Journal of Statistical
Theory & Methods
1997, Vol. 8, No. 1, pp. 41~48

Empirical Bayes Interval Estimation by a Sample Reuse Method

Kil-Ho Cho · Dal-Woo Choi · Hyeon-Sook Chae[1]

Abstract We construct the empirical Bayes(EB) confidence intervals that attain a specified level of EB coverage for the unknown scale parameter in the Weibull distribution with the known shape parameter under the type II censored data. Our general approach is to use an EB bootstrap samples introduced by Larid and Louis(1987). Also, we compare the coverage probability and the expected interval length for these bootstrap intervals with those of the naive intervals through Monte Carlo simulation.

Keywords: Bootstrap, Empirical Bayes confidence interval, Type II censored data Weibull distribution.

1. Introduction

There are two main classes of EB estimators. The first class consists of methods that attempt to approximate the Bayes estimator without explicitly estimating the unknown prior distribution. The second class consists of methods in which the unknown prior distribution is explicitly estimated. The smooth EB methods developed by Maritz(1966,1967), Lemon and Krutchkoff(1969), and Couture and Martz(1972) are based on the second method. The posterior distribution is easily obtained for the second class of EB estimators and EB confidence intervals can be obtained from it. In this paper, we consider the smooth EB estimations of the scale parameter for the Weibull distribution given by

$$f\left(t_{ij}\middle|\theta_i,\beta\right)=\theta_i\beta t_{ij}^{\beta-1}exp\left(-\theta_i t_{ij}^{\beta}\right),\quad t_{ij}\geq 0,\ \theta_i,\ \beta>0. \tag{1}$$

where θ_i is the scale parameter and β is the shape parameter whose value is assumed to be known. We assume that a sequence of $N(\geq 2)$ life test experiments has been conducted in which n_i items in the i -th experiment are placed on life

---

[1] Kil Ho Cho is an Associate Professor, Department of Statistics, Kyungpook National University, Taegu, 702-701, Korea. Dal-Woo Choi and Hyeon Sook Chae are Part-time Lecturers, Department of Statistics, Kyungpook National University, Taegu, 702-701, Korea.

test which is terminated at the time when the number of failures is r_i, $1 \le r_i \le n_i$. The ordered failure times $t_{i1} \le t_{i2} \le \cdots \le t_{ir_i}$ are recorded in each experiment. The statistic

$$S_i = \sum_{j=1}^{r_i} t_{ij}^{\beta_i} + (n_i - r_r) t_{ir_i}^{\beta_i} \tag{2}$$

is the sufficient statistic for the scale parameter θ_i. And conditional on θ_i, S_i has a gamma distribution $G(r_i, \theta_i)$. It is also easily shown that the maximum likelihood(ML) estimator of θ_i, $\hat{\theta}_i = r_i / S_i$, has an inverse gamma distribution $IG(r_i, \theta_i r_i)$. Lemon and Krutchkoff(1969) proposed a smoothing procedure which may be interpreted the approximation of the prior distribution by a step function having steps of equal height $1/N$ at each of the ML estimates $\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_N$. They also suggested a second possible interaction with each ML estimate $\hat{\theta}$ replaced by the corresponding EB estimate from the first iteration. Such EB estimators will be referred to the iterated EB estimators.

2. EB Confidence Intervals

We consider the iterated EB estimations of the scale parameter for the Weibull distribution under the type II censoring scheme. Suppose θ is univariate. If $\hat{\theta}_j$ denotes the ML estimate of θ_j and has an $IG(r_j, \theta_j r_j)$ in the j-th experiment, $j = 1, 2, \cdots, N$, then the EB estimator of θ_j for the first iteration is given by

$$\begin{aligned} \tilde{\theta}_j &= \frac{\sum_{k=1}^{N} \hat{\theta}_k f(\hat{\theta}_j | \hat{\theta}_k)}{\sum_{k=1}^{N} f(\hat{\theta}_j | \hat{\theta}_k)}, \; j = 1, 2, \cdots, N \\ &= \frac{\sum_{k=1}^{N} \hat{\theta}_k^{r_j+1} \exp(-\hat{\theta}_k r_j / \hat{\theta}_j)}{\sum \hat{\theta}_k^{r_j} \exp(-\hat{\theta}_k r_j / \hat{\theta}_j)} \end{aligned} \tag{3}$$

where

$$f(\hat{\theta}_j | \hat{\theta}_k) = \frac{(\hat{\theta}_k r_j)^{r_j}}{\Gamma(r_j)} \left(\frac{1}{\hat{\theta}_j}\right)^{r_j} \exp\left(-\frac{\hat{\theta}_k r_j}{\hat{\theta}_j}\right), \; j = 1, 2, \cdots, N \tag{4}$$

and

$$\hat{\theta}_j = r_j / S_j = r_j \Big/ \left[\sum_{k=1}^{r_j} t_{jk}^{\beta_j} + \left(n_j - r_j\right) t_{jr_j}^{\beta_j} \right]. \tag{5}$$

Using $\tilde{\theta}_j$'s in equation (3), the iterared EB estimator of θ_i is given by

$$\hat{\theta}_{LK_i} = \frac{\sum_{j=1}^{N} \tilde{\theta}_j f\left(\hat{\theta}_i \middle| \tilde{\theta}_j\right)}{\sum_{j=1}^{N} f\left(\hat{\theta}_i \middle| \tilde{\theta}_j\right)}, \quad i = 1, 2, \cdots, N$$

$$= \frac{\sum_{j=1}^{N} \tilde{\theta}_j^{r_i+1} \exp\left(-\tilde{\theta}_j r_i / \hat{\theta}_i\right)}{\sum_{j=1}^{N} \tilde{\theta}_j^{r_i} \exp\left(-\tilde{\theta}_j r_i / \hat{\theta}_i\right)}. \tag{6}$$

2.1 Naive EB Confidence Interval

Martz and Waller(1982) obtained the iterated EB naive confidence interval of θ_i as follows. Let $g_N(\tilde{\theta}_j|\hat{\theta}_j)$ denote the estimated posterior probability mass function, where $\tilde{\theta}_j$ is the EB point estimate of θ_j according to equation (4). Then the estimated posterior distribution of θ_i according to equation (5) is given by

$$g_N\left(\tilde{\theta}_j \middle| \hat{\theta}_i\right) = \frac{f\left(\hat{\theta}_i \middle| \tilde{\theta}_j\right)}{\sum_{j=1}^{N} f\left(\hat{\theta}_i \middle| \tilde{\theta}_j\right)}, \quad j = 1, 2, \cdots, N$$

$$= \frac{\tilde{\theta}_j^{r_i} \exp\left(-\tilde{\theta}_j r_i / \hat{\theta}_i\right)}{\sum_{j=1}^{N} \tilde{\theta}_j^{r_i} \exp\left(-\tilde{\theta}_j r_i / \hat{\theta}_i\right)} \tag{7}$$

where $\tilde{\theta}_j$ is given by equation (4). Therefore, they constructed the symmetric $100(1-2\alpha)\%$ iterated EB naive confidence interval of θ_i based upon $g_N(\tilde{\theta}_j|\hat{\theta}_i)$ given by

$$\left(\tilde{\theta}_{(j^-+1)},\ \tilde{\theta}_{(j^+)}\right), \tag{8}$$

where $\tilde{\theta}_{(1)}, \cdots, \tilde{\theta}_{(N)}$ denote the ordered (smallest to largest) sequence of EB estimates of $\theta_1, \cdots, \theta_N$, j^+ is the smallest integer satisfying the relation

$$\sum_{j=1}^{j^+} g_N\left(\tilde{\theta}_{(j)} \,\middle|\, \hat{\theta}_i\right) = \frac{\sum_{j=1}^{j^+} \tilde{\theta}_{(j)}^{r_i} \exp\left(-\tilde{\theta}_{(j)} r_i / \hat{\theta}_i\right)}{\sum_{j=1}^{N} \tilde{\theta}_j^{r_i} \exp\left(-\tilde{\theta}_j r_i / \hat{\theta}_i\right)}$$

$$\geq 1 - \alpha \tag{9}$$

and j^- is the largest integer satisfying the relation

$$\sum_{j=1}^{j^-} g_N\left(\tilde{\theta}_{(j)} \,\middle|\, \hat{\theta}_i\right) = \frac{\sum_{j=1}^{j^-} \tilde{\theta}_{(j)}^{r_i} \exp\left(-\tilde{\theta}_{(j)} r_i / \hat{\theta}_i\right)}{\sum_{j=1}^{N} \tilde{\theta}_j^{r_i} \exp\left(-\tilde{\theta}_j r_i / \hat{\theta}_i\right)}$$

$$\leq \alpha . \tag{10}$$

If no such integer j^- exists between 1 and N, then j^- sets to zero. Thus the lower limit of the estimate will be $\tilde{\theta}_{(1)}$.

2.2 Marginal EB Bootstrap Interval

We construct the marginal EB bootstrap interval to correct the bias of the naive interval using the type II nonparametric(smoothed) bootstrap which is a sample reuse method introduced by Laird and Louis(1987). The marginal EB bootstrap procedure for the symmetric $100(1-2\alpha)\%$ iterated EB confidence interval of θ_i may be described as follows:

(1) Generate the bootstrap samples $t_{ij}^*, j = 1, 2, \cdots, n_i$ from the sample distribution

$$f\left(t_{ij} \,\middle|\, \hat{\theta}_i, \beta\right) = \hat{\theta}_i \beta t_{ij}^{\beta-1} \exp\left(-\hat{\theta}_i t_{ij}^{\beta}\right), \quad i = 1, 2, \cdots, N, \tag{11}$$

where $\hat{\theta}_i$ is the ML estimator of θ_i.

(2) Construct the bootstrap sufficient statistic S_i^*. That is,

$$S_i^* = \sum_{j=1}^{r_i} t_{ij}^{*\beta_i} + (n_i - r_i) t_{ir_i}^{*\beta_i} . \tag{12}$$

(3) Compute the bootstrap ML estimator θ_i^* from the S_i^*'s, $i = 1, 2, \cdots, N$. That is,

$$\theta_i^* = \frac{r_i}{S_i^*} . \tag{13}$$

(4) Compute the first EB bootstrap estimator $\hat{\theta}_j^{*b}$ in the same way as $\tilde{\theta}_j$ in

equation (3) from the b-th bootstrap samples. That is,

$$\hat{\theta}_j^{*b} = \frac{\sum_{j=1}^{N} \theta_j^{(r_i+1)*} \exp\left(-\theta_j^* r_i / \hat{\theta}_i\right)}{\sum_{j=1}^{N} \theta_j^{(r_i)*} \exp\left(-\theta_j^* r_i / \hat{\theta}_i\right)} . \tag{14}$$

(5) Construct the bootstrap interval $[\hat{\theta}_{(j^-+1)}^{*b}, \hat{\theta}_{(j^+)}^{*b}]$ using $\tilde{\theta}_j$ replaced by $\hat{\theta}_j^{*b}$ in equation (9) and equation (10), where $\hat{\theta}_{(j^+)}^{*b}$ is the smallest value $\hat{\theta}_{(j)}^{*b}$ satisfying the inequality

$$\sum_{j=1}^{j^+} g_N\left(\hat{\theta}_{(j)}^{*b} \mid \hat{\theta}_i\right) = \frac{\sum_{j=1}^{j^+} \left\{\hat{\theta}_{(j)}^{*b}\right\}^{r_i} \exp\left(-\hat{\theta}_{(j)}^* r_i / \hat{\theta}_i\right)}{\sum_{j=1}^{N} \left\{\hat{\theta}_{j=1}^{*b}\right\}^{r_i} \exp\left(-\hat{\theta}_j^* r_i / \hat{\theta}_i\right)}$$

$$\geq 1-\alpha , \tag{15}$$

and $\hat{\theta}_{(j^-)}^{*b}$ is the largest value $\hat{\theta}_{(j)}^{*b}$ satisfying the inequality

$$\sum_{j=1}^{j^-} g_N\left(\hat{\theta}_{(j)}^{*b} \mid \hat{\theta}_i\right) = \frac{\sum_{j=1}^{j^-} \left\{\hat{\theta}_{(j)}^{*b}\right\}^{r_i} \exp\left(-\hat{\theta}_{(j)}^* r_i / \hat{\theta}_i\right)}{\sum_{j=1}^{N} \left\{\hat{\theta}_j^{*b}\right\}^{r_i} \exp\left(-\hat{\theta}_j^* / \hat{\theta}_i\right)}$$

$$\leq \alpha , \tag{16}$$

where $\hat{\theta}_{(1)}^{*b}, \cdots, \hat{\theta}_{(N)}^{*b}$ denote the ordered (smallest to largest) EB bootstrap estimates of $\theta_1, \cdots, \theta_N$.

(6) Repeat this process B times.

(7) Obtain the iterated EB bootstrap interval (θ_L^* , θ_U^*) of θ_i, where

$$\theta_{L_i}^* = \frac{1}{B}\sum_{b=1}^{B} \hat{\theta}_{(j^-+1)}^{*b} \quad \text{and} \quad \theta_{U_i}^* = \frac{1}{B}\sum_{b=1}^{B} \hat{\theta}_{(j^+)}^{*b} .$$

Thus

$$\Pr\left[\theta_{L_i}^* \leq \theta_i \leq \theta_{U_i}^* \mid \hat{\theta}_i\right] = 1-2\alpha .$$

2.3 Percentile Bootstrap Confidence Interval

Efron(1981a, b) introduced the percentile method which constructs approximate confidence interval for θ based on the sampling distribution of bootstrap estimator of θ. This approximate confidence interval by percentile method is called percentile interval. Let us construct the EB percentile interval of θ_i using the type II nonparametric bootstrap. Then $100(1-2\alpha)\%$ percentile interval for is θ_i approximated by the following steps.

(1) Construct the first EB bootstrap estimator $\hat{\theta}_i^{*b}$ as the same marginal procedure(Laird and Louis(1987)) in equation (14) from the b-th bootstrap samples. That is,

$$\hat{\theta}_i^{*b} = \frac{\sum_{i=1}^{N} \theta_i^{(r_i+1)*} \exp\left(-\theta_i^* r_i / \hat{\theta}_i\right)}{\sum_{i=1}^{N} \theta_i^{(r_i)*} \exp\left(-\theta_i^* r_i / \hat{\theta}_i\right)}.$$

Construct the iterated EB bootstrap estimator $\theta_{LK_i}^{*b}$ in the same way as $\hat{\theta}_{LK_i}$ in equation (6). That is,

$$\theta_{LK_i}^{*b} = \frac{\sum_{j=1}^{N} \left\{\hat{\theta}_j^{*b}\right\}^{r_i+1} \exp\left(-\hat{\theta}_j^* r_i / \hat{\theta}_i\right)}{\sum_{j=1}^{N} \left\{\hat{\theta}_j^{*b}\right\}^{r_i} \exp\left(-\hat{\theta}_j^* r_i / \hat{\theta}_i\right)}. \tag{17}$$

(3) Repeat this process B times.

(4) Compute the 100α empirical percentiles $\theta_{LK_i}^{*b}(\alpha)$ and the $100(1-\alpha)$ empirical percentile $\theta_{LK_i}^{*b}(1-\alpha)$ of $\theta_{LK_i}^{*b}$ respectively, where $\theta_{LK_i}^{*b}(\alpha)$ is the $B\alpha$-th value in the ordered list of the B replications of $\theta_{LK_i}^{*b}$, $b = 1, 2, \cdots, B$.

(5) Approximate $100(1-2\alpha)\%$ percentile interval for θ_i by

$$\left(\theta_{LK_i}^{*b}(\alpha),\ \theta_{LK_i}^{*b}(1-\alpha)\right).$$

3. Comparisons and Conclusions

The iterated EB confidence intervals are approximated by Monte Carlo method. In each iteration, we generate θ_i, $i = 1, \cdots, N(=10)$, from IMSL subroutine RNGAM. Given the θ_i's, we generate t_{ij}, $j = 1, \cdots, n(=10, 20, 40)$, from IMSL subroutine RNWIB and a fixed shape parameter $\beta = 2$. We order the variables $t_{i1} \le t_{i2} \le \cdots \le t_{ir}$, and compute $S_i = (10-r)t_{ir}^{\beta} + \sum_j^r t_{ij}^{\beta}$ (Note that we assume $r_i = r$ for all i). We consider the censoring rates (CR) defined by $100(1-r/n)\%$ of $0\%, 30\%$, and 50%. For given independent random samples the iterated EB confidence intervals are computed by each method with bootstrap replications $B = 1000$ times. And the Monte Carlo samplings are repeated 1000 times. The criteria used to compare the EB interval are coverage probability and expected interval length, where the probability and length refer to the distribution of S in equation (2). Let CP denote the coverage probability for θ and we consider nominal coverage probability of $0.90(\alpha = 0.05)$. Let $\hat{\theta}_{lo}$ and $\hat{\theta}_{up}$ be the lower limit and the upper limit of EB confidence interval for θ, respectively. And we define EL by

$$EL = \frac{1}{R}\sum_{j=1}^{R}\left(\hat{\theta}_{j,up} - \hat{\theta}_{j,lo}\right), \tag{18}$$

where R is the number of Monte Carlo simulation replications. The results of these simulations are presented in Table 1. To asses the CP and the EL of the naive and two bootstrap intervals for the sample sizes $n = 10, 20, 40$, when the CR changes, we observe Table 1 for the shape parameter $\beta = 2$. We can observe the followings:

(1) For all the sample sizes and censoring rates, the CP's of two bootstrap intervals are more accurate than those of the naive intervals.

(2) As n increases, the percentile bootstrap interval tends to achieve the nominal coverage probability.

(3) For all the sample sizes, the EL's of two bootstrap intervals are longer than those of naive intervals.

(4) As expected, the naive interval fails to achieve nominal coverage probability and is very poor for EL's even if n increases.

Table 1. Comparisons of EB Naive and Bootstrap Intervals When $\beta = 2$ (Raleigh case) and $\alpha = 0.05$

Censoring rate = 0%

Interval method	n=10 Coverage	Length	n=20 Coverage	Length	n=40 Coverage	Length
Naive	0.432	1.952	0.572	1.604	0.601	1.072
Laird & Louis	0.678	2.790	0.784	1.842	0.841	1.177
Percentile	0.891	3.805	0.902	2.619	0.903	1.854

Censoring rate = 30%

Interval method	n=10 Coverage	Length	n=20 Coverage	Length	n=40 Coverage	Length
Naive	0.525	2.274	0.620	1.822	0.604	1.336
Laird & Louis	0.769	3.942	0.823	2.388	0.848	1.521
Percentile	0.865	5.431	0.898	3.169	0.909	2.215

Censoring rate = 50%

Interval method	n=10 Coverage	Length	n=20 Coverage	Length	n=40 Coverage	Length
Naive	0.545	2.341	0.574	1.894	0.560	1.493
Laird & Louis	0.780	4.556	0.827	2.804	0.851	1.812
Percentile	0.858	6.319	0.892	3.851	0.902	2.555

References

1. Efron, B. (1981a). Nonparametric standard errors and confidence intervals(with comment), *Canadian Journal of Statistics*, 9, 139-172.

2. Efron, B. (1981b). Nonparametric estimates of standard error: The Jackknife, the bootstrap, and other methods, *Biometrika*, 68, 589-599.

3. Laird, N. M. and Louis, T. A. (1987). Empirical Bayes confidence intervals based on bootstrap samples(with comments), *Journal of American Statistical Association*, 82, 739-757.

4. Lemon, G. H. and Krutchkoff, R. G. (1969). An empirical Bayes smoothing technique, *Biometrika*, 56, 2, 361-365.

5. Maritz, J. S. (1966). Smooth empirical Bayes estimation for one-parameter discrete distributions, *Biometrika*, 53, 417-429.

6. Maritz, J. S. (1967). Smooth empirical Bayes estimation for continuous distributions, *Biometrika*, 54, 435-450.

7. Martz, H. F. and Waller, R. A. (1982). *Bayesian Reliability Analysis*, John Wiley & Sons, New York.