Austin Biometrics and Biostatistics

Open Access

Austin Publishing Group

Research Article

Inference on the Loglogistic Model with Right Censored Data

Chris B. Guure*
Department of Biostatistics, School of Public Health, University of Ghana, Ghana

***Corresponding author:** Chris B. Guure, Department of Biostatistics, School of Public Health, University of Ghana, Ghana

Received: November 07, 2014; **Accepted:** April 12, 2015; **Published:** April 15, 2015

Abstract

Survival Analysis Methods are commonly used to analyze clinical trial data. In most clinical studies, the time until the occurrence of an event is the main outcome of significance. Clinical trials are conducted to assess the worth of new treatment regimens. The major events that the trial subjects seek to determine are either death, development of an undesirable reaction, relapse from remission, or the progress of a new disease entity. In order to model time-to-event data or clinical trials data, a parametric distribution can be assumed. We have in this study assumed that the data follow a log-logistic distribution. To estimate the parameters of this lifetime distribution, the Bayesian estimation approach is considered under the assumption of informative (gamma) priors as well as the frequentist estimation method. The Bayes estimators cannot be obtained in close forms; therefore, approximate Bayesian estimates are computed using the idea of Lindley. The clinical trial data considered in this study is either randomly or non-informatively censored. These types of data occur when each subject has a censoring time that is statistically independent of their failure times. A simulation study is carried out and also three different sets of real data have been analyzed in order to examine our methods. The Bayesian methods are considered under squared error and linear exponential loss functions.

Keywords: Bayesian Inference; Maximum likelihood; Squared Error and LINEX Loss Functions

Introduction

The log-logistic survival model is a lifetime distributional model which can be used as an alternative to the well-known and used Weibull distribution in lifetime or clinical trials data analysis. The shape parameter of the log-logistic distribution performs similar functions as that of the Weibull distribution. It is important that sometimes we model the survival or clinical trial data using a distribution that has a non-monotone hazard rate. According to [1], when the shape parameter is say p > 1, the hazard function becomes unimodal and when p ≤ 1, the hazard decreases monotonically. The fact that the cumulative distribution function can be written in closed form unlike the lognormal distribution makes it useful for analyzing survival data. The loglogistic model has the distribution, density and survival functions respectively as

$$F(t;\theta,p)=\left[1+\left(\frac{t}{\theta}\right)^{-p}\right]^{-1}$$

$$f(t;\theta p)=\frac{p}{\theta}\left(\frac{t}{\theta}\right)^{p-1}\left[1+\left(\frac{t}{\theta}\right)^{p}\right]^{-2}\quad \forall t>0,\theta,p>0$$

$$s(t;\theta,p)=\left[1+\left(\frac{t}{\theta}\right)^{p}\right]^{-1}$$

where p is the shape parameter and θ the scale parameter.

The log logistic distribution is a continuous probability distribution which has non-negative random variables, hence, it can be used in survival analysis as a parametric model for events whose rate increases initially and decreases consequently, For instance, mortality of cancer patients following diagnoses or treatments. See for instance, [2-5].

According to [6], the log logistic distribution has been shown to be a suitable model in analyzing survival or clinical data was considered by Cox, Cox and Oakes, Bennet and others. [7], employed the log logistic distribution on lung cancer data and in their study, they estimated the mortality ratio at which it reached a maximum level. They determined the parameters of the log logistic model by making use of maximum likelihood estimate and bootstrap methods and observed the proximity of the results. A study conducted by [8], on the spread of HIV virus in San Francisco between 1978 and 1986 indicated that, the log logistic model was most suitable among other models to use with half censored data.

Under random or non-informative censoring, sample of say n elements are followed for a specified time say, T, the number of elements that is experiencing the event is considered to be random, but the entire length of study is fixed. Since the time is fixed, there are certain practical advantages with regards to designing a follow-up study. In a straightforward overview of this scheme, which is known as fixed time censoring, each element has a maximum inspection time say T_i, for $i = 1, \ldots, n$, which may possibly vary from one situation to another. $S(t)$represents the probability that a unit i will be alive at the end of the inspection time. Consider an experiment where we start with an observation of 50 cancer patients that have died or survived at the specified time. The survival of the patients may be due to withdrawal, inadequate monitoring mechanism or deaths which are not related to the purpose of the study.

Citation: Guure CB. Inference on the Loglogistic Model with Right Censored Data. Austin Biom and Biostat. 2015;2(1): 1015.

Maximum Likelihood Estimator (MLE) has been used frequently in determining the parameters of most of the lifetime distributions such as Weibull, lognormal, generalized exponential and others. Some of the works can be found in [11], they studied, generalized exponential distribution: Bayesian estimations. Other estimation procedures related to the above were considered by [12]. Determined the Bayes estimates of the reliability function and the hazard rate of the Weibull failure time distribution by employing squared error loss function [13]. Applied Bayesian to the parameter and reliability estimate of Weibull failure time distribution [14], studied the approximate Bayesian estimates for the Weibull reliability function and hazard rate from censored data by employing a new method that has the potential of reducing the number of terms in Lindley's approximation procedure. Others include; [15-20].

The main objective of this study is to apply the Bayesian estimator's procedure using Lindley's approximation method with two loss functions for the unknown parameters of the log logistic distribution against the classical maximum likelihood estimator with different sample sizes and parameter values using simulation study. Since both parameters of the distribution are non-negative, we assume that both take on the gamma prior distributions which are not necessarily the conjugate priors for the parameters.

Maximum Likelihood Estimation

Consider a set of n independently and identically distributed random pairs of (t_i, δ_i), where $t_i = min(X_i, T_i)$ and $\delta_i = I(X_i \leq T_i)$ indicating whether the observation is censored or not for $I = 1, 2, \ldots, n$. in an independent random censored model, it is assumed that the survival time X_i and the censoring time T_i are independent and from the same distribution. The score vectors are

$$h(\theta) = \frac{\partial Log\ L(\theta, p, t, \delta)}{\partial \theta} \quad and \quad h(p) = \frac{\partial Log\ L(\theta, p, t, \delta)}{\partial p}$$

where the score becomes a vector of the first partial derivatives of (θ, p). When using maximum likelihood to estimate unknown parameters that cannot be obtained in close form, one always requires that an iterative (eg, Newton-Raphson) procedure be implemented, such that, one can consider evaluating MLEs of α with a trial value say α_0 using a first order Taylor series as

$$h(\hat{\alpha}) \approx h(\alpha_0) + \frac{\partial h(\alpha)}{\partial \alpha}(\hat{\alpha} - \alpha_0) \quad (1)$$

Setting the left hand side of equation (1) to zero and solving for $\hat{\alpha}$, we have

$$\hat{\alpha} = \alpha_0 - H^{-1}(\alpha_0) h(\alpha_0) \quad (2)$$

where $H(\alpha_0)$ is the Hessian matrix and $h(\alpha_0)$ the score vector.

Considering the two parameters of the log logistic distribution, the Hessian matrix can be obtained as follows for the parameters estimates. The score vector of

$$h(\theta) = -\frac{p\sum_{i=1}^{n}\delta_i}{\theta} + \frac{p\sum_{i=1}^{n}\delta_i\left(\frac{t_i}{\theta}\right)^p}{\theta\sum_{i=1}^{n}(1+\left(\frac{t_i}{\theta}\right)^p)} \quad (3)$$

$$h(p) = \frac{\sum_{i=1}^{n}\delta_i}{p} - In(\theta)\sum_{i=1}^{n}\delta_i + \sum_{i=1}^{n}\delta_i In(t_i) - \frac{\sum_{i=1}^{n}\delta_i\left(\frac{t_i}{\theta}\right)^p In\left(\frac{t_i}{\theta}\right)}{\sum_{i=1}^{n}(1+\left(\frac{t_i}{\theta}\right)^p)} - \frac{\sum_{i=1}^{n}\left(\frac{t_i}{\theta}\right) In\left(\frac{t_i}{\theta}\right)}{\sum_{i=1}^{n}(1+\left(\frac{t_i}{\theta}\right)^p)} \quad (4)$$

From above, the partial derivatives for both θ and p is

$$\frac{\partial^2 \log L(\theta, p; t_i, \delta_i)}{\partial p\ \partial\theta} = \frac{\partial^2 \log L(\theta, p; t_i, \delta_i)}{\partial\theta\ \partial p} = \frac{\sum_{i=1}^{n}\delta_i}{\theta} + \frac{p\sum_{i=1}^{n}\delta_i\left(\frac{t_i}{\theta}\right)^p \sum_{i=1}^{n} In\left(\frac{t_i}{\theta}\right)^p}{\theta\sum_{i=1}^{n}(1+\left(\frac{t_i}{\theta}\right)^p)} + \frac{\sum_{i=1}^{n}\delta_i\left(\frac{t_i}{\theta}\right)^p}{\theta\sum_{i=1}^{n}(1+\left(\frac{t_i}{\theta}\right)^p)} - \frac{p\sum_{i=1}^{n}\left[\delta_i\left(\frac{t_i}{\theta}\right)^p\right]^2 \sum_{i=1}^{n} In\left(\frac{t_i}{\theta}\right)^p}{\theta\sum_{i=1}^{n}\left[1+\left(\frac{t_i}{\theta}\right)^p\right]^2} + \frac{p\sum_{i=1}^{n}\left(\frac{t_i}{\theta}\right)^p \sum_{i=1}^{n} In\left(\frac{t_i}{\theta}\right)^p}{\theta\sum_{i=1}^{n}(1+\left(\frac{t_i}{\theta}\right)^p)} + \frac{\sum_{i=1}^{n}\left(\frac{t_i}{\theta}\right)^p}{\theta\sum_{i=1}^{n}(1+\left(\frac{t_i}{\theta}\right)^p)} \quad (5)$$

Where $\frac{\partial h(p)}{\partial p}$ and $\frac{\partial h(\theta)}{\partial\theta}$ are easy to obtain. Equations (3), (4) and (5) can be substituted into equation (2), from which an iterative procedure could be implemented to obtain the parameter estimates under maximum likelihood.

Bayesian Inference of the Unknown Parameters

In this section, we consider Bayesian inference of the unknown parameters of the log logistic distribution. In order to employ the Bayesian methods, a prior needs to be defined. A prior is simply one's knowledge or an expert's opinion on the parameters being estimated. We have little prior information for all the parameters being estimated, and so we want our data information to dominate the prior distribution by assuming reasonably non-informative priors for all the parameters in this model. It is assumed that the two parameters follow a vague Gamma (a, b) and Gamma (c, d) prior distributions. These prior models are chosen because both the scale and shape parameters of the log logistic distribution are non-negative.

$$\Pi_1(\theta)\alpha\theta^{a-1}\exp(-\theta b), \theta > 0 \quad (6)$$

$$\Pi_2(p)\alpha p^{c-1}\exp(-pd), \theta > 0 \quad (7)$$

The Bayesian posterior distribution based on which inferences are drawn is

$$\pi^*(\theta, p|t)\alpha\frac{L(data|\theta, p)\pi_1(\theta)\pi_2(p)}{\int\int L(data|\theta, p)\pi_1(\theta)\pi_2(p)d\theta dp} \quad (8)$$

Squared-error Loss

The squared error loss is the loss incurred by adapting action say, $\hat{a}$ when the true value is say, a.

In other words, it implies the cost obtained by replacing the actual value of the parameter with the parameter estimate. Let the Bayesian estimator say, β_{se} be the posterior mean. If u (θ, p) is considered as the function of interest, then:

$$\beta_{se} = E\{u(\theta,p)|t,\delta\} = \frac{\int\int u(\theta,p)L(data|\theta,p)\pi_1(\theta)\pi_2(p)d\theta dp}{\int\int L(data|\theta,p)\pi_1(\theta)\pi_2(p)d\theta dp} \quad (9)$$

Note; the function of interest in our study is the loss function which measures the distribution parameters of θ and p. It is observed that equation (9) cannot be computed explicitly even if we take some specific priors on the parameters, as a result [21] proposed an approximation procedure to compute the ratio of two integrals similar to equation (9). The approximation procedure is adopted in this paper.

Lindley Approximation

The posterior Bayes estimator of an arbitrary function u(α) given by [21] is

$$E\{u(\alpha)|x\} = \frac{\int \omega(\alpha)\exp[l(\alpha)]d\alpha}{\int v(\alpha)\exp[l(\alpha)]d\alpha} \quad (10)$$

Where $l(\alpha)$ is the log-likelihood and $\omega(\alpha)$, $v(\alpha)$ are arbitrary functions of α. We assume that $v(\alpha)$ is the prior distribution for α and $\omega(\alpha)= u(\alpha).v(\alpha)$ with $u(\alpha)$ being some function of interest. The posterior expectation according to [12] is

$$E\{u(a)|t\} = \frac{\int u(a)\exp\{l(a)+\rho(\alpha)\}d\alpha}{\int \exp\{l(a)+\rho(\alpha)\}d\alpha} \quad (11)$$

Where $\rho(\alpha)=\log\{v(\alpha)\}$.

An asymptotic expansion of Lindley's approach of equation (11) according to [18] is

$$\hat{u} = u(\hat{\theta},\hat{p}) + \frac{1}{2}[(u_{11}\sigma_{11})+(u_{11}\sigma_{11})] + u_1\rho_1\sigma_{11} + u_2\rho_2\sigma_{22} + \frac{1}{2}[(l_{30}u_1\sigma_{11}^2)+(l_{03}u_2\sigma_{22}^2)] \quad (12)$$

where l stands for the log-likelihood function.

Considering the Bayesian estimator via Lindley, the following are obtained with u_1, u_{11} and u_2, u_{22} representing the first and second derivatives of θ and p respectively under the squared error loss which is referred to as the posterior mean.

$$u(\theta)=\theta,\ u_1 = 1,\ u_{11} = 0,\ u(p) = p,\ u_2 = 1,\ u_{22} = 0\ \rho = \text{In}\ \pi_1(\theta) + \text{In}\ \pi_2(p),$$

$$\rho_1 = \frac{a-1}{\theta} - b,\ \rho_2 = \frac{c-1}{\theta} - d\ \ \sigma_{11} = (-l_{20})^{-1}, \sigma_{22} = (-l_{02})^{-1}$$

Let l_{20} and l_{30} represent the second and third derivatives of the log-likelihood function with respect to the scale parameter θ, then

$$l_{20} = -\frac{p\sum_{i=1}^n \delta_i}{\theta^2} - \frac{p^2\sum_{i=1}^n \delta_i\left(\frac{t_1}{\theta}\right)^p}{\theta^2\sum_{i=1}^n\left[1+\left(\frac{t_1}{\theta}\right)^p\right]} - \frac{p\sum_{i=1}^n \delta_i\left(\frac{t_1}{\theta}\right)^p}{\theta^2\sum_{i=1}^n\left[1+\left(\frac{t_1}{\theta}\right)^p\right]} +$$

$$\frac{p^2\sum_{i=1}^n \delta_i\left[\left(\frac{t_i}{\theta}\right)^p\right]^2}{\theta^2\sum_{i=1}^n\left[1+\left(\frac{t_i}{\theta}\right)^p\right]^2} - \frac{p^2\sum_{i=1}^n \delta_i\left(\frac{t_i}{\theta}\right)^p}{\theta^2\sum_{i=1}^n\left[1+\left(\frac{t_i}{\theta}\right)^p\right]} -$$

$$\frac{p\sum_{i=1}^n \left(\frac{t_i}{\theta}\right)^p}{\theta^2\sum_{i=1}^n\left[1+\left(\frac{t_i}{\theta}\right)^p\right]} + \frac{p\left[\sum_{i=1}^n \left(\frac{t_i}{\theta}\right)^p\right]^2}{\theta^2\sum_{i=1}^n\left[1+\left(\frac{t_i}{\theta}\right)^p\right]^2}$$

$$l_{30} = -\frac{2p\sum_{i=1}^n \delta_i}{\theta^3} + \frac{p^3\sum_{i=1}^n \delta_i\left(\frac{t_i}{\theta}\right)^p}{\theta^3\sum_{i=1}^n\left[1+\left(\frac{t_i}{\theta}\right)^p\right]} +$$

$$\frac{3p^2\sum_{i=1}^n \delta_i\left(\frac{t_1}{\theta}\right)^p}{\theta^3\sum_{i=1}^n\left[1+\left(\frac{t_1}{\theta}\right)^p\right]} - \frac{3p^3\sum_{i=1}^n \delta_i\left[\left(\frac{t_1}{\theta}\right)^p\right]^2}{\theta^3\sum_{i=1}^n\left[1+\left(\frac{t_1}{\theta}\right)^p\right]^2} + \frac{2p\sum_{i=1}^n \delta_i\left(\frac{t_1}{\theta}\right)^p}{\theta^3\sum_{i=1}^n\left[1+\left(\frac{t_1}{\theta}\right)^p\right]} -$$

$$\frac{3p^2\sum_{i=1}^n \delta_i\left[\left(\frac{t_i}{\theta}\right)^p\right]^2}{\theta^3\sum_{i=1}^n\left[\left(1+\left(\frac{t_i}{\theta}\right)^p\right)\right]^2} + \frac{2p^9\sum_{i=1}^n \delta_i\left[\left(\frac{t_i}{\theta}\right)^p\right]^3}{\theta^3\sum_{i=1}^n\left[\left(1+\left(\frac{t_i}{\theta}\right)^p\right)\right]^3} +$$

$$\frac{p^3\sum_{i=1}^n \left(\frac{t_i}{\theta}\right)^p}{\theta^3\sum_{i=1}^n\left[\left(1+\left(\frac{t_i}{\theta}\right)^p\right)\right]} + \frac{3p^2\sum_{i=1}^n \left(\frac{t_i}{\theta}\right)^p}{\theta^3\sum_{i=1}^n\left[\left(1+\left(\frac{t_i}{\theta}\right)^p\right)\right]} - \frac{3p^3\sum_{i=1}^n \left[\left(\frac{t_i}{\theta}\right)^p\right]^2}{\theta^3\sum_{i=1}^n\left[\left(1+\left(\frac{t_i}{\theta}\right)^p\right)\right]^2} +$$

$$\frac{2p\sum_{i=1}^n \left(\frac{t_i}{\theta}\right)^p}{\theta^3\sum_{i=1}^n\left[1+\left(\frac{t_i}{\theta}\right)^p\right]} - \frac{3p^2\sum_{i=1}^n \delta_i\left[\left(\frac{t_i}{\theta}\right)^p\right]^2}{\theta^3\sum_{i=1}^n\left[\left(1+\left(\frac{t_i}{\theta}\right)^p\right)\right]^2} + \frac{2p^3\sum_{i=1}^n \delta_i\left[\left(\frac{t_i}{\theta}\right)^p\right]^3}{\theta^3\sum_{i=1}^n\left[\left(1+\left(\frac{t_i}{\theta}\right)^p\right)\right]^3}$$

If we let l_{02} and l_{03} represent the second and third derivatives of the log-likelihood function with respect to the shape parameter p, we will have

$$l_{02} = -\frac{\sum_{i=1}^n \delta_i}{p^2} - \frac{\sum_{i=1}^n \delta_i\left(\frac{t_i}{\theta}\right)^p \text{In}\left(\frac{t_i}{\theta}\right)^2}{\sum_{i=1}^n\left[1+\left(\frac{t_i}{\theta}\right)^p\right]} + \frac{\sum_{i=1}^n \delta_i\left[\left(\frac{t_i}{\theta}\right)^p\right]^2 \text{In}\left(\frac{t_i}{\theta}\right)^2}{\sum_{i=1}^n\left[1+\left(\frac{t_i}{\theta}\right)^p\right]^2} -$$

$$\frac{\sum_{i=1}^n \left(\frac{t_i}{\theta}\right)^p \text{In}\left(\frac{t_i}{\theta}\right)^2}{\sum_{i=1}^n\left[1+\left(\frac{t_i}{\theta}\right)^p\right]} + \frac{\left[\sum_{i=1}^n \left(\frac{t_i}{\theta}\right)^p\right]^2 \text{In}\left(\frac{t_i}{\theta}\right)^2}{\sum_{i=1}^n\left[1+\left(\frac{t_i}{\theta}\right)^p\right]^2}$$

$$l_{03} = \frac{2\sum_{i=1}^n \delta_i}{p^3} - \frac{\sum_{i=1}^n \delta_i\left(\frac{t_i}{\theta}\right)^p \text{In}\left(\frac{t_i}{\theta}\right)^3}{\sum_{i=1}^n\left[1+\left(\frac{t_i}{\theta}\right)^p\right]} + \frac{3\sum_{i=1}^n \delta_i\left[\left(\frac{t_i}{\theta}\right)^p\right]^2 \text{In}\left(\frac{t_i}{\theta}\right)^2}{\sum_{i=1}^n\left[1+\left(\frac{t_i}{\theta}\right)^p\right]^2} -$$

$$\frac{2\sum_{i=1}^n \delta_i\left[\left(\frac{t_i}{\theta}\right)^p\right]^3 \text{In}\left(\frac{t_i}{\theta}\right)^3}{\sum_{i=1}^n\left[1+\left(\frac{t_i}{\theta}\right)^p\right]^3} - \frac{\sum_{i=1}^n \delta_i\left(\frac{t_i}{\theta}\right)^p \text{In}\left(\frac{t_i}{\theta}\right)^3}{\sum_{i=1}^n\left[1+\left(\frac{t_i}{\theta}\right)^p\right]} +$$

$$\frac{3\left[\sum_{i=1}^n \left(\frac{t_i}{\theta}\right)^p\right]^2 \text{In}\left(\frac{t_i}{\theta}\right)^3}{\sum_{i=1}^n\left[1+\left(\frac{t_i}{\theta}\right)^p\right]^2} - \frac{2\left[\sum_{i=1}^n \left(\frac{t_i}{\theta}\right)^p\right]^3 \text{In}\left(\frac{t_i}{\theta}\right)^3}{\sum_{i=1}^n\left[1+\left(\frac{t_i}{\theta}\right)^p\right]^3}$$

Linear exponential loss function

This loss function measures the degree of overestimation and underestimation of the parameters being examined. Let k represent the shape parameter of the LINEX loss function. Refer to [13] for the posterior expectation of the LINEX loss function. The Bayes estimator $\widehat{u}_{BL}$ of a function u=u[*exp*(-*kθ*),exp(-*kp*)] under LINEX is given as

$$\widehat{u_{BL}} = \frac{\int\int u\pi_1(\theta)\pi_2(p)L(\delta,t_i;\theta,p)d\theta dp}{\int\int \pi_1(\theta)\pi_2(p)L(\delta,t_i;\theta,p)d\theta dp} \quad (13)$$

With Lindley's approach, u_1, u_{11} and u_2, u_{22} are the first and second derivatives for θ and p respectively under the linear exponential loss function, hence

$$u(\theta) = \exp(-k\theta), u_1 = \frac{\partial u}{\partial \theta} = -k\exp(-k\theta)$$

$$u_{11} = \frac{\partial^2 u}{\partial \theta^2} = k^2\exp(-k\theta),\ u_2 = u_{22=0}$$

$$u(p) = \exp(-kp),\ u_2 = \frac{\partial u}{\partial p} = -k\exp(-kp)$$

$$u_{22} = \frac{\partial^2 u}{\partial p^2} = k^2\exp(-kp),\ u_1 = u_{11} = 0$$

Real Data Analysis

Example 1

The data for this example are from survival of patients with cervical cancer recruited to a randomised clinical trial that was aimed at analysing the effect of an addition of a radio sensitizer to radiotherapy (New therapy- "treatment B") compared to using radiotherapy alone (Control - "treatment A"). Treatment A and B were given to 16 and 14 patients respectively. The data are in days since the start of the study, the event of interest was death caused by this cancer. Our interest is on patients under treatment A to illustrate the proposed methods in this paper. The data is obtained from [22], and asterisked observations are censored.

Using the iterative procedure suggested in this paper and basing on comparison criterion on standard errors as well as their average confidence/credible lengths, we have for the MLEs of $\hat{\theta}$ and $\hat{p}$ to be 770.5429 and 1.90488 with their corresponding standard errors as 48.15893 and 0.11906 respectively. Since we do not have any prior information on the hyper-parameters, we assume a = b = c = d = 0.0001. The Bayes estimators under squared error loss for $\hat{\theta}$ and $\hat{p}$ have respectively the following parameters estimates and standard errors, 770.5429, 1.90206 and 48.15893, 0.11888.

Computing the Bayes estimates of $\hat{\theta}$ and $\hat{p}$ and that of their standard errors via the linear exponential loss function with a loss parameter of 0:7 we have, 859.7094, 1.78586 and 53.73182, 0.11162. With the loss parameter of 0:7, we have, 909.4092, 1.82677 and 56.83807, 0.11417 respectively.

What has been observed here is, both the maximum likelihood and Bayes under squared error loss function have the same scale parameter estimates and standard errors which are smaller than that of Bayes under the linear exponential loss function. For the shape parameter, Bayes under LINEX loss function with the loss parameter of 0.7 has the smallest standard error. This implies that overestimation is more serious than underestimation.

Considering a 95% confidence interval under MLE, we have $\hat{\theta}$ = (679.1514,864.9344) and that of $\hat{p}$ = (1.67153, 2.13823). The Bayesian credible intervals via the squared error loss function for $\hat{\theta}$ and $\hat{p}$ are (679.1514, 864.9344) and (1.66906, 2.13506) respectively. The Bayes credible intervals with respect to the linear exponential loss function with a loss parameter of 0.7 for $\hat{\theta}$ and $\hat{p}$ are (754.3950, 965.2380) and (1.56709, 2.00463) and that of the 0:7 are (798.0065, 1020.8120) and (1.60299, 2.05055) respectively.

Observing from above, LINEX loss function with a positive loss parameter had narrower credible intervals as compared to squared error loss function and maximum likelihood for the shape parameter. For the scale parameter, maximum likelihood's confidence interval and Bayes credible interval with squared error loss were narrower than Bayes using LINEX.

Example 2

In this example, we analyse another data set which is considered moderate to obtain the parameter estimates and their standard errors in order to compare the methods employed in this paper. The data shown in Table 6, Example 2, are obtained from [22] and refer to remission times, in weeks, for a group of 30 patients with leukaemia who received similar treatment. Asterisks denote censoring times.

From Table 1, the Bayes estimator under squared error loss for the scale parameter of the log logistic model has the same estimate and standard error as compared to that of the classical maximum likelihood estimator but with the shape parameter (*p*), Bayes under LINEX loss function has a smaller standard error as compared to the others but underestimates it. With the confidence/credible intervals as shown in Table 1, Bayes using LINEX loss function with 0:7 as the loss parameter has the narrowest credible interval for the shape parameter whiles for the scale parameter both Bayes using squared error and maximum likelihood had the narrowest.

Example 3

The data in Table 6, which is considered large, are obtained from [22]. The data represent survival times for 121 breast cancer patients who were treated over the period 1929-1938. Times are in months and asterisks denote censoring times.

With this sample as depicted in Table 1, both maximum likelihood and Bayes under squared error had almost the same parameter estimates as well as standard errors for both the scale and shape parameters of the log logistic survival model. Both standard errors for the shape parameter are better than Bayes using the linear exponential loss function but for the scale parameter Bayes had the smallest standard error. Bayes had the narrowest credible intervals than maximum likelihood for the shape parameter via LINEX and that of Bayes under squared error loss function and maximum likelihood for the scale parameter.

Simulation Study

Due to the difficulties in comparing the performances of the methods theoretical, extensive simulations are performed and the proposed estimators compared via standard errors and absolute errors. A sample size of n = 25, 50 and 100, are considered and the steps below followed to generate the data.

Table 1: Standard Errors and Confidence/ Credible Intervals for $\widehat{\theta}$ *and* $\widehat{p}$.

n = 30		MLE	BS	BL k = 0.7	BL k = -0.7
	$\widehat{\theta}$	17.45724	17.45724	17.83051	20.37725
	s.e $(\widehat{\theta})$	0.58191	0.58191	0.59435	0.67924
	CI $(\widehat{\theta})$	(16.3617, 18.5978)	(16.3617, 18.5978)	(16.6656, 18.9954)	(19.0459,21.7086)
	$\widehat{p}$	1.37594	1.37316	1.45007	1.27455
	s.e $(\widehat{p})$	0.04586	0.04577	0.04834	0.04248
	CI $(\widehat{p})$	(1.2861, 1.4658)	(1.2835, 1.4629)	(1.3553, 1.5448)	(1.19123, 1.3578)
n=121					
	$\widehat{\theta}$	60.49719	60.49719	59.57932	58.96793
	s.e $(\widehat{\theta})$	0.4500	0.4500	0.49239	0.44601
	CI $(\widehat{\theta})$	(59.5172, 61.4771)	(59.5172, 61.4771)	(58.6142, 60.5444)	(58.01275, 59.92311)
	$\widehat{p}$	1.60268	1.60256	1.77885	1.76418
	s.e $(\widehat{p})$	0.01325	0.01324	0.01470	0.01458
	CI $(\widehat{p})$	(1.57672, 1.62864)	(1.57660, 1.62851)	(1.75004, 1.80767)	(1.73560, 1.79276)

ML: Maximum Likelihood; BL: LINEX; BS: Squared Error; CI: Confidence/Credible Interval; s.e: Standard Error

We generated the survival time X from the sample sizes shown above from the log logistic model. The values of the assumed actual shape parameter (p) of the log logistic distribution were taken to be, 0.8, 1.8 and 2.8. The scale parameter *(θ)* was considered through out to be 1 without loss of generality.

The same sample size was generated from the Uniform distribution for the censored time T with (0; b), where the value of b depends solely on the proportion of the observations that are censored. In this study, the percentage of censoring was considered to be 20. t = min(X; T) is taken as the minimum of the failure time and that of the censored time.

To compute the Bayes estimates, an assumption was made such that _ and p take respectively Gamma (a, b) and Gamma (c, d) priors. The hyper-parameters were set to 0.0001, i.e. a = b = c = d = 0.0001, as suggested by [20].

For the purpose of illustration, the linear exponential loss parameter was taken to be k= ±0.7.The following were considered in choosing the parameter value. The sign of the loss parameter k represents the direction and its magnitude. If k = 0.7, the LINEX loss function will be asymmetric about zero with overestimation more costly than underestimation. If k < 0, the loss function rises exponentially but linearly with k > 0. Note also that if the parameter equal zero, the LINEX loss function turn to be approximately the squared error loss and therefore symmetric. See [23], for detailed discussions on the loss functions.

The objective of this study is to obtain the parameter estimates and to compare the methods proposed in this study. To examine the estimates of the parameters, the absolute errors and standard errors of the estimates are obtained and presented below.

When we consider Table 2, which contains the standard errors and the absolute errors of the estimated shape parameter $(\widehat{p})$, it was noticed that LINEX loss function has the smallest standard errors and absolute errors as compared to the others. It is observed that, LINEX with a positive loss parameter overall performed better than the others which is an indication of overestimation. As the sample size increases, Bayes using squared error and maximum likelihood performed equally better. Though, Bayes estimator using LINEX loss function seems to have the smallest standard errors and minimum absolute errors, it must be stated that, all the estimators standard errors and minimum absolute errors are close.

From Table 3, Bayesian using the linear exponential loss function (LINEX) again had the smallest standard errors and minimal absolute errors for the scale parameter $\widehat{\theta}$. In almost all the cases, the standard errors and the absolute errors of the two estimators, i.e., maximum likelihood and Bayes under squared error loss function turn to have the same values. This may be expected, in that, the priors used for the Bayesian analysis are non-informative. Again as the sample size increases, there is a corresponding decrease in standard errors for all the estimators with respect to the two parameters.

Extension

We have discussed Bayesian inference of the two-parameter log logistic distribution, but our method can be extended for many other cases also. We briefly describe an extension of our methods to Weibull distribution that was introduced by Waloddi Weibull in 1939 and has the probability density function and survival function as

$$f(x;\theta,p) = p/\theta(x/\theta)^{p-1}\exp[-(x/\theta)]^{p} \qquad (14)$$

Table 2: Average parameter estimates, standard errors and absolute errors for the shape parameter $(\hat{p})$.

n		p =0.8 $\hat{p}\ (\hat{p}_{se}, \hat{p}_{ae})$	p = 1.8 $\hat{p}\ (\hat{p}_{se}, \hat{p}_{ae})$	p= 2.8 $\hat{p}\ (\hat{p}_{se}, \hat{p}_{ae})$
25	ML	0.77618 (0.03068, 0.13305)	1.61270 (0.03325, 0.28271)	2.76125 (0.03846, 0.42696)
	BS	0.79193 (0.03063, 0.12370)	1.61091 (0.04001, 0.29231)	2.76123 (0.03846, 0.42703)
	BL(k = 0.7)	0.78761 (0.03059, 0.12110)	1.53437 (0.03387, 0.27571)	2.45155 (0.03770, 0.40630)
	BL (k = -0.7)	0.70829 (0.03057, 0.12067)	1.56760 (0.03351, 0.28183)	2.99242 (0.03784, 0.43115)
50	ML	0.61674 (0.03028, 0.10956)	1.37336 (0.03161, 0.23066)	2.16261 (0.03343, 0.33746)
	BS	0.62393 (0.03027, 0.10600)	1.37174 (0.03063, 0.23174)	2.16256 (0.03343, 0.33748)
	BL(k = 0.7)	0.62240 (0.03030, 0.10629)	1.68089 (0.03145, 0.22913)	2.26032 (0.03372, 0.32222)
	BL (k = -0.7)	0.76862 (0.03030, 0.10526)	1.64116 (0.03136, 0.22765)	2.63234 (0.03341, 0.33057)
100	ML	0.69397 (0.03013, 0.09996)	1.47272 (0.03070, 0.21096)	2.75933 (0.03169, 0.28657)
	BS	0.69512 (0.03013, 0.09865)	1.47244 (0.03070, 0.21118)	2.75933 (0.03169, 0.28657)
	BL(k = 0.7)	0.69465 (0.03014, 0.09982)	1.46329 (0.03068, 0.20685)	2.39198 (0.03170, 0.28763)
	BL (k = -0.7)	0.62615 (0.03013, 0.09964)	1.37476 (0.03072, 0.21498)	2.41555 (0.03171, 0.28681)

ML: Maximum Likelihood; BL: LINEX; BS: Squared Error; CI: Confidence/Credible Interval; s.e: Standard Error

Table 3: Average parameter estimates, standard errors and absolute errors for the shape parameter $(\hat{\theta})$.

n	e	p = 0.8 $\hat{\theta}\ (\hat{\theta}_{se}, \hat{\theta}_{ae})$	p = 1.8 $\hat{\theta}\ (\hat{\theta}_{se}, \hat{\theta}_{ae})$	p= 2.8 $\hat{\theta}\ (\hat{\theta}_{se}, \hat{\theta}_{ae})$
25	ML	2.66381 (0.40940, 0.11132)	1.78798 (0.04604, 0.50727)	0.92318 (0.00151, 0.24499)
	BS	2.66381 (0.40940, 0.11132)	1.78798 (0.04604, 0.50727)	0.92318 (0.00151, 0.24499)
	BL(k = 0.7)	2.66381 (0.35546, 0.11120)	1.54740 (0.04606, 0.48758)	0.96129 (0.00146, 0.23054)
	BL (k = -0.7)	2.74731 (0.39465, 0.11130)	2.07995 (0.04362, 0.50721)	1.07539 (0.00157, 0.24836)
50	ML	2.02794 (0.10703, 0.03951)	1.35288 (0.01514, 0.45403)	1.14866 (0.00072, 0.22216)
	BS	2.02794 (0.10703, 0.03951)	1.35288 (0.01514, 0.45403)	1.14866 (0.00072, 0.22216)
	BL(k = 0.7)	2.02779 (0.10098, 0.03761)	1.44093 (0.01276, 0.44474)	1.34871 (0.00072, 0.21737)
	BL (k = -0.7)	2.16347 (0.12558, 0.03981)	1.58097 (0.01346, 0.45824)	1.24848 (0.00072, 0.21009)
100	ML	2.58336 (0.04295, 0.01637)	1.61304 (0.00457, 0.41951)	1.15175 (0.00033, 0.20293)
	BS	2.58336 (0.04295, 0.01637)	1.61304 (0.00457, 0.41951)	1.15175 (0.00033, 0.20293)
	BL(k = 0.7)	2.58336 (0.04386, 0.01679)	1.39433 (0.00437, 0.42054)	1.19133 (0.00034, 0.19923)
	BL (k = -0.7)	2.58959 (0.03540, 0.01719)	1.69649 (0.00509, 0.42749)	1.13503 (0.00034, 0.20220)

ML = Maximum Likelihood, BL = LINEX, BS= Squared Error, CI = Confidence/Credible Interval, s.e= Standard Error

$$S(x;\theta,p)=\exp[-(x/\theta)]^{p} \qquad (15)$$

Note that the likelihood function can be obtained from equation (1). Here p and θ are the shape and scale parameters of the Weibull distribution. The log-likelihood function of the Weibull distribution from which maximum likelihood and Bayes estimates are obtained is

$$l=\sum_{i=1}^{n}\left[\delta_i[\text{In}(p)-p\ \text{In}(\theta)+(p-1)\text{In}(x_i)]-\left(\frac{x_i}{\theta}\right)^{p}\right] \qquad (16)$$

To compute the Bayes estimates of the unknown parameters, {θ, p}, it is assumed that p and θ have the same priors as described in equations (8) and (9) respectively. Based on the observed sample of $\{x_1,\ldots,x_n\}$, the posterior density function with respect to p,θ and the data is

$$\pi^{*}(\theta,p|x)\alpha\frac{L(data|\theta,p)\pi_1(\theta)\pi_2(p)}{\int_0^{\infty}\int_0^{\infty}L(data|\theta,p)\pi_1(\theta)\pi_2(p)d\theta dp} \qquad (17)$$

Note that equations (17) and (8) are of the same form except that t is replaced in equation (17) by x. Therefore, the methods used to obtain the posterior estimates with respect to equation (8), can be applied here also. It implies that for the squared error loss and linear exponential loss functions, we can make use of equations (9) to obtain the parameter estimates.

For the purpose of illustrations, we have analysed the data in Example 2 and have also obtained mean squared errors for both the scale and shape parameters via simulation which enable us to compare the methods proposed in this paper under the Weibull distribution. Standard errors were obtained for each parameter of the Weibull distribution parameters for the purpose of comparison. The standard errors are presented in Table 4 for the real data, it is noticed that Bayesian estimation method is more robust and give better estimates with corresponding smaller errors for both parameters than the traditional maximum likelihood method under the LINEX loss function. It is also clear from Table 5, that the Bayesian method

Table 4: Standard errors (se) for $\hat{\theta}$ and $\hat{\theta}$ with n =30 for the Weibull distribution.

	MLE	BS	BL k = 0.7	BL k = -0.7
$\hat{\theta}$	32.81261	32.85129	30.49252	32.06486
s.e $\hat{\theta}$	1.093754	1.095043	0.976417	1.068829
$\hat{\theta}$	1.085563	1.085555	0.972909	1.084810
s.e $\hat{p}$	0.036185	0.036185	0.032454	0.036127

Table 5: Mean squared errors for the estimated scale $(\hat{\theta})$ and shape $\hat{\theta}$ parameters.

n	θ	p	$\hat{\theta}_{ML}$	$\hat{\theta}_{BS}$	$\hat{\theta}_{BL}$ k = 1	$\hat{\theta}_{BL}$ k = -1	$\hat{\theta}_{BL}$ k = 2	$\hat{\theta}_{BL}$ k = 2
25	0.5	0.8	0.1619	0.164	0.1572	0.1414	0.1659	0.1608
		1.2	0.0534	0.0539	0.0573	0.0625	0.0576	0.058
	1.5	0.8	1.5896	1.6096	1.4967	1.4848	1.5067	1.4325
		1.2	0.5157	0.5208	0.5412	0.5359	0.5078	0.5262
50	0.5	0.8	0.1116	0.112	0.1127	0.1119	0.1115	0.1056
		1.2	0.0448	0.0449	0.0442	0.0435	0.0436	0.0437
	1.5	0.8	1.0158	1.0194	0.9678	0.9721	0.9342	0.9624
		1.2	0.3992	0.4003	0.3891	0.3861	0.3791	0.3879
100	0.5	0.8	0.0859	0.086	0.0844	0.0826	0.0826	0.0857
		1.2	0.0362	0.0363	0.0357	0.0358	0.0349	0.0355
	1.5	0.8	0.7818	0.7826	0.7608	0.7435	0.7524	0.7397
		1.2	0.3317	0.3319	0.3266	0.3109	0.3265	0.3171
n	θ	p	$\hat{p}_{ML}$	$\hat{p}_{BS}$	$\hat{p}_{BL}$ k = 1	$\hat{p}_{BL}$ k = -1	$\hat{p}_{BL}$ k = 2	$\hat{p}_{BL}$ k = 2
25	0.5	0.8	0.0242	0.0242	0.0213	0.0226	0.0216	0.0226
		1.2	0.0527	0.0527	0.0489	0.0485	0.0484	0.0497
	1.5	0.8	0.0256	0.0256	0.0224	0.0216	0.0211	0.0206
		1.2	0.0489	0.0489	0.0489	0.0457	0.0501	0.0497
50	0.5	0.8	0.0095	0.0095	0.0095	0.0088	0.0093	0.0096
		1.2	0.0199	0.0199	0.0221	0.0199	0.0211	0.0191
	1.5	0.8	0.0095	0.0095	0.0088	0.0092	0.0093	0.0091
		1.2	0.0199	0.0199	0.0203	0.0202	0.0197	0.0208
100	0.5	0.8	0.0044	0.0044	0.0046	0.0041	0.0041	0.0038
		1.2	0.0097	0.0097	0.0094	0.0091	0.0092	0.0092
	1.5	0.8	0.004	0.004	0.0043	0.0043	0.0043	0.0041
		1.2	0.0093	0.0093	0.0094	0.0101	0.0093	0.0087

has the smallest mean squared errors especially with the linear exponential loss function as compared to the MLE method. Note that the true parameters assumed for the Weibull model are 0.5 and 1.5 for the scale parameter and 0.8 and 1.2 for the shape parameter. We have also considered the LINEX loss parameter to be ± 1and ±2 without loss of generality.

Conclusion

In this paper, we have addressed the problem of Bayesian estimation for the loglogistic survival model, under linear exponential and squared error loss functions and that of maximum likelihood estimation. Bayes estimators are obtained using Lindley approximation whiles MLE are obtained using Newton-Raphson. A simulation study was conducted to examine and compare the performance of the estimates for different sample sizes with different values for the loss parameter. Three real dataset are also analysed.

From the above discussions in relation to both the real data analysis and the simulation study, we can conclude that all the estimators are presumably good for estimating the scale and shape parameters of the log logistic survival model, since the standard errors and the absolute errors of the estimates under the simulation study are

Table 6: Real Data for examples 1 to 3.

Example 1	90, 890*, 142, 1037, 150, 1090*, 269, 1113*, 291, 1153, 468*, 1297, 680, 1429, 837, 1577*
Example 2	1, 1, 2, 4, 4, 6, 6, 6, 7, 8, 9, 9, 10, 12, 13, 14, 18, 19, 24, 26, 29, 31*, 42, 45*, 50*, 57, 60, 71*, 85*, 91
Example 3	0.3, 0.3*, 4.0*, 5.0, 5.6, 6.2, 6.3, 6.6, 6.8, 7.4*, 7.5, 8.4, 8.4, 10.3, 11.0, 11.8, 12.2, 12.3, 13.5, 14.4, 14.4, 14.8, 15.5*, 15.7,16.2, 16.3, 16.5, 16.8, 17.2, 17.3, 17.5, 17.9, 19.8, 20.4, 20.9, 90, 93*,96*, 103*, 105*,109*, 109*, 111*, 115*, 117*, 125*, 126, 127*, 129*, 129*, 139*, 154*54,55*, 56, 57*, 58*, 59*, 60, 60*, 60*, 61*, 62*, 65*, 65*, 67*, 67*,68*, 69*, 78,80, 83*, 88*, 89,39*, 40, 40*, 40*, 41, 41, 41*, 42, 43*, 43*, 43*, 44, 45*, 45*, 46*,46*, 47*, 48, 49*, 51, 51, 51*, 52,23.0, 23.4*, 23.6, 24.0, 24.0, 27.9, 28.2,29.1, 30, 31, 31, 32, 35, 35, 37*, 37*, 37*, 38, 38*, 38*, 39*,21.0, 21.0, 21.1,

close. Maximum likelihood estimator and Bayes using squared error loss with respect to both the simulation and the real data analysis had the same estimate for the scale parameter of the distribution. With respect to the shape parameter, Bayes using squared error loss performed better than maximum likelihood but as the sample size increased regarding the simulation study and the real data application both had their standard errors converging to the same values. The LINEX loss function had the smallest standard error with the fairly small and moderate samples for the shape parameter but when the data became large, maximum likelihood and Bayes using the squared error loss also performed better.

Based on the results and discussions given above, we agree with [9], whose study suggests that maximum likelihood estimation is a suitable method in estimating the parameters when performing analyses using log logistic distribution on grouped data such as half censored data but for Weibull distribution the Bayesian approach outperform that of the MLE. The Bayesian approach adopted in this paper is a good alternative to MLE but is computational intensive as compared to maximum likelihood method which is simple to implement.

Acknowledgement

The authors are grateful to the reviewers and the editor-in-chief for their comments which helped improve the paper significantly.

References

1. Bennett S. Log-Logistic Regression Models for Survival Data", Applied Statistics, Journal of the Royal Statistical Society. Series C (Appl Stat). 1983; 32: 165-171.
2. Cox DR, Oakes D. Analysis of survival data. Chapman and Hall, New York. 1984.
3. Cox DR. The analysis of binary data. Methenn & Co., London. 1970.
4. Mann NR, Schafer, RE Singpurwalla ND. Methods for Statistical Analysis of Reliability and Life Data. New York. 1974.
5. Lawless JF. Statistical Models and Methods for Lifetime Data, Wiley, New York. 2003.
6. Gupta RD, Kundu D. Generalized Exponential Distributions, Aus Nz J Stat. 1999; 41: 173-188.
7. Gupta RD, Akman O, Lvin SA. Study of Log-Logistic Model in Survival Analysis. BiometricalJ. 1999; 41: 431-443.
8. Byers RH, Morgan WM, Darrow WW, Doll L, Jaffe HW, Rutherford G, Hessol N. Estimating AIDS infection rates in the San Francisco cohort. AIDS. 1988; 2: 207-210.
9. Zhou YY, Mi J, Guo S. Estimation of parameters in logistic and log-logistic distribution with grouped data. Lifetime Data Anal. 2007; 13: 421-429.
10. Elvan AH, Asli S, Burak U, Omer D, Mehmet NO, Gui KMD. Comparison of Five Survival Models: Breast Cancer Registry Data from Ege University Cancer Research Centre, Turkiye Klinikleri Journal of Medical Sciences. 2010; 30: 1665-1674.
11. Kundu D, Gupta R. Generalized exponential distribution: Bayesian estimations. Comput Stat Data Analysis. 2008; 52: 1873-1883.
12. Sinha SK. Bayes Estimation of the Reliability Function and Hazard Rate of a Weibull Failure Time Distribution, Tranbajos De Estadistica. 1986; 1: 47-56.
13. Guure CB, Ibrahim NA, Adam MB, Bosomprah S, Ahmed AOM. Bayesian Parameter and Reliability Estimate of Weibull Failure Time Distribution. Bull Malays Math Sc Soc. In Press.
14. Abdel-Wahid AA, Winterbottom A. Approximate Bayesian Estimates for the Weibull Reliability Function and Hazard Rate from Censored Data, J Stat Plan Infer. 1987; 16, 277 283.
15. Soliman AA, Abd Ellah AH, Ryan SB. Comparison of Estimates Using Record Statistics FromWeibull Model: Bayesian and Non-Bayesian Approaches. Comput Stat Data Analysis. 2006; 51: 2065-2077.
16. Vasile P, Eugeine P, Alina C. Bayes Estimators of Modified Weibull Distribution Parameters Using Lindley's Approximation, WSEAS Transactions on Mathematics. 2010; 7: 539-549.
17. Berger JO, Sun D. Bayesian analysis for the Poly-Weibull distribution, J Am Stat Assoc. 1993; 88: 1412-1418.
18. Guure CB, Ibrahim NA, Adam MB. Bayesian inference of the Weibull model based on interval-censored survival data. Comput Math Methods Med. 2013; 2013: 849520.
19. Lye LM, Hapuarachchi KP, Ryan S. Bayes Estimation of the Extreme-Value Reliability Function. IEEE Trans Reliab. 2010; 42: 641-644.
20. Press JS. The subjectivity of scientists and the Bayesian approach, New York; Wiley. 2001.
21. Lindley DV. Approximate Bayesian Methods, Trabajos Estadist. 1980; 31: 223-237.
22. Coolen FPA, Yan KJ. Nonparametric predictive inference with right-censored data J Stat Plan Infer. 2004; 126: 25-54.
23. Calabria R. and Pulcini G. Point Estimation under Asymmetric Loss Function for Left Truncated Exponential Samples, Commun Stat-Theor M. 1996; 25: 585-600.

Austin Biom and Biostat - Volume 2 Issue 1 - 2015
Submit your Manuscript | www.austinpublishinggroup.com

Citation: Guure CB. Inference on the Loglogistic Model with Right Censored Data. Austin Biom and Biostat. 2015;2(1): 1015.