabulated by Cohen in 1961 as a function of a and b. The proportion of censored observations, b, is calculated as

$$b = \frac{n-r}{n}$$

and

$$a = \frac{1 - Y(Y-c)}{(Y-c)^2}$$

where $Y = [b/(1-b)]f(c)/F(c)$, $f(c)$ and $F(c)$ being the density and distribution functions, respectively, of the standard normal distribution, evaluated at:

$$c = \left(logt_{(r)} - \mu\right)/\sigma$$

2.3 The Bayesian estimation procedure

Let $x_1, x_2, x_3, \ldots, x_n$ be a random sample from the density $f(x;\theta)$. Before taking the sample, the distribution of θ, $g(\theta)$ is assumed known. Hence, $g(\theta)$ is called a prior distribution. The task is to know the distribution $f(\theta|x)$, after taking the sample. Hence, $f(\theta|x)$ is called a posterior distribution.

Let us consider the conditional distribution

$$f(x|\theta) = \frac{f(x;\theta)}{g(\theta)}$$

$$\Rightarrow f(x;\theta) = f(x|\theta)g(\theta) \quad (20)$$

$$\Rightarrow f(\theta|x) = \frac{f(x;\theta)}{h(x)} \quad (21)$$

Substituting for equation (20) in equation (21) gives,

$$\Rightarrow f(\theta|x) = \frac{f(x|\theta)g(\theta)}{h(x)} \quad (22)$$

But $\int_\Omega f(\theta|x)\, d\theta = 1$

Therefore,

$$\int_\Omega f(\theta|x)\, d\theta = \int_\Omega \frac{f(x|\theta)g(\theta)}{h(x)} d\theta = 1$$

$$\Rightarrow 1 = \frac{1}{h(x)} \int_\Omega f(x|\theta)g(\theta)\, d\theta$$

$$\Rightarrow h(x) = \int_\Omega f(x|\theta)g(\theta)\, d\theta \quad (23)$$

Putting equation (23) into equation (22) gives

$$\Rightarrow f(\theta|x) = \frac{f(x|\theta)g(\theta)}{\int_\Omega f(x|\theta)g(\theta)d\theta} \quad (24)$$

Since we are taking a random sample of this distribution

$$f(x|\theta) = L(x|\theta) = \prod_{i=1}^{n} f(x_i|\theta)$$

Hence, equation (24) becomes:

$$\Rightarrow f(\theta|x) = \frac{L(x|\theta)g(\theta)}{\int_\Omega L(x|\theta)g(\theta)d\theta} \quad (25)$$

The above equation (25) gives $f(\theta|x)$ as the posterior Bayes distribution with respect to the prior distribution $g(\theta)$.

Hence,

$$E[\tau(\theta)] = \int_{\Omega} \tau(\theta) f(\theta|x) d\theta \qquad (26)$$

is called the posterior Bayes estimator with respect to the prior distribution $g(\theta)$; where $\tau(\theta)$ is any function of θ.

2.4 Proposed procedure of the study

The proposed Bayesian alternative will be implemented with the procedure below.

Step 1: Determine an appropriate prior $\pi(\mu)$.

The appropriate prior for the log-normal distribution is a normal distribution of μ with mean u and v^2. That is,

$$\pi(\mu) = \frac{1}{v\sqrt{2\pi}} exp\left[-\frac{1}{2v^2}(\mu - u)^2\right]$$

Step 2:Obtain the Bayesian estimates of μ and σ^2 for data with censored observations.

Step 2(a): Deduce the mean remission time in this instance.

Step 2(b): Deduce the variance of the remission time in this instance.

3. RESULTS

3.1 Theorem 1 (censored case of log-normal survival distribution)

Suppose that a random sample of size n is drawn from a log-normal distribution with unknown mean μ and known variance σ^2. Also, suppose that the prior distribution of μ is also log-normal with mean u and variance v^2. Then the posterior distribution of μ is also log-normal, with mean and variance given by:

$$\hat{\mu} = \frac{u\sigma^2 + v^2 \sum_{i=1}^{n} logt_i + v^2 \sum_{i=r+1}^{n} logt_i^+}{\sigma^2 + nv^2}; \qquad \hat{\sigma}^2 = \frac{\sigma^2 v^2}{\sigma^2 + nv^2}$$

Proof 1:

$$\pi(\mu|t) = \frac{f(t|\mu)\pi(\mu)}{\int_{-\infty}^{\infty} f(t|\mu)\pi(\mu) d\mu}$$

$$\pi(\mu|t) \propto f(t|\mu)\pi(\mu)$$

The likelihood function is given by:

$$\prod_{i=1}^{n} f(t|\mu) = \prod_{i=1}^{r} \left\{\frac{1}{t_i\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(logt_i - \mu)^2}\right\} \prod_{i=r+1}^{n} \left\{\frac{1}{t_i^+\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(logt_i^+ - \mu)^2}\right\}$$

$$= \frac{1}{\prod_{i=1}^{r} t_i\, \sigma^r (2\pi)^{\frac{r}{2}}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{r}(logt_i-\mu)^2} \frac{1}{\prod_{i=r+1}^{n} t_i^+\, \sigma^{n-r} (2\pi)^{\frac{n-r}{2}}} e^{-\frac{1}{2\sigma^2}\sum_{i=r+1}^{n}\left(logt_i^+-\mu\right)^2}$$

$$= \frac{1}{\prod_{i=1}^{r} t_i\, \sigma^r (2\pi)^{\frac{r}{2}}} \frac{1}{\prod_{i=r+1}^{n} t_i^+\, \sigma^{n-r} (2\pi)^{\frac{n-r}{2}}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{r}(logt_i-\mu)^2} e^{-\frac{1}{2\sigma^2}\sum_{i=r+1}^{n}\left(logt_i^+-\mu\right)^2}$$

$$= \frac{1}{\left[\prod_{i=1}^{r} t_i\, \sigma^r (2\pi)^{\frac{r}{2}}\right]\left[\prod_{i=r+1}^{n} t_i^+\, \sigma^{n-r} (2\pi)^{\frac{n-r}{2}}\right]} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{r}(logt_i-\mu)^2-\frac{1}{2\sigma^2}\sum_{i=r+1}^{n}\left(logt_i^+-\mu\right)^2}$$

$$= \frac{1}{\prod_{i=1}^{r} t_i \prod_{i=r+1}^{n} t_i^+\, \sigma^n (2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{r}(logt_i-\mu)^2+\sum_{i=r+1}^{n}\left(logt_i^+-\mu\right)^2\right]}$$

The prior distribution of μ is given by:

$$\pi(\mu) = \frac{1}{v\sqrt{2\pi}} exp\left[-\frac{1}{2v^2}(\mu-u)^2\right]$$

The posterior distribution is:

$$\pi(\mu|t) \propto exp\left[-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{r}(logt_i-\mu)^2 + \sum_{i=r+1}^{n}(logt_i^+-\mu)^2\right] - \frac{(\mu-u)^2}{2v^2}\right]$$

$$\pi(\mu|t) \propto exp\left[-\frac{1}{2}\left[\frac{\sum_{i=1}^{r}(logt_i-\mu)^2 + \sum_{i=r+1}^{n}(logt_i^+-\mu)^2}{\sigma^2} + \frac{(\mu-u)^2}{v^2}\right]\right]$$

$$\propto exp\left[-\frac{1}{2}\left[\frac{\sum_{i=1}^{r}(logt_i)^2 - 2\mu\sum_{i=1}^{r} logt_i + r\mu^2 + \sum_{i=r+1}^{n}(logt_i^+)^2 - 2\mu\sum_{i=r+1}^{n} logt_i^+ + (n-r)\mu^2}{\sigma^2} + \frac{\mu^2-2\mu u+u^2}{v^2}\right]\right]$$

$$\propto exp\left[-\frac{1}{2}\left[\frac{v^2\sum_{i=1}^{r}(logt_i)^2 - 2\mu v^2\sum_{i=1}^{n} logt_i + n\mu^2 v^2 + v^2\sum_{i=r+1}^{n}(logt_i^+)^2 - 2\mu v^2\sum_{i=r+1}^{n} logt_i^+ + \mu^2\sigma^2 - 2\mu u\sigma^2 + u^2\sigma^2}{\sigma^2 v^2}\right]\right]$$

Dropping all terms that do not involve μ gives:

$$\pi(\mu|t) \propto exp\left[-\frac{1}{2}\left[\frac{\mu^2\sigma^2 + n\mu^2 v^2 - 2\mu v^2\sum_{i=1}^{n} logt_i - 2\mu v^2\sum_{i=r+1}^{n} logt_i^+ - 2\mu u\sigma^2}{\sigma^2 v^2}\right]\right]$$

$$\pi(\mu|t) \propto exp\left[-\frac{1}{2}\left[\frac{\mu^2(\sigma^2+nv^2)-2\mu(v^2\sum_{i=1}^{n} logt_i + v^2\sum_{i=r+1}^{n} logt_i^+ + u\sigma^2)}{\sigma^2 v^2}\right]\right]$$

Dividing the numerator and denominator by $\sigma^2 + nv^2$ gives:

$$\pi(\mu|t) \propto exp\left[-\frac{1}{2}\left[\frac{\mu^2 - 2\mu\left(\frac{u\sigma^2+v^2\sum_{i=1}^{n} logt_i+v^2\sum_{i=r+1}^{n} logt_i^+}{\sigma^2+nv^2}\right)}{\frac{\sigma^2 v^2}{\sigma^2+nv^2}}\right]\right]$$

Completing the square in μ gives:

$$\propto exp\left[-\frac{1}{2}\left[\frac{\mu^2 - 2\mu\left(\frac{u\sigma^2+v^2\sum_{i=1}^{n} logt_i+v^2\sum_{i=r+1}^{n} logt_i^+}{\sigma^2+nv^2}\right)+\left(\frac{u\sigma^2+v^2\sum_{i=1}^{n} logt_i+v^2\sum_{i=r+1}^{n} logt_i^+}{\sigma^2+nv^2}\right)^2}{\frac{\sigma^2 v^2}{\sigma^2+nv^2}}\right]\right]$$

$$\pi(\mu|t) \propto exp\left[-\frac{1}{2}\left[\frac{\left(\mu - \frac{u\sigma^2+v^2\sum_{i=1}^{n} logt_i+v^2\sum_{i=r+1}^{n} logt_i^+}{\sigma^2+nv^2}\right)^2}{\frac{\sigma^2 v^2}{\sigma^2+nv^2}}\right]\right]$$

This implies that μ is normally distributed with:

$$\hat{\mu} = \frac{u\sigma^2 + v^2\sum_{i=1}^{n} logt_i + v^2\sum_{i=r+1}^{n} logt_i^+}{\sigma^2+nv^2}; \qquad \hat{\sigma}^2 = \frac{\sigma^2 v^2}{\sigma^2+nv^2}$$

Therefore, the following axiom is established:

<u>Axiom 1:</u>

(a) The mean remission time (that is, the mean of T) is given as:

$$\mu_T = exp\left[\frac{u\sigma^2 + v^2\sum_{i=1}^{n} logt_i + v^2\sum_{i=r+1}^{n} logt_i^+}{\sigma^2+nv^2} + \frac{1}{2}\left(\frac{\sigma^2 v^2}{\sigma^2+nv^2}\right)\right]$$

(b) The variance of the remission time (that is, the variance of T) is given as:

$$\sigma_T^2 = \left[exp\left(\frac{\sigma^2 v^2}{\sigma^2+nv^2}\right) - 1\right] exp\left[2\left(\frac{u\sigma^2 + v^2\sum_{i=1}^{n} logt_i + v^2\sum_{i=r+1}^{n} logt_i^+}{\sigma^2+nv^2}\right) + \frac{\sigma^2 v^2}{\sigma^2+nv^2}\right]$$

3.2 Simulation

3.2.1 Log-normal distribution with censored observations via MLE

Suppose that in a study of the efficacy of a new drug, 12 mice with tumors are given the drug. The experimenter decides to terminate the study after 9 mice have died. The survival times are in weeks, 5, 8, 9, 10, 12, 15, 20, 21, 25, 25+, 25+, and 25+, as shown in Table 1. Assume that the times to death of these mice follow the log-normal distribution. In this case $n = 12$, $r = 9$, and $n - r = 3$. Using equations (14) and (15) (where the coefficients a_i and b_i are as calculated and tabulated by Saharan and Greenberg for $n \leq 20$), $\hat{\mu}$ and $\hat{\sigma}$ can be calculated as:

$$\hat{\mu} = 0.036 log5 + 0.0581 log8 + 0.0682 log9 + 0.0759 log10 + 0.0827 log12 + 0.0888 log15 + 0.0948 log20 + 0.1006 log21 + 0.3950 log25$$

$$\Rightarrow \hat{\mu} = 2.811$$

$$\hat{\sigma} = -0.2545 log5 - 0.1487 log8 - 0.1007 log9 - 0.0633 log10 - 0.0308 log12 - 0.0007 log15 + 0.0286 log20 + 0.0582 log21 + 0.5119 log25$$

$$\Rightarrow \hat{\sigma} = 0.747$$

TABLE 1

Survival Times (In Weeks) of 12 Mice Test Drug Subjects

t_i	$logt_i$		$logt_i^+$
5	0.6970		
8	0.9031		
9	0.9542		
10	1.0000		
12	1.0792		
15	1.1761		
20	1.3010		
21	1.3222		
25	1.3979		
25+			1.3979
25+			1.3979
25+			1.3979
Total	9.8307		4.1937
Mean		1.1687	
Variance		0.6269	

3.2.2 Log-normal distribution with censored observations via Bayesian Estimation

Using the same case study in 3.2.2 we compute the $\hat{\mu}$ and $\hat{\sigma}^2$, at say $u = 0$ and $v = 1$. Thus, we have that:

$$\hat{\mu} = \frac{u\sigma^2 + v^2 \sum_{i=1}^{n} logt_i + v^2 \sum_{i=r+1}^{n} logt_i^+}{\sigma^2 + nv^2} = \frac{(0)(0.6269) + (1)(9.8307) + (1)(4.1937)}{(0.6269) + (5)(1)}$$

$$\Rightarrow \hat{\mu} = 2.4924$$

$$\hat{\sigma}^2 = \frac{(0.6269)(1)}{(0.6269) + (5)(1)} \cong 0.1114$$

3.2.3 Discussion of results

Table 4 summarizes the simulation of the results. But the results of this study are summarized in Table 2 and Table 3. Table 2 showed the established result from the stated theorem, in which case the parameter estimates of μ and σ, using the maximum likelihood estimation and Bayesian estimation procedures under uncensored circumstance was obtained. But Table 3 showed the established results from deduced axioms in which case the values of μ_T and σ_T^2, using the maximum likelihood estimation and Bayesian estimation procedures under censored circumstance is obtained.

Our study confirms the existence of μ and σ for both the MLE and Bayesian procedures under a censored circumstance, particularly at the specific choice of $(u, v) = (0, 1)$, a standard normal instance of our prior used for the simulation in the study. Moreso, our results confirm the non-existence of μ_T and σ_T^2 for both the MLE and Bayesian procedures under censored circumstances.

TABLE 2

Established Results from Stated Theorems

Case	$\hat{\mu}$	$\hat{\sigma}^2$
Censored	Maximum Likelihood Estimate	
	$\sum_{i=1}^{r} a_i logt_{(i)}$	$\sum_{i=1}^{r} b_i logt_{(i)}$
	Bayesian Estimate	
	$\frac{u\sigma^2 + v^2 \sum_{i=1}^{n} logt_i + v^2 \sum_{i=r+1}^{n} logt_i^+}{\sigma^2 + nv^2}$	$\frac{\sigma^2 v^2}{\sigma^2 + nv^2}$

TABLE 3

Established Results from Deduced Axioms

Case	μ_T	σ_T^2
Censored	Maximum Likelihood Estimate	
	Nil	Nil
	Bayesian Estimate	
	Nil	Nil

TABLE 4

Simulation Results

Case	$\hat{\mu}$	$\hat{\sigma}^2$
Uncensored	Maximum Likelihood Estimate	
	2.811	0.747
	Bayesian Estimate	
	2.4924	0.1114

4. CONCLUSION

To conclude, our study proposed and implemented a Bayesian alternative estimation procedure on the log-normal survival distribution (without covariates, under censored circumstance) with which parameters, μ and σ were estimated under. For both of the estimated parameters, two axioms were deduced about the mean and variance of the survival time. Our results showed that one could obtain the parameter estimates of μ and σ, via maximum likelihood estimation as well as Bayesian estimation procedures under censored circumstance; it also confirmed that parameters of the log-normal distribution existed whether through the MLE or Bayesian procedure, under censored circumstance, especially for the case of a standard normal prior. Regardless, our study also confirmed that μ_T and σ_T^2 did not exist for both the MLE and Bayesian procedures under censored circumstance.

REFERENCES

1. Abrikah, S. A., Farraj, G. & Hammad, M. A. (2023). Bayes estimators of the log-normal distribution parameters using Lindley's approximation. International Conference on Information Technology, 2023, 149-152.

2. Ajagbe, O. B., Kabir, Z. & O'Connor, T (2014). Survival analysis of adult tuberculosis disease. *Plos One, 9*(11).

3. Amico, M. & Keilegom, I. V. (2018). Cure models in survival analysis. *Annual Review of Statistics and its Applications, 5,* 311-342.

4. Baghestani, A. R., Oghaddam, S. S., Majd, H. A., Akbari, M. E., Nafissi, N. & Gohari, K. (2015). Survival analysis of patients with breast cancer using Weibull parametric model. *Asian Pacific Journal of Cancer Prevention, 16*(18), 8567-8571.
5. Clark, T. G., Bradburn, M. J., Love, S. B. & Altman, D. G. (2003). Survival analysis part 1: Basic concepts and first analyses. *British Journal of Cancer, 89*, 232-238.

6. Danrimi, M. L. & Abubakar, H. (2023). A Bayesian framework for estimating Weibull distribution parameters: Application in finance, insurance, and natural disaster analysis. *UMYU Journal of Accounting and Finance Research, 5*(1), 3-15.

7. Fawei, T. E. L., Orumie, U. C. & Biu, E. O. (2023). Log-normal distribution approach for forecasting the spread of COVID-19 in Nigeria. *American Journal of Applied Mathematics and Statistics, 11*(1), 13-21.

8. Jackson, C. H., Sharples, L. D. & Thompson, S. G. (2010). Survival models in health economics evaluations: Balancing fit and parsimony to improve prediction. *The International Journal of Biostatistics, 6*(1), 1-31.

9. Khan, N., Akhtar, T. & Khan, A. A. (2016). A Bayesian approach to survival analysis of inverse Gaussian model with Laplace approximation. *International Journal of Statistics and Applications, 6*(6), 391-398.

10. Kim, J. (2016). Survival analysis. *Research and Statistics, 33*(4), 172-174.

11. Lin, J., Sinha, D., Lipsitz, S. & Polpo, A. (2012). Semiparametric Bayesian survival analysis using models with log-linear median. *Biometrics, 68*(14), 1136-1145.

12. Muse, A. H., Mwalili, S., Ngesa, O., Almalki, S. J. & Abd-Elmougod, G. A. (2021). Bayesian and classical inference for the

generalized log-logistic distribution with applications to survival data. *Computational Intelligence and Neuroscience,* 1-24.

13. Paolucci, I., Lin, Y., Silva, J. A. M., Brock, K. K. & Odisio, B. C. (2023). Bayesian parametric models for survival prediction in medical applications. *BMC Medical Research Methodology, 23*(250), 1-25.

14. Srivastava, A. K. (2019). Bayesian approach in estimation of shape and scale parameter of Log-Weibull model. *International Journal of Innovative Science, Engineering and Technology, 6*(12), 185-196.

15. Tomazella, V. L. D., De Jesus, S. R., Louzada, F., Nadarajah, S. & Ramo, P. L. (2020). Reference Bayesian analysis for the generalized lognormal distribution with application to survival data. *Statistics and Its Interface, 13*, 139-149.

16. Triana, Y. & Purwadi, J. (2019). Exponential distribution parameter estimation with Bayesian SELF method in survival analysis. *Journal of Physics: Conference Series, 1373*, 1-8.