ScienceAsia 39S (2013): **75–79**

RESEARCH ARTICLE

doi: 10.2306/scienceasia1513-1874.2013.39S.075

Generalized Bayesian non-informative prior estimation of Weibull parameter with interval censoring

Chris Bambey Guure[a,b,*], **Noor Akma Ibrahim**[b,c]

[a] Department of Biostatistics, School of Public Health, University of Ghana, Accra, Ghana
[b] Institute for Mathematical Research, Universiti Putra Malaysia, 43400 Selangor, Malaysia
[c] Department of Mathematics, Faculty of Science, Universiti Putra Malaysia, 43400 Selangor, Malaysia

*Corresponding author, e-mail: chris@chrisguure.com

Received 7 Jan 2013
Accepted 5 Apr 2013

ABSTRACT: Interval-censored data consist of adjacent inspection times that surround an unknown failure time. We seek to determine the best estimator for the Weibull scale parameter using interval-censored survival data. Consideration is given to the classical maximum likelihood and Bayesian estimation under squared error loss with interval censoring using non-informative prior and a proposed generalization of non-informative prior. The study is based on simulation and comparisons are made using mean squared error and absolute bias. We find that the proposed generalized non-informative prior is the preferred estimator of the scale parameter.

KEYWORDS: interval censored data, maximum likelihood, Bayesian inference, simulation study, squared error loss

INTRODUCTION

One distinctive part of failure-time data is censoring, and there is a lot of literature on right censoring[1–7]. Here our focus is on interval censoring, which is more challenging than right censoring, and as a result the approaches developed for right censoring do not generally apply.

Interval censoring has to do with a study subject or failure time processes of interest that is not under regular observation. With interval censoring one only knows a range, i.e, an interval, inside of which one can say the survival event has occurred. Left or right-censored failure times are special cases of interval-censored failure-time data. As stated in Ref. 8, one could define an interval-censored observation as a union of several non-overlapping windows or intervals.

According to Ref. 9, interval-censored failure time data occur in many areas including demographical, epidemiological, financial, medical, sociological, and engineering studies. A typical example of interval-censored data occurs in medical or health studies that entail periodic follow-ups, and many clinical trials and longitudinal studies fall into this category. In such situations, interval-censored data may arise in several ways. For instance, an individual may miss one or more observation times that have been scheduled to clinically observe possible changes in disease status and then return with a changed status[9].

When we consider individuals who visit clinical centres at times convenient to them rather than at predetermined observation time or in a situation where a mechanical system is under observation at some time schedule for which there is no control over the time. In these situations, the data that is obtained are interval censored. If all study subjects or units follow the predetermined observation schedule time exactly, it is still not possible to observe the exact time of the occurrence of the change even if we assume that it is a continuous variable. In the last situation, one has grouped failure time data, that is, interval-censored data for which the observation for each subject is a member of a collection of non-overlapping intervals.

Further discussions on interval censoring using the classical statistical approach can be obtained in Refs. 10–14. No one has examined interval-censored data using the Bayesian estimation approach with regards to Weibull distribution, which is the essence of this study.

Next in this paper we give the derivatives of the parameters under the maximum likelihood estimator followed by the Bayesian estimator. We also present the derivatives of the Bayesian technique using the Lindley approximation procedure after which is the simulation study followed by the results.

MAXIMUM LIKELIHOOD

Interval-censored data arise in a natural way when n items on test are not monitored but are inspected for the number of surviving items where two data values surround an unknown variable. Let $t_1, \dots, t_n$ be a random sample of size n where the probability density function is represented by $f(t, \alpha, \beta)$ and the cumulative distribution function is $F(t, \alpha, \beta)$. Then

$$f(t;\alpha,\beta) = \frac{\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1}\exp\left[-\left(\frac{t}{\alpha}\right)^{\beta}\right] \tag{1}$$

and

$$F(t;\alpha,\beta) = 1-\exp\left[-\left(\frac{t}{\alpha}\right)^{\beta}\right] \tag{2}$$

where β is the shape parameter and α the scale parameter. Let $[L_i, R_i]$ denote the interval-censored data and T represent the unknown time, i.e., $L_i \leqslant T_i \leqslant R_i$, where L_i is the last inspection time, R_i the state end time. If censoring occurs non-informatively and if the law governing L and R does not involve any of the parameters of interest, we can base our inferences on the likelihood function $L(L_i, R_i|\alpha, \beta)$ as stated in Ref. 15 which is given by

$$L(L_i, R_i|\alpha,\beta) = \prod_{i=1}^{n}[F(R_i,\alpha,\beta) - F(L_i,\alpha,\beta)]. \tag{3}$$

This implies that

$$\prod_{i=1}^{n}[S(L_i) - S(R_i)] = \prod_{i=1}^{n} P(L_i \leqslant T_i \leqslant R_i),$$

where P is the probability. From this we have

$$L(.) = \prod_{i=1}^{n}\left\{\exp\left[-\left(\frac{L_i}{\alpha}\right)^{\beta}\right] - \exp\left[-\left(\frac{R_i}{\alpha}\right)^{\beta}\right]\right\}. \tag{4}$$

Taking the log of (4) we have

$$\ell = \sum_{i=1}^{n}\ln\left\{\exp\left[-\left(\frac{L_i}{\alpha}\right)^{\beta}\right] - \exp\left[-\left(\frac{R_i}{\alpha}\right)^{\beta}\right]\right\}. \tag{5}$$

To find α and β that maximizes (5), differentiate (5) with respect to α and β and set the resulting equations to zero. The score vectors are

$$\frac{\partial \ell}{\partial \beta} = 0 \text{ and } \frac{\partial \ell}{\partial \alpha} = 0. \tag{6}$$

Hence

$$\frac{\partial \ell}{\partial \beta} = \sum_{i=1}^{n}\left[\left(-\frac{\left(\frac{L_i}{\alpha}\right)^{\beta}\ln\left(\frac{L_i}{\alpha}\right)\exp\left[-\left(\frac{L_i}{\alpha}\right)^{\beta}\right]}{\exp\left[-\left(\frac{L_i}{\alpha}\right)^{\beta}\right] - \exp\left[-\left(\frac{R_i}{\alpha}\right)^{\beta}\right]}\right) - \left(\frac{\left(\frac{R_i}{\alpha}\right)^{\beta}\ln\left(\frac{R_i}{\alpha}\right)\exp\left[-\left(\frac{R_i}{\alpha}\right)^{\beta}\right]}{\exp\left[-\left(\frac{L_i}{\alpha}\right)^{\beta}\right] - \exp\left[-\left(\frac{R_i}{\alpha}\right)^{\beta}\right]}\right)\right],$$

$$\frac{\partial \ell}{\partial \alpha} = \sum_{i=1}^{n}\left[\left(-\frac{\left(\frac{L_i}{\alpha}\right)^{\beta}\left(\frac{L_i}{\alpha}\right)\exp\left[-\left(\frac{L_i}{\alpha}\right)^{\beta}\right]}{\exp\left[-\left(\frac{L_i}{\alpha}\right)^{\beta}\right] - \exp\left[-\left(\frac{R_i}{\alpha}\right)^{\beta}\right]}\right) - \left(\frac{\left(\frac{R_i}{\alpha}\right)^{\beta}\left(\frac{R_i}{\alpha}\right)\exp\left[-\left(\frac{R_i}{\alpha}\right)^{\beta}\right]}{\exp\left[-\left(\frac{L_i}{\alpha}\right)^{\beta}\right] - \exp\left[-\left(\frac{R_i}{\alpha}\right)^{\beta}\right]}\right)\right].$$

The shape parameter β is assumed known. As a result, the scale parameter α can easily be obtained with simple substitutions.

BAYESIAN ESTIMATION

Bayesian estimation approach has received a lot of attention for analysing failure time data. It makes use of ones prior knowledge about the parameters and takes into consideration the data available. In a situation where the researcher has a previous knowledge or can obtain an information from experts that is closely related to the current study, then a suitable prior to use is the informative prior if not a non-informative prior will be an alternative to use by assuming lack of previous knowledge. The Bayesian estimation is considered under squared error loss function which is simply the posterior mean. We have in this study considered Jeffreys vague non-informative prior and have also proposed generalized non-informative prior since we are assuming a lack of information about the parameters. Given a sample $[L_i, R_i]$, the likelihood function L follows (4).

Non-informative prior

If we consider a likelihood function of $L(\theta)$, with its Fisher information $I(\theta) = -E(\partial^2 \log L(\theta)/\partial\theta^2)$. Jeffreys suggested that $\pi(\theta) \propto \det(I(\theta))^{1/2}$ be considered as a prior for the likelihood function $L(\theta)$. The Jeffreys prior is justified on the grounds of its invariance under parametrization[16]. With one parameter, the Jeffreys vague prior is[16]

$$u(\alpha) \propto \frac{1}{\alpha}. \tag{7}$$

Let the likelihood equation which is $L(L_i, R_i|\alpha, \beta)$ be the same as (4). The posterior distribution with respect to the Jeffreys prior is found by using the conditional distribution which depends on the joint probability density function and the marginal probability density function. The conditional distribution is given as

$$\pi(L_i, R_i; \alpha, \beta) = \frac{H(L_i, R_i; \alpha, \beta)}{P[L_i, R_i]} \tag{8}$$

where the joint probability density function is

$$H(L_i, R_i; \alpha, \beta) = L(L_i, R_i|\alpha, \beta)u(\alpha) \tag{9}$$

and the marginal distribution function is

$$P[L_i, R_i] = \int_0^\infty L(L_i, R_i|\alpha, \beta)u(\alpha)\,\mathrm{d}\alpha. \tag{10}$$

Hence the posterior probability density function of α given the data $[L_i, R_i]$ is obtained by dividing the joint posterior density function over the marginal distribution function as

$$\pi^*(\alpha, \beta|L_i, R_i) = \frac{L(L_i, R_i|\alpha, \beta)u(\alpha)}{\int_0^\infty L(L_i, R_i|\alpha, \beta)u(\alpha)\,\mathrm{d}\alpha}. \tag{11}$$

When we therefore consider the squared error loss function with the scale parameter, we have

$$\pi^*(.) = \frac{\int_0^\infty u(\alpha)\alpha \prod_{i=1}^{n} \mathrm{e}^{-(L_i/\alpha)^\beta} - \mathrm{e}^{-(R_i/\alpha)^\beta}\,\mathrm{d}\alpha}{\int_0^\infty u(\alpha) \prod_{i=1}^{n} \mathrm{e}^{-(L_i/\alpha)^\beta} - \mathrm{e}^{(R_i/\alpha)^\beta}\,\mathrm{d}\alpha}. \tag{12}$$

Generalized non-informative prior

We propose a generalized non-informative prior for the parameter α which is given as

$$v(\alpha) \propto \frac{1}{\alpha^k}, \qquad k > 0. \tag{13}$$

This is a generalization of Jeffreys vague prior. When $k = 1$, we have Jeffreys' vague non-informative prior. The posterior probability density function of α given the data L_i, R_i under squared error loss function is obtained by making use of (12) and simply substituting $u(\alpha)$ for $v(\alpha)$ as given below.

$$\pi^*(.) = \frac{\int_0^\infty v(\alpha)\alpha \prod_{i=1}^{n} \mathrm{e}^{-(L_i/\alpha)^\beta} - \mathrm{e}^{-(R_i/\alpha)^\beta}\,\mathrm{d}\alpha}{\int_0^\infty v(\alpha) \prod_{i=1}^{n} \mathrm{e}^{-(L_i/\alpha)^\beta} - \mathrm{e}^{(R_i/\alpha)^\beta}\,\mathrm{d}\alpha}. \tag{14}$$

It may be noted that (12) and (14) do not simplify to nice closed forms. This is due to the complex form of the likelihood function given in (4). We therefore propose to use the Lindley approximation method to evaluate the integrals involved.

LINDLEY APPROXIMATION

According to Ref. 17, the posterior SELF-Bayes estimator of an arbitrary function $u(\theta)$ given by Lindley is

$$E\{(\theta)|x\} = \frac{\int u(\theta)v(\theta)\exp[L(\theta)]\,\mathrm{d}\theta}{\int v(\theta)\exp[L(\theta)]\,\mathrm{d}\theta}. \tag{15}$$

This can be approximated asymptotically by

$$\begin{aligned} E\{u(\theta)|x\} = u &+ \frac{1}{2}\sum_i\sum_j (u_{ij} + 2u_i.\rho_j).\sigma_{ij} \\ &+ \frac{1}{2}\sum_i\sum_j\sum_k\sum_l L_{ijk}.\sigma_{ij}.\sigma_{kl}.u_l \end{aligned} \tag{16}$$

where $i, j, k, l = 1, 2, \dots, n$ and $\theta = (\theta_1, \theta_2, \dots, \theta_m)$ and m is the number of parameters. Taking only the scale parameter of the Weibull distribution into consideration, (16) reduces to Ref. 18

$$\hat{u} = u(\theta) + \tfrac{1}{2}[(u_{11}\sigma_{11})] + u_1\rho_1\sigma_{11} + \tfrac{1}{2}[(L_{30}u_1\sigma_{11}^2)] \tag{17}$$

where L is the log-likelihood function in (5). When we consider first the non-informative prior with the squared error loss function, we have

$$\rho_1 = -\frac{1}{\alpha}, \quad \sigma_{11} = (-L_{20})^{-1},$$
$$u = \alpha,\ u_1 = \frac{\partial u}{\partial \alpha} = 1,\ u_{11} = \frac{\partial^2 u}{\partial \alpha^2} = 0.$$

Let the following definitions hold:

$$x = \exp\left[-\left(\frac{L_i}{\alpha}\right)^\beta\right], \quad y = \exp\left[-\left(\frac{R_i}{\alpha}\right)^\beta\right],$$
$$e = -\left(\frac{L_i}{\alpha}\right)^\beta, \quad f = -\left(\frac{R_i}{\alpha}\right)^\beta.$$

Let L_{20} and L_{30} represent the second and third derivatives of the log-likelihood function such that

$$\begin{aligned} L_{20} = \frac{\partial^2 \ell}{\partial \alpha^2} = \sum_{i=1}^{n} \frac{1}{x-y}\Bigg\{ & e\left(\frac{\beta^2}{\alpha^2}\right)\ x + e\left(\frac{\beta}{\alpha^2}\right)x \\ & + e^2\left(\frac{\beta^2}{\alpha^2}\right)x - f\left(\frac{\beta^2}{\alpha^2}\right)y - f\left(\frac{\beta}{\alpha^2}\right)y \\ & - f^2\left(\frac{\beta^2}{\alpha^2}\right)y\Bigg\} - \frac{\left[e\left(\frac{\beta}{\alpha}\right)x + f\left(\frac{\beta}{\alpha}\right)y\right]^2}{[x-y]^2}, \end{aligned}$$

Table 1 Average mean squared errors for the scale parameter.

Size	α	β	ML	BN	BGN		
					$k=2$	$k=3$	$k=4$
25	0.5	0.8	0.09839	0.09733	0.09645	0.09497	0.10673
		1.2	0.06645	0.06576	0.06464	0.06283	0.06196
	1.5	0.8	0.89040	0.88025	0.87713	0.85366	0.86518
		1.2	0.58682	0.58018	0.58212	0.58275	0.58034
50	0.5	0.8	0.09686	0.09666	0.09608	0.09283	0.09071
		1.2	0.06335	0.06321	0.06363	0.06276	0.05867
	1.5	0.8	0.87340	0.87144	0.87276	0.85265	0.85782
		1.2	0.57503	0.57352	0.58065	0.58054	0.57389
100	0.5	0.8	0.09655	0.09650	0.09520	0.08125	0.07740
		1.2	0.06302	0.06300	0.06319	0.06074	0.05462
	1.5	0.8	0.85686	0.85647	0.86170	0.85258	0.84474
		1.2	0.57348	0.57319	0.57343	0.56889	0.57294

ML = maximum likelihood, BN = Bayes non-informative prior, BGN = Bayes generalized non-informative prior

$\sigma_{11} = (-L_{20})^{-1}$, $\sigma_{22} = (-L_{02})^{-1}$, and

$$L_{30} = \frac{\partial^3 \ell}{\partial \alpha^3} = \sum_{i=1}^{n} \left[\frac{1}{x-y}\right] \left\{-e\left(\frac{\beta^3}{\alpha^3}\right)x - 3e\left(\frac{\beta^2}{\alpha^3}\right)x - 3e^2\left(\frac{\beta^3}{\alpha^3}\right)x - 2f\left(\frac{\beta}{\alpha^3}\right)x - 3e^2\left(\frac{\beta^2}{\alpha^3}\right)x - e^3\left(\frac{\beta^3}{\alpha^3}\right)x + f\left(\frac{\beta^3}{\alpha^3}\right)y + 3f\left(\frac{\beta^2}{\alpha^3}\right)y + 3f^2\left(\frac{\beta^3}{\alpha^3}\right)y + 2f\left(\frac{\beta}{\alpha^3}\right)y + 3f^2\left(\frac{\beta^2}{\alpha^3}\right)y + f^3\left(\frac{\beta^3}{\alpha^3}\right)y\right\} - \left[\frac{1}{x-y}\right]^2 \times \left\{3\left[e\left(\frac{\beta^2}{\alpha^2}\right)x + e\left(\frac{\beta}{\alpha^2}\right)x + e^2\left(\frac{\beta^2}{\alpha^2}\right)x - f\left(\frac{\beta^2}{\alpha^2}\right)y - f\left(\frac{\beta}{\alpha^2}\right)y - f^2\left(\frac{\beta^2}{\alpha^2}\right)y\right] \times \left[-e\left(\frac{\beta}{\alpha}\right)x + f\left(\frac{\beta}{\alpha}\right)y\right]\right\} + \frac{2\left[-e\left(\frac{\beta}{\alpha}\right)x + f\left(\frac{\beta}{\alpha}\right)y\right]^3}{[x-y]^3}.$$

SIMULATION STUDY

A simulation study was carried out to determine the best estimator for the scale parameter of the Weibull distribution with interval censoring and sample sizes of $n = 25$, 50, and 100. The coding and the analysis were performed using the R programming language. The scale parameter was estimated with maximum likelihood and Bayesian using non-informative prior and a generalized non-informative prior approach. The values for the generalized non-informative prior are $k = 2$, 3, and 4. These were iterated $R = 1000$ times. The mean squared errors and that of the absolute bias are determined and presented in Table 1 and Table 2 for the purpose of comparison. The mean squared error and absolute bias are given, respectively, by

$$\mathrm{MSE}(\hat{\theta}) = \frac{1}{R}\sum_{r=1}^{1000}(\hat{\theta}^r - \theta)^2,$$

$$\text{Abs Bias}(\hat{\theta}) = \frac{1}{R}\sum_{r=1}^{1000}|\hat{\theta}^r - \theta|.$$

Table 2 Average absolute bias for the scale parameter.

Size	α	β	ML	BN	BGN		
					$k=2$	$k=3$	$k=4$
25	0.5	0.8	0.30904	0.30896	0.30818	0.30650	0.29946
		1.2	0.25188	0.25051	0.25117	0.24932	0.24760
	1.5	0.8	0.92221	0.91634	0.92597	0.91869	0.93131
		1.2	0.74837	0.74328	0.75283	0.75284	0.75335
50	0.5	0.8	0.30772	0.30736	0.30695	0.30123	0.26214
		1.2	0.24853	0.24824	0.24994	0.24770	0.23898
	1.5	0.8	0.92348	0.92239	0.92231	0.91184	0.91974
		1.2	0.74926	0.74824	0.75257	0.74999	0.74842
100	0.5	0.8	0.30623	0.30448	0.30135	0.27205	0.21823
		1.2	0.24957	0.24952	0.24657	0.24040	0.18939
	1.5	0.8	0.92048	0.92026	0.91219	0.90179	0.88705
		1.2	0.75290	0.75271	0.74512	0.74603	0.74564

CONCLUSIONS

Bayesian non-informative prior and the proposed generalized non-informative prior estimators of the Weibull scale parameter are obtained using a squared error loss function via the Lindley approximation. Comparisons are made between the estimators based on a simulation study. The performance of MLE and Bayes using the non-informative priors are examined and the following conclusions made.

Table 1 shows the mean squared error values of the scale parameter. It is observed that the proposed generalized Bayes non-informative prior estimator has smaller mean squared errors than the others except at $\alpha = 1.5$ and $\beta = 1.2$ with $n = 25$ where Bayes using the Jeffreys vague non-informative prior gave the smallest mean squared error. As the sample size increased there was a corresponding decrease in mean square error values for all the estimators.

The absolute bias of the estimated values are presented in Table 2. We observe that all the estimators are not very reliable since they all exhibit some level of bias but the generalized Bayes non-informative prior has a lower absolute bias than the other estimators. The lowest absolute biases occur mostly at $k = 4$.

REFERENCES

1. Al-Aboud FM (2009) Bayesian estimation for the extreme value distribution using progressive censored data and asymmetric loss. *Int Math Forum* **4**, 1603–22.
2. Al-Athari FM (2011) Parameter estimation for the double-Pareto distribution. *J Math Stat* **7**, 289–94.
3. Syuan-Rong H, Shuo-Jye W (2011) Bayesian estimation and prediction for Weibull model with progressive censoring. *J Stat Comput Simulat* **81**, 1607–20.
4. Guure CB, Ibrahim NA, Adam MB (2013) Bayesian inference of the Weibull model based on interval-censored survival data. *Comput Math Meth Med* **2013**, 849520.
5. Pandey BN, Dwividi N, Pulastya B (2011) Comparison between Bayesian and maximum likelihood estimation of the scale parameter in Weibull distribution with known shape under linex loss function. *J Sci Res* **55**, 163–72.
6. Soliman AA, Abd Ellah AH, Ryan SB (2006) Comparison of estimates using record statistics from Weibull model: Bayesian and Non-Bayesian Approaches. *Comput Stat Data Anal* **51**, 2065–77.
7. Abdel-Wahid AA, Winterbottom A (1987) Approximate Bayesian estimates for the Weibull reliability function and hazard rate from censored data. *J Stat Plann Infer* **16**, 277–83.
8. Turnbull BW (1976) The empirical distribution with arbitrarily grouped censored and truncated data. *J Roy Stat Soc B* **38**, 290–5.
9. Jianguo S (2006) *The Statistical Analysis of Interval-Censored Failure Time Data*, Springer, New York.
10. Lawless JF (2003) *Statistical Models and Methods for Lifetime Data*, Wiley, New York.
11. Flygare ME, Buckwalter JA (2010) Maximum likelihood estimation for the 2-parameter Weibull distribution based on interval-data. *IEEE Trans Reliab* **R-34**, 57–60.
12. Lindsey JK (1998) A study of interval censoring in parametric regression models. *Lifetime Data Anal* **4**, 329–54.
13. Scallan AJ (1999) Regression modelling of interval censored failure time data using the Weibull distribution. *J Appl Stat* **26**, 613–8.
14. Odell PM, Anderson KM, D'Agostino RB (1992) Maximum likelihood estimation for interval-censored data using a Weibull-based accelerated failure time model. *Biometrics* **48**, 951–9.
15. Gomez G, Luz Calle M, Ramon O (2004) Frequentist and Bayesian approaches for interval-censored data. *Stat Paper* **45**, 139–73.
16. Sinha SK (1986) Bayes estimation of the reliability function and hazard rate of a weibull failure time distribution. *Tranbajos de Estadistica* **1**, 47–56.
17. Lye LM, Hapuarachchi KP, Ryan S (2010) Bayes estimation of the extreme-value reliability function. *IEEE Trans Reliab* **42**, 641–4.
18. Guure CB, Ibrahim NA, Al Omari MA (2012) Bayesian estimation of two-parameter Weibull distribution using extension of Jeffreys' prior information with three loss functions. *Math Probl Eng* **2012**, 589640.