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

International Journal of Advance Research Publication and Reviews

Vol 02, Issue 07, pp 460-471, July 2025

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

Ogar Obol Tom[1*]; Okoi Essien Jasper[2]; Ettah Otu Otu[3]; Egbe James Egbe[4]

[1]Department of Statistics and Research, College of Nursing Sciences, Calabar, Cross River State, Nigeria
[2]Department of Statistics, University of Cross River State, Calabar, Cross River State, Nigeria
[3]Department of Physiology, College of Nursing Sciences, Calabar, Cross River State, Nigeria
[4]Department of Microbiology, College of Nursing Sciences, Calabar, Cross River State, Nigeria
**ogardbishop@gmail.com*

ABSTRACT

In this article, we proposed and implemented a Bayesian approach to estimating parameters of the log-normal survival distributions (without covariates) under uncensored circumstance. The log-normal survival distributions are only one out of several survival distributions used for tackling problems in survival analysis. The research reviewed several recent literatures which unraveled the absence of scholarly works addressing the need for an alternative procedure that considered these peculiar cases of a log-normal survival distribution; in turn, this gap in literature birthed a necessity for this research. Thereon, the Bayesian estimation technique was used to estimate the parameters of the log-normal survival distributions, and to also deduce the mean and variance of the survival time. The results of the research showed how one could obtain the parameter estimates of μ and σ, through maximum likelihood estimation and Bayesian estimation procedures under uncensored circumstance. The results of the study also confirmed that μ and σ existed for both MLE and Bayesian procedures under uncensored circumstance, particularly at $(u, v) = (0, 1)$ – which is a standard normal instance of our prior used for the simulation in the study.

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

1. INTRODUCTION

1.1 Background of the study

In statistics, survival analysis is a branch which estimates the expected duration of time until one, or more, occurrences (such as: death in biological organisms, and failure in mechanical systems) happen [1; 3; 5]. Survival analysis has found applications, in recent times, in the fields of: clinical sciences, engineering, and industry, ultimately implying that survival analysis applications in most real-life scenarios cannot be overemphasized [6; 9].

Whereas in the field of engineering, process engineers aim to estimate the time which is available until a machine fails, clinical scientists may be interested in estimating the time available until cancer reoccurs, or until a badly-injured individual dies due to his injuries, or until a terminal illness overwhelms an individual, etc. [4; 8; 9] In order to make such estimations, these experienced professionals may rely on intuition for making such estimations [7; 8]. In this instance, survival analysis becomes a preferred way for making non-intuitive based estimation of such times described, with high level of precision [2; 12; 13].

Overtime, in performing survival analysis, survival analysts have developed what are referred to as "survival distributions". Survival distributions are statistical distributions specifically redesigned to attend to the demands of survival analysis, such as obtaining the survivorship functions, hazard functions, etc. [4; 13; 14] Since the development of survival analysis in the early 1940s, as a very vital bio-statistical tool, about six statistical distributions such as the: exponential, Weibull, log-normal, gamma, generalized gamma, and log-logistic distributions, have been used for studies on survival analysis, with the exponential survival distribution being the most commonly used distribution out of the lot [2; 3]. Notwithstanding, the log-normal survival distribution is also a survival distribution that has been extensively used because of its efficacy in reliability applications to model failure times [1; 10; 11].

The log-normal survival distribution is a survival distribution redesigned from the known log-normal statistical distribution with a slight interpretational variation [3; 5; 9]. Compared to the conventional log-normal statistical distribution, the log-normal survival distribution may be used to deduce the hazard function and survivorship function. For all the survival distributions, and of course the log-normal survival distribution, all parameters have been estimated overtime via the general maximum likelihood estimation (MLE) procedure, and this has no doubt yielded formidable results [4; 10; 13]. Notwithstanding, in this research, attempt was made at adopting a Bayesian estimation alternative as this could lead to deeper theoretical and practical results, comparatively.

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 emergence of attempts at estimating the parameters of survival distributions using Bayesian technique, MLE had remained the only estimation approach. Notwithstanding, 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 uncensored cases seem not be considered at all. [14] was one (among several) attempts to overcome this pitfall, as they used 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 and uncensored 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 utilized a Bayesian approach to estimating parameters of the log-normal survival distribution with applications. In line with achieving 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 $l(\boldsymbol{b})$:

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

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 $g(\boldsymbol{b}_{\boldsymbol{i}})$ is its respective range R on the confidence interval in equation (6), that is,

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

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

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

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

If the survival times $t_1, t_2, \ldots, t_n$ 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

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

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 $\hat{\boldsymbol{b}}$ 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 $t_1, t_2, t_3, \ldots, t_n$ 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

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

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:

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

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

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

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

and

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

Then the MLEs of μ and σ^2 are

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

and

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

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 $x_1, x_2, x_3, \dots, x_n$ 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

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

$$\Rightarrow f(x;\theta) = f(x|\theta)g(\theta) \tag{20}$$

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

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

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

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

Therefore,

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

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

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

Putting equation (23) into equation (22) gives

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

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:

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

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 without 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 (uncensored 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 normal with mean u and variance v^2. Then the posterior distribution of μ is log-normal, with mean and variance given by:

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

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:

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

$$\Rightarrow \prod_{i=1}^{n} f(t|\mu) = \frac{1}{\prod_{i=1}^{n} t_i \, \sigma^n (2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(logt_i - \mu)^2}$$

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:

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

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

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

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

Dropping all terms that do not involve μ gives:

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

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

Dividing the numerator and denominator by σ^2+nv^2 gives:

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

Completing the square in μ gives:

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

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

This implies that μ is normally distributed with:

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

Therefore, the following axiom is established:

Axiom 1:

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

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

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

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

3.2 Simulation

3.2.1 Log-normal distribution without censored observations via MLE

Five melanoma (resected) patients receiving immunotherapy BCG are followed. The remission durations in weeks are, in order of magnitude, 8, 16, 23, 27, and 28, as shown in Table 1. Suppose that the remission times follow a lognormal distribution. Estimate of the parameters may be obtained as follows.

$$\hat{\mu} = \frac{14.615}{5} = 2.923$$

$$\hat{\sigma}^2 = \frac{1}{5}\left[43.806 - \frac{1}{5}(14.615)^2\right] = 0.217$$

$$s^2 = \frac{5\hat{\sigma}^2}{5-1} = 0.271$$

The mean remission time is $exp\left(2.923 + \frac{0.217}{2}\right)$, or 20.728 weeks, and the standard deviation of the remission times is $\left[[exp(0.217) - 1]exp(5.846 + 0.217)\right]^{\frac{1}{2}}$ or 10.204 weeks. A 95% confidence interval for μ is:

$$2.973 - 2.776\left(\frac{0.521}{\sqrt{4}}\right) < \mu < 2.923 + 2.776\left(\frac{0.521}{\sqrt{4}}\right) = (2.200, 3.646)$$

A 95% confidence interval for σ^2 is:

$$\frac{5(0.217)}{11.1433} < \sigma^2 < \frac{5(0.217)}{0.4844} = (0.097, 2.240)$$

TABLE 1

Remission Durations of Melanoma Patients

t_i	$logt_i$	$(logt_i)^2$
8	2.079	4.322
16	2.773	7.690
23	3.135	9.828
27	3.296	10.864
28	3.332	11.102
Total	14.615	43.806
Mean	2.923	
Variance	0.217	

3.2.2 Log-normal distribution without censored observations via Bayesian Estimation

Using the same case study in 3.2.1 we compute the $\hat{\mu}$ and $\hat{\sigma}^2$, at say $u = 0$ and $v = 1$. Thus, we have that:

$$\hat{\mu} = \frac{u\sigma^2 + v^2 \sum_{i=1}^{n} logt_i}{\sigma^2 + nv^2} = \frac{(0)(0.217) + (1)(14.615)}{(0.217) + (5)(1)} \cong 2.8014$$

$$\hat{\sigma}^2 = \frac{\sigma^2 v^2}{\sigma^2 + nv^2} = \frac{(0.217)(1)}{(0.217) + (5)(1)} \cong 0.0416$$

$$s^2 = \frac{5\hat{\sigma}^2}{5-1} = \frac{5(0.0416)}{5-1} = 0.052$$

The mean remission time is $exp\left(2.8014 + \frac{0.0416}{2}\right)$, or 16.814 weeks, and the standard deviation of the remission times is $\left[[exp(0.0416) - 1]exp(5.6028 + 0.0416)\right]^{\frac{1}{2}}$ or 3.465 weeks. A 95% confidence interval for μ is:

$$2.8014 - 2.776\left(\frac{0.228}{\sqrt{4}}\right) < \mu < 2.8014 + 2.776\left(\frac{0.228}{\sqrt{4}}\right) = (2.485, 3.117864)$$

A 95% confidence interval for σ^2 is:

$$\frac{5(0.0416)}{11.1433} < \sigma^2 < \frac{5(0.0416)}{0.4844} = (0.019, 0.429)$$

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 uncensored circumstance is obtained.

Our study confirms the existence of μ and σ for both the MLE and Bayesian procedures under an uncensored 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.

TABLE 2

Established Results from Stated Theorems

Case	$\hat{\mu}$	$\hat{\sigma}^2$
Uncensored	Maximum Likelihood Estimate	
	$\frac{1}{n}\sum_{i=1}^{n} logt_i$	$\frac{1}{n}\left[\sum_{i=1}^{n}(logt_i)^2 - \frac{(\sum_{i=1}^{n} logt_i)^2}{n}\right]$
	Bayesian Estimate	
	$\frac{u\sigma^2 + v^2 \sum_{i=1}^{n} logt_i}{\sigma^2 + nv^2}$	$\frac{\sigma^2 v^2}{\sigma^2 + nv^2}$

TABLE 3

Established Results from Deduced Axioms

Case	μ_T	σ_T^2
Uncensored	Maximum Likelihood Estimate	
	$exp\left(\hat{\mu}+\frac{1}{2}\hat{\sigma}^2\right)$	$[exp(\hat{\sigma}^2)-1]exp(2\hat{\mu}+\hat{\sigma}^2)$
	Bayesian Estimate	
	$exp\left[\frac{u\sigma^2+v^2\sum_{i=1}^{n}logt_i}{\sigma^2+nv^2}+\frac{1}{2}\left(\frac{\sigma^2v^2}{\sigma^2+nv^2}\right)\right]$	$\left[exp\left(\frac{\sigma^2v^2}{\sigma^2+nv^2}\right)-1\right]exp\left[2\left(\frac{u\sigma^2+v^2\sum_{i=1}^{n}logt_i}{\sigma^2+nv^2}\right)+\frac{\sigma^2v^2}{\sigma^2+nv^2}\right]$

TABLE 4

Simulation Results

Case	$\hat{\mu}$	$\hat{\sigma}^2$
Uncensored	Maximum Likelihood Estimate	
	2.923	0.217
	Bayesian Estimate	
	2.8014	0.0416

4. CONCLUSION

In conclusion, this research has proposed and implemented a Bayesian alternative estimation procedure on the log-normal survival distributions with which the parameters, μ and σ, have been estimated under uncensored circumstances, without covariates. From the estimated parameters, two axioms have also been deduced about the mean and variance of the survival time. The results of the study showed that one could obtain the parameter estimates of μ and σ, via maximum likelihood estimation as well as Bayesian estimation procedures under uncensored circumstance; it also confirmed that parameters of the log-normal distribution existed whether through MLE or Bayesian procedure, under uncensored circumstances, especially for the case of a standard normal prior.

REFERENCES

1. Aelen, O. O. (1994). Effects of frailty in survival analysis. *Statistical Methods in Medical Research, 12*(5), 34-56.
2. Aliyu, S. & Bemanga, M. A. (2024). Estimation and comparison of Weibull-Normal distribution with some other probability models using Bayesian method of estimation. *Science World Journal, 19*(1), 92-99.
3. Aristizabal, R. J. (2012). Estimating the parameters of the three-parameter log-normal distribution. *FIU Electronic Thesis and Dissertation,* 575.
4. Bartos, F., Aust, F. & Haaf, J. M. (2022). Informed Bayesian survival analysis. *BMC Medical Research Methodology, 22*(1), 1-22.

5. Consul, J., Osaisai, E., Bunakiye, J. & Erho, J. (2023). Parametric modeling using Bayesian approach. *UNILAG Journal of Mathematics and Statistics, 2*(1), 23-36.

6. Emmert-Streib, F. & Dehmer, M. (2019). Introduction to survival analysis in practice. *Machine learning and knowledge extraction, 1*, 1013-1038.

7. Iddrisu, A., Alhassan, A. & Amidu, N. (2019). Survival analysis of birth defect infants and children with Pneumonia mortality in Ghana. *Hidawi – Advances in Public Health*, 1-7.

8. Khaleel, B. J. & Ibrahim, W. S. (2023). Comparison of estimating the survival function for the two-parameter Lindley distribution by the Bayesian and maximum likelihood methods. *Journal of Statistical Science, 20*, 85-100.

9. Kharazmi, O., Hamedani, G. G. & Cordeiro, G. M. (2023). Log-mean distribution: applications to medical data, survival regression, Bayesian and non-Bayesian discussion with MCMC algorithm. *Journal of Applied Statistics, 50*(5), 1152-1177.

10. Lee, E. T. & Go, O. T. (1997). Survival analysis in public health research. *Annual Review of Public Health 18*(1), 105-134.

11. Molinares, C. A. (2011). Parametric and Bayesian modeling of reliability and survival analysis. *USF Tampa Graduate Theses and Dissertation,* 1-164.

12. Ngari, M. M., Schmitz, S., Maronga, C., Mramba, L. K. & Vaillant, M. (2021). A systematic review of the quality of conduct and reporting of survival analyses of tuberculosis outcomes in Africa. *BMC Medical Research Methodology, 21*(89), 1-14.

13. Skelly, B. P., Clipp, H. L., Landry, S. M., Rogers, R., Phelps, Q., Anderson, J. T. & Rota, C. T. (2023). A flexible Bayesian approach for estimating survival probabilities from age-at-harvest data. *Methods in Ecology and Evolution,* 1061-1073.

14. 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.