Bayesian Estimation of a Weibull Distribution in a Highly Censored and Small Sample Setting

Mostafa Bacha, Gilles Celeux

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Bayesian Estimation of a Weibull distribution in a highly censored and small sample setting

Mostafa Bacha, Gilles Celeux

N° 2993

Octobre 1996

THÈME 4

Rapport de recherche

ISSN 0249-6399

Bayesian Estimation of a Weibull distribution in a highly censored and small sample setting

Mostafa Bacha, Gilles Celeux

Thème 4 — Simulation et optimisation
de systèmes complexes
Projet ISIS

Rapport de recherche n° 2993 — Octobre 1996 — 16 pages

Abstract: We propose and investigate through Monte Carlo simulations two methods for Bayesian inference for the shape and the scale parameters of a Weibull distribution in a small and highly censored sample setting. The first method, WLB-SIR, is the Sampling Importance Resampling-adjusted Weighted Likelihood Bootstrap of Newton and Raftery. The second one is a new Bayesian Restoration Maximization (BRM) new algorithm working along the same line but replacing the bootstrap weighting step by a stochastic simulation step of the censored failure times. The advantage of the BRM method is that it takes account of the prior distribution in its first step. As a consequence, the Sampling Importance Resampling-adjusted version of BRM, BRM-SIR, is less fragile than WLB-SIR as it appears from our numerical experiments for Weibull parameters estimation. This article also includes a flexible procedure to transform prior knowledge into prior distributions on the Weibull parameters.

Key-words: Weibull distribution, Bayesian estimation, BRM algorithm, WLB algorithm, Importance Sampling, censored data, Monte Carlo simulation.

(Résumé : tsvp)

Unité de recherche INRIA Rhône-Alpes
655, avenue de l' Europe, 38330 MONTBONNOT ST MARTIN (France)
Téléphone : (33) 76 61 52 00 – Télécopie : (33) 76 61 52 52

Estimation bayésienne des paramètres d'une loi de Weibull pour de petits échantillons très censurés.

Résumé : Nous proposons deux méthodes d'inférence bayésienne pour estimer le paramètre de forme et d'échelle d'une loi de Weibull pour de petits échantillons très censurés. Nous analysons leurs performances par des simulations de Monte Carlo. Ces deux méthodes sont le *Weighted Likelihood Bootstrap* (WLB) de Newton et Raftery et une nouvelle méthode *Bayesian Restoration Maximization* (BRM) qui travaille dans le même esprit, mais remplace les tirages de poids à l'aide du bootstrap par la simulation des données censurées. L'avantage de BRM sur WLB est qu'il tient compte dans sa phase initiale de la loi a priori des paramètres. Ainsi, il fournit des résultats plus fiables et plus stables comme le montrent nos expérimentations numériques. Cet article propose de plus une procédure souple pour traduire les simples connaissances a priori en distributions a priori.

Mots-clé : Loi de Weibull, estimation bayésienne, algorithme BRM, algorithme WLB, échantillonnage préférentiel, simulation de Monte Carlo.

1 Introduction

In this paper, we are concerned with the problem of estimating the shape and scale parameters, β and η, of a Weibull distribution from small and highly censored samples. In many circumstances, the Maximum Likelihood (ML) approach provides satisfactory estimates of these parameters and can be regarded as a reference technique. However, ML estimation turns out to be imprecise or even unreliable for small or highly censored samples [4]. On an other hand, Bayesian inference is desirable for estimating industrial lifetime models involving Weibull distributions, since available samples are small and highly censored and, generally, industrial experts have prior information on the Weibull distribution parameters. But, direct Bayesian analysis for Weibull distribution is quite difficult, especially for the shape parameter (see, for instance, [14]). Even Markov Chain Monte Carlo (MCMC) methods appear to be deceptive in this context.

In this paper, we consider two methods using Importance Sampling to implement the Bayes paradigm. Those methods are the Importance Sampling-adjusted Weighted Likelihood Bootstrap (WLB-IS) of Newton and Raftery [17] and a new method, the Importance Sampling-adjusted Bayesian Restoration Maximization (BRM-IS), which takes advantage of the incomplete data structure of the estimatiom problem from censored lifetimes.

The paper is organized as follows. Section 2 reviews the maximum likelihod and Bayesian attempts for estimating the parameters of a Weibull distribution and sketch Importance Sampling techniques. The two methods, WLB and BRM, are presented in Section 3. Section 4 is devoted to the formulation of an expert opinion in a prior distribution on the parameters. In Section 5, numerical Monte Carlo experiments are reported. They highlight the practical usefulness of both methods and show their superiority over ML estimation in the particular context under consideration. They also show that the BRM methodology can be preferred to the WLB methodology.

2 Importance Sampling for Bayesian inference

We consider a Type I right censored sample of size N, $\boldsymbol{t} = \{t_1, \ldots, t_N\}$. For convenience and simplicity, we will assume that the sample has been rearranged so that the corresponding m ($m > 0$) uncensored failure times are $t_i, i = 1, \ldots, m$, where each t_i arises from a Weibull distribution with density

$$f(t|\theta) = \frac{\beta}{\eta}(\frac{t}{\eta})^{\beta-1} \exp[-(\frac{t}{\eta})^{\beta}],\ t > 0,\ \beta > 0,\ \eta > 0, \tag{2.1}$$

where $\theta = (\beta, \eta)$, β denoting the shape parameter and η the scale parameter. And, $(t_{m+1}, \ldots, t_N)$ denote the $(N - m)$ right censored data.

The likelihood of the observed sample is

$$L(\boldsymbol{t}; \theta) = \prod_{i=1}^{m} f(t_i \mid \theta) \prod_{i=m+1}^{N} R(t_i \mid \theta)$$

where $R(t \mid \theta) = \exp\left(-(t/\eta)^{\beta}\right)$ is the Weibull survival function.

The maximum likelihood estimate $\widehat{\beta}$ of the shape parameter is the solution of the implicit equation

$$\frac{1}{\beta} + \frac{\sum_{i=1}^{m} \log(t_i)}{m} - \left[\frac{\sum_{i=1}^{N} t_i^{\beta} \log(t_i)}{\sum_{i=1}^{N} t_i^{\beta}}\right] = 0$$

and

$$\widehat{\eta} = \left[\sum_{i=1}^{N} t_i^{\widehat{\beta}}/m\right]^{\frac{1}{\widehat{\beta}}}.$$

It is well known that $(\widehat{\beta}, \widehat{\eta})$ are biased for small samples and for censored samples, even with large N. Moreover, the asymptotic variances of the maximum likelihood estimates are increasing functions of the censoring rate m/N [4].

In Engineering applications, Bayesian methods can be expected to be useful since, in many circumstances, there exists some knowledge on the underlying failure mechanism; also, any amount of prior information is useful in a highly censored setting.

In Bayesian inference, a prior probability distribution $\pi(\theta)$ is specified and leads to the posterior distribution $\pi(\theta \mid \boldsymbol{t}) \propto L(\boldsymbol{t}; \theta)\pi(\theta)$ for the parameters

$\theta = (\beta, \eta)$. The Bayesian estimation of any function of interest $\Phi(\theta)$, under the quadratic loss function, is the posterior mean

$$I(\Phi) = \int \Phi(\theta)\pi(\theta \mid \boldsymbol{t})\, d\theta.$$

Calculating this integral analytically is rarely possible for realistic choice of prior distributions. Berger [5], Ellis & Tummala [9] and Evans & Nigm [11] consider discrete prior distribution for the shape parameter to obtain explicit formula for posterior estimates; but if continuous prior distribution is choosen, numerical approximations are needed. There exists many procedures producing such approximations. Achcar & Louzada [1], Singpurwalla & Song [21] and Sinha & Sloan [22] use Laplace approximation [24] which is valid, theoretically, only for large data sets. Bogdanoff & Pierce [7], Calabria, Guida and Pulcini [8], Erto [10], Naylor & Smith [15], Reilly [18] and Singpurwalla [20] use numerical integration methods as Gauss-Hermite formulae. These methods are known to be unstable (see [6, 25]) and require some preliminary analysis, as reparametrisation on which the efficiency of the method depend. More recently, Gibbs sampling [12] has been intensively used in many statistic fields. This iterative method generates an approximate sample from the posterior distribution if simulation from conditional distributions is easily attainable. But, this is not the case for Weibull distribution. Implementing Gibbs sampling need the incorporation of a Hasting-Metropolis step inside each Gibbs sampling iteration. And, this Hasting-Metropolis step can produce prohibitively slow convergence because most of the proposed moves are rejected, especially when there are a small amount of observed failure times. In this paper, we investigate two methods closely related to Importance Sampling methods that we sketch now.

In Monte Carlo integration via Importance Sampling [13], $I(\Phi)$ is rewritten as

$$I(\Phi) = \int \Phi(\theta)\frac{\pi(\theta \mid \boldsymbol{t})}{\rho(\theta)}\rho(\theta)\, d\theta$$

where ρ is a tractable density. And, $I(\Phi)$ is approximated by

$$I_M(\Phi) = \sum_{i=1}^{M} w_i \Phi(\widetilde{\theta}^i) / \sum_{i=1}^{M} w_i.$$

where $(\widetilde{\theta}^1, \ldots, \widetilde{\theta}^M)$ is an i.i.d. sample from ρ and $w_i \propto \pi(\widetilde{\theta}^i \mid \boldsymbol{t})/\rho(\widetilde{\theta}^i)$. Under mild conditions [13], $I_M(\Phi)$ converges to $I(\Phi)$ with probability one and follows approximatively a normal distribution $\mathcal{N}(I(\Phi), \sigma^2/M)$ where

$$\sigma^2 = \left\{ \int [\Phi - I(\Phi)]^2 \frac{[L(\boldsymbol{t};\theta)\pi(\theta)]^2}{\rho(\theta)} \, d\theta \right\} / \left\{ \int L(\boldsymbol{t};\theta)\pi(\theta) \, d\theta \right\}^2 .$$

A variant of Importance Sampling allowing to produce an i.i.d. sample of size ℓ following approximatively $\pi(\theta \mid \boldsymbol{t})$ is the Sampling Importance Resampling (SIR) algorithm [19]. The SIR algorithm consists in three steps.

1. `sample` $(\widetilde{\theta}^1, \ldots, \widetilde{\theta}^M)$ `i.i.d. from` ρ;
2. `calculate the importance weights` $w_i \propto \pi(\widetilde{\theta}^i \mid \boldsymbol{t})/\rho(\widetilde{\theta}^i)$ `and the probabilities` $p_i = w_i / \sum_{j=1}^{M} w_j$;
3. `draw` $(\theta^1, \ldots, \theta^\ell)$ `i.i.d. from the discrete distribution over` $(\widetilde{\theta}^1, \ldots, \widetilde{\theta}^M)$ `affecting mass` p_i `on` $\widetilde{\theta}^i$.

In practical situations, IS and SIR can appear to be fragile methods. If importance weights w_i are very variable then an enormous initial sample will be needed for adequate richness in the resampled sample. Thus, two practical conditions are required for good performances for both methods :

- the proposal density $\rho(\theta)$ has to be a reasonable approximation of the posterior density $\pi(\theta \mid \boldsymbol{t})$.
- $\rho(\theta)$ must have heavier tails than $\pi(\theta \mid \boldsymbol{t})$.

The main purpose of the two methods that we present now is precisely to propose a good proposal density ρ for the Importance Sampling algorithm.

3 The WLB and BRM algorithms

Introduced by Newton and Raftery([16], [17]), the Weighted Likelihood Bootstrap (WLB) is a bootstrap procedure for sampling from posterior distributions. This algorithm involves the following steps:

1. `Sample` M `weight vectors` $\widetilde{\omega}^j = (\widetilde{\omega}_1^j, \ldots, \widetilde{\omega}_N^j)$ `from some pre-chosen distribution` $\mathcal{P}(\omega_1, \ldots, \omega_N)$`, eventually power up the weights to some fixed` $\alpha > 0$. `(the defaults are uniform Dirichlet distribution and` $\alpha = 1$`.)`

2. `For each` $\widetilde{\omega}^j$ `maximize the weighted likelihood to obtain`

$$\widetilde{\theta}^j = \arg\max_{\theta} \left[\prod_{i=1}^{m} f(t_i \mid \theta)^{\widetilde{\omega}_i^j}\right] \left[\prod_{i=m+1}^{N} R(t_i \mid \theta)^{\widetilde{\omega}_i^j}\right]$$

   `solution of equations`

$$\frac{1}{\beta} + \frac{\sum_{i=1}^{m} \widetilde{\omega}_i^j \log(t_i)}{\sum_{i=1}^{m} \widetilde{\omega}_i^j} - \frac{\sum_{i=1}^{N} \widetilde{\omega}_i^j t_i^\beta \log(t_i)}{\sum_{i=1}^{N} \widetilde{\omega}_i^j t_i^\beta} = 0 \tag{3.2}$$

   `and`

$$\eta = \left[\frac{\sum_{i=1}^{N} \widetilde{\omega}_i^j t_i^\beta}{\sum_{i=1}^{m} \widetilde{\omega}_i^j}\right]^{\frac{1}{\beta}}.$$

3. `Use the sample of WLB parameters values to calculate a kernel density estimate` $\widehat{\rho}$ `at each` $\widetilde{\theta}^j$.

Estimating the kernel density is an important step. Following [17], we use the maximal smoothing principle of Terrel [23] with normal kernel in our numerical experiments reported in Section 5.

The WLB-SIR algorithm consists finally in resampling, with replacement, the $\{\widetilde{\theta}\}$ using importance weights $w_j \propto \pi(\widetilde{\theta}^j \mid \boldsymbol{t})/\widehat{\rho}(\widetilde{\theta}^j)$. The WLB-IS algorithm consists in performing the Importance Sampling algorithm with these importance weights.

Numerical experiments reported in [16] and [17] show good performances of WLB-SIR especially for intricated problems where MCMC algorithms are prohibitively slow. However, the WLB methodology has several disadvantages:

- The parameter α has been introduced to provide a proposal distribution $\widehat{\rho}$ having heavier tails than the target posterior distribution [3], but there is no convincing general guidelines for choosing α.

- The possibility that the fitting algorithm, used in maximizing the weighted likelihood, does not converge for some selections of weights and where parameters drift off to infinity is present in a small sample setting (see Section 5 for illustrations).

- To build an importance sampling density which approximate reasonably the posterior distribution $\pi(\theta \mid \boldsymbol{t})$, WLB takes no account of the prior information contained in $\pi(\theta \mid \boldsymbol{t})$.

The Bayesian Restoration Maximization (BRM) methodology, that we present now, does not suffer such disadvantages. It takes profit of the incomplete data structure of the problem to create artificial completed data by simulating the missing values according to conditional distribution computed with the parameter θ drawn from the prior distribution $\pi(\theta)$. More precisely, it consists in replacing steps 1 and 2 of WLB with:

- **(B)** `sample` θ `from the prior distribution` $\pi(\theta)$;

- **(R)** `sample the missing data from its conditional distribution knowing the observed sample and the parameter` θ `drawn in` **B**`-step`;

- **(M)** `obtain` $\widetilde{\theta}$ `maximizing the likelihood of the completed sample`.

In our context of right censored failure times, the missing data are the $(N-m)$ unknown failures times occurring after the censoring times $t_{m+1}, \ldots, t_N$. And, the B, R, and M steps of BRM are as follows :

(B) `sample` $\bar{\theta}$ `from` $\pi(\theta)$;

(R) `generate` $\widetilde{t}_i \sim \dfrac{f(\cdot \mid \bar{\theta})}{R(t_i \mid \bar{\theta})} \mathbb{1}_{\{]t_i, \infty[\}}, i = m+1, \ldots, N$;

(M) `obtain` $\widetilde{\theta} = \arg\max_{\theta} \left[\prod_{i=1}^{m} f(t_i;\theta)\right] \left[\prod_{i=m+1}^{N} f(\widetilde{t}_i;\theta)\right]$ `maximum likelihood estimate for the completed sample` $(t_1, \ldots, t_m, \widetilde{t}_{m+1}, \ldots, \widetilde{t}_N)$.

By repeating steps (B), (R) and (M) we get a sample $(\widetilde{\theta}^1, \dots, \widetilde{\theta}^M)$. We calculate a kernel density estimate $\widehat{\rho}$ at each $\widetilde{\theta}^j$ and conclude using IS or SIR algorithm with the importance sampling density $\widehat{\rho}$ as in the WLB methodology.

The main advantage of BRM over WLB, is that BRM takes account of the prior distribution π to construct the proposal distribution $\widehat{\rho}$. As a consequence, this proposal distribution can be expected to reasonably approximate the posterior distribution, and, by the way, there is no need to introduce an over dispersion parameter α and numerical incidents in the **M**-step are quite unlikely to occur. However, notice that the BRM methodology is limited to incomplete data models while WLB is not.

4 Designing the prior distributions

In this section we address the problem of designing prior distribution for parameters (β, η) taking account of the expert prior knowledge. Inspired by some authors [5], [10] and owing to our own experience [2], we assume that expert's prior knowledge about θ consists in:

- the shape parameter β is supposed to be in an interval $[\beta_\ell, \beta_r]$, and eventually a prior guess β_g is available,
- the scale parameter η is supposed to be in an interval $[\eta_\ell, \eta_r]$, and a prior guess η_g is available.

The prior density choosen for the shape parameter is the Beta density $\mathcal{B}(p,\, q,\, [\beta_\ell, \beta_r])$ on $[\beta_\ell, \beta_r]$:

$$\pi(\beta) \propto \frac{(\beta - \beta_\ell)^{p-1}(\beta_r - \beta)^{q-1}}{(\beta_r - \beta_\ell)^{p+q-1}} \mathbb{1}_{[\beta_\ell, \beta_r]}(\beta)$$

with mode, mean and variance

$$\mathbb{M}_\pi[\beta] = \beta_\ell + \frac{(p-1)(\beta_r - \beta_\ell)}{p+q-2} \text{ if } p,\ q > 1,$$

$$\mathbb{E}_\pi[\beta] = \beta_\ell + \frac{p(\beta_r - \beta_\ell)}{p+q},$$

$$\mathrm{Var}_{\pi}[\beta] = \frac{pq(\beta_r - \beta_\ell)^2}{(p+q+1)(p+q)^2}.$$

This choice is motivated by some interesting features of the Beta density. It is uniforme if $p = q = 1$ and unimodale if $p, q > 1$ with its mode in the right half (resp. middle, left half) of the interval if $q < p$ (resp. $p = q$, $q > p$). Moreover, $\mathbb{E}_\pi[\beta]$ depends on p and q by the mean of the ratio p/q, so we can increase the variance by decreasing p and q keeping the value of the mean constant.

As prior distribution for the scale parameter, we choose the Gamma density $\mathcal{G}\,(a, b)$

$$\pi(\eta) \propto \eta^{a-1} \exp\left(-\frac{\eta}{b}\right) \mathbb{1}_{]0,+\infty[}(\eta)$$

with mode, mean and variance

$$\mathbb{M}_\pi[\eta] = (a-1)b \text{ if } a \geq 1,$$

$$\mathbb{E}_\pi[\eta] = ab,$$

$$\mathrm{Var}_\pi[\eta] = ab^2.$$

The property of scale invariance of Gamma distribution (i.e. if $X \sim \mathcal{G}(a, b)$ then $cX \sim \mathcal{G}(a, bc)$ for any positive constant c) is convenient for modelling prior knowledge about a scale parameter.

To give values to the hyperparameters p, q, a and b, we propose the following procedure.

At first, we suppose that the central value β_g is available. Let β_m denote the middle of $[\beta_\ell, \beta_r]$.

1. Locate β_g in $[\beta_\ell, \beta_r]$:

    1.1 if $\beta_g \in [\beta_\ell, \beta_m]$, choose $q > p > 1$,

    1.2 if $\beta_g \in [\beta_m, \beta_r]$, choose $p > q > 1$,

    1.3 if β_g is in the vicinity of β_m, choose $p = q > 1$.

2. What β_g is representing ?

    2.1 if β_g stands for prior mean, we get the equation $\beta_g = \beta_\ell + \dfrac{p(\beta_r - \beta_\ell)}{p+q}$, 2.2 if β_g stands for prior mode, we get the equation $\beta_g = \beta_\ell + \dfrac{(p-1)(\beta_r - \beta_\ell)}{p+q-2}$,

  2.3 if the expert is unable to qualify β_g as mode or mean, use any of the two above equations with p and q large enough to merge mode and mean.

3. If the confidence level in β_g is strong, choose large p and q to have a small variance, else choose small p and q, but greater than 1, to have a large variance.

4. What η_g is representing ?

  4.1 if η_g stands for prior mean, we get the equation $b = \eta_g / a$,

  4.2 if η_g stands for prior mode, we get the equation $b = \eta_g/(a-1)$,

  4.3 if the expert is unable to qualify β_g as mode or mean, choose a and b to merge mean and mode.

5. Choose a and b to have $\pi(\eta)$ concentrated in $[\eta_\ell, \eta_r]$ with probability one approximately.

If β_g is not available, the first step is modified as follows:

- If the true shape parameter is more likely in the left half of $[\beta_\ell, \beta_r]$, set $\beta_g = (\beta_\ell + \beta_m)/2$ and choose $q > p > 1$.

- If it is in the right half of $[\beta_\ell, \beta_r]$, set $\beta_g = (\beta_m + \beta_r)/2$ and choose $p > q > 1$.

- If there is no information on its position in $[\beta_\ell, \beta_r]$, set $\beta_g = \beta_m$ and choose $p = q > 1$.

Indeed, our procedure is deliberately flexible, it aims to provide some guidelines to apprehend crude prior informations. It can easily be adapted to work in an empirical Bayesian perspective to take advantage of maximum likelihood estimates or any sample-based estimates.

5 Numerical experiments

We now present Monte Carlo simulations performed to compare the behavior of WLB and BRM with the maximum likelihood estimation in a small sample and highly censored setting. For those experiments, we randomly draw 50 replications of each of the four situations described below. We generated right censored samples of size $N = 25$ from a Weibull distribution $\mathcal{W}(\beta, \eta)$ with shape parameter β and scale parameter η. The scale parameter was $\eta = 100$ and we considered an unique censoring time $c = 40$. We considered four different shape parameter values $\beta = 0.5; 1.2; 2; 3$.

As prior information, for each of the four situations, we suppose that β and η are known to be in the intervals $[0.5,\ 3]$ and $[70,\ 170]$ respectively and that $\eta_g = 120$ is a prior guess for η. No prior guess was supposed for β. Following the procedure described in Section 4, the specified prior distribution are $\pi(\beta) = \mathcal{B}(1.5,\ 1.5,\ [0.5,\ 3])$ and $\pi(\eta) = \mathcal{G}(51.8,\ 2.3)$, the hyperparameters for this last distribution has been calculated using the normal approximation to have $\pi(\eta)$ concentrated on $[70,\ 170]$ with probability 0.99. The WLB algorithm has been applied with uniform Dirichlet as weight distributions and $\alpha = 1$, and for both algorithms, the maximal smoothing principle of Terrel [23] with normal kernel was used to calculate a kernel density estimate.

The average number, over the 50 replications, of uncensored observations (m in the notation) was 12 for $\mathcal{W}(0.5, 100)$, 7 for $\mathcal{W}(1.2, 100)$, 4 for $\mathcal{W}(2, 100)$ and 2 for $\mathcal{W}(3, 100)$.

Coupled with the IS algorithm, WLB and BRM, with $M = 5000$ iterations, give the results summarized in Table 1.

In Table 1, E (resp. Std) stands for average (resp. standard deviation) of the parameter estimations over the 50 replications. The first column gives the true parameter values. In the column ML, the maximum likelihood estimates are displayed.

Before discussing the results diplayed in Table 1, it is worth noting that with $\beta = 2$ and $\beta = 3$, two, respectively nine, samples among 50, doesn't contain uncensored failure times. Those samples has been omitted for maximum likelihood method and WLB-IS algorithm since we cannot maximize the weighted likelihood when no observed failure time is available. On the contrary, those samples have been included for the BRM algorithm.

Table 1: Comparison of the maximum likelihood, the WLB-IS and the BRM-IS parameter estimates for a right censored Weibull distribution.

	ML	WLB-IS	BRM-IS
	E (Std)	E (Std)	E (Std)
$\beta = 0.5$	0.583 (0.196)	0.628 (0.096)	0.593 (0.098)
$\eta = 100$	110.534 (112.237)	108.342 (8.244)	83.724 (21.108)
$\beta = 1.2$	1.810 (2.540)	1.648 (2.534)	1.327 (0.325)
$\eta = 100$	296.287 (987.475)	109.290 (14.517)	108.138 (6.505)
$\beta = 2$	2.994 (2.738)	2.733 (2.774)	1.898 (0.284)
$\eta = 100$	300.355 (1237.941)	109.030 (24.281)	116.157 (4.932)
$\beta = 3$	7.683 (20.783)	2.877 (0.180)	2.232 (0.176)
$\eta = 100$	341.242 (1288.989)	102.777 (28.502)	121.143 (3.898)

The maximum likelihood estimates are correct when the number of observed failure times is sufficient. The overestimation phenomena for the shape parameter is pronounced as soon as $\beta = 1.2$ and increases with the true β value. The important values of standard deviations are also notable. They are essentially due to "pathological" samples causing great variability in parameter estimation. For example, $(\widehat{\beta}, \widehat{\eta}) = (13.51,\ 46.52)$ and $(\widehat{\beta}, \widehat{\eta}) = (130.33,\ 40.99)$ have been obtained as estimates of true parameters $(\beta, \eta) = (2, 100)$ and $(\beta, \eta) = (3, 100)$ respectively.

Comparing WLB-IS and BRM-IS estimators, it appears that BRM provides more accurate results. For the two limit cases $\beta = 0.5$ and $\beta = 3$, WLB seems to perform better than BRM, especially for $\beta = 3$. In fact, this superiority is an artefact. Actually, only few values obtained by maximization of weighted likelihood in the WLB step lay in the interval $[\beta_\ell, \beta_r]$ near the boundaries. This is due to a numerical problem arising when maximizing the weighted likelihood. The optimisation algorithm at hand (Newton Raphson in our case) does not converge for some simulated vectors of weights $\widetilde{\omega}_i^j$. This lack of convergence typically occurs when the number of uncensored observations is very small as for $\beta = 3$ where 21 samples over 50 have been discarded from the WLB numerical experiment. Thus, the IS step is based on a few selected reliables estimates. On the other hand, the underestimation with BRM-IS with $\beta = 3$, is due to the very limited number of observed failures (2 failures in average) and to the fact that β prior mean is 1.75.

Finally, note that there is a little to choose between IS and SIR. From numerical experiments [2] not reported here, (WLB or BRM)-IS and (WLB or BRM)-SIR produce the same estimates.

In conclusion, we think that Bayesian inference is highly preferable to maximum likelihood inference for estimating the parameters of a Weibull distribution in a small sample and severely censored setting. In that context, both the WLB and BRM methodologies are powerful alternatives to MCMC algorithms which can be prohibitively slow. At last, BRM seems to be more accurate and less fragile than WLB since it takes account of the prior distribution when designing the Importance Sampling proposal distribution while WLB does not.

Acknowledgements

We thank A. Lannoy and D. Vasseur from Électricité de France for helpful discussion.

References

[1] J. Achcar and F. Louzada-Neto, "A Bayesian approach for accelerated life tests considering the Weibull distribution", *Comput. Statist.*, pp. 355–369, 1992.

[2] M. Bacha, *Statistical Inference for Life Span Models and Applications (in french)*, PhD thesis, Rouen University, France, 1996.

[3] M. Bacha and G. Celeux, "Comment on:"approximate Bayesian inference with the weighted likelihood bootstrap"", *J. Roy. Statist. Soc. (Ser. B)*, vol. 56, pp. 33–34, 1994.

[4] L. Bain and M. Engelhardt, *Statistical Analysis of Reliability and Life Testing Models*, Marcel Dekker, 2 edition, 1991.

[5] J. Berger, "Bayesian reliability for Weibull data", in *XXII* $^{\text{es}}$ *Journées de Statistiques. Tours*, 1990, pp. 15–17.

[6] J. Bernardo and A. Smith, *Bayesian Theory*, John Wiley, 1994.

[7] D. Bogdanoff and D. Pierce, "Bayes-fiducial inference for the Weibull distribution", *J. Amer. Statist. Assoc.*, vol. 68, pp. 659–664, 1973.

[8] R. Calabria, M. Guida, and G. Pulcini, "A Bayes procedure for estimation of current system reliability", *IEEE Transactions on Reliability*, vol. 41, pp. 616–620, 1992.

[9] W. Ellis and V. Tummala, "Minimum expected loss estimators of the shape & scale parameters of the Weibull distribution", *IEEE Transactions on Reliability*, vol. R-35, pp. 212–215, 1986.

[10] P. Erto, "New practical Bayes estimators for the 2-parameter Weibull distribution", *IEEE Transactions on Reliability*, vol. R-31, pp. 194–197, 1982.

[11] I. Evans and A. Nigm, "Bayesian 1-sample prediction for the 2-parameter Weibull distribution", *IEEE Transactions on Reliability*, vol. R-29, pp. 410–413, 1980.

[12] A. Gelfand and A. Smith, "Sampling based approaches to calculating marginal densities", *J. Amer. Statist. Assoc.*, vol. 85, pp. 398–409, 1990.

[13] J. Geweke, "Bayesian inference in econometric models using Monte Carlo integration", *Econometrica*, vol. 57, pp. 1317–1339, 1989.

[14] H. Martz and R. Waller, *Bayesian Reliability Analysis*, John Wiley, 1982.

[15] J. Naylor and R. Smith, "A comparison of maximum likelihood and Bayesian estimators for the three-parameter Weibull distribution", *Appli. Statist.*, vol. 36, pp. 358–369, 1987.

[16] M. Newton, *The Weighted Likelihood Bootstrap and an Algorithm for Prepivoting*, PhD thesis, University of Washington, 1991.

[17] M. Newton and A. Raftery, "Approximate Bayesian inference with the Weighted Likelihood Bootstrap", *J. Roy. Statist. Soc. (Ser. B)*, vol. 56, pp. 3–48, 1994.

[18] P. Reilly, "The numerical computation of posterior distributions in Bayesian statistical inference", *Appli. Statist.*, vol. 25, pp. 201–209, 1976.

[19] D. Rubin, "Comment on:"the calculation of posterior distributions by data augmentation"", *J. Amer. Statist. Assoc.*, vol. 82, pp. 543–546, 1987.

[20] N. Singpurwalla, "An interactive pc-based procedure for reliability assessment incorporating expert opinion and survival data", *J. Amer. Statist. Assoc.*, vol. 83, pp. 43–51, 1988.

[21] N. Singpurwalla and M. Song, "The analysis of Weibull lifetime data incorporating expert opinion", in *Probability and Bayesian Statistics*, R. Viertl, Ed. Plenum Publishing Corporation, 1987.

[22] S. Sinha and J. Sloan, "Bayes estimation of the parameters and reliability function of the 3-parameter Weibull distribution", *IEEE Transactions on Reliability*, vol. 37, pp. 364–368, 1988.

[23] G. Terrel, "The maximal smoothing principle in density estimation", *J. Amer. Statist. Assoc.*, vol. 85, pp. 470–477, 1990.

[24] L. Tierney and J. Kadane, "Accurate approximations for posterior moments and marginal densities", *J. Amer. Statist. Assoc.*, vol. 81, pp. 82–86, 1986.

[25] R. Wolpert, "Monte Carlo integration in Bayesian statistical analysis", in *Statistical Multiple Integration*, N. Flournoy and R. Tsutakawa, Eds., Philadelphia, 1991, Contemporary Mathematics, 115, Proceedings of a Joint Summer Research Conference held at Humboldt University, June, 1989.

Unité de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unité de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 655, avenue de l' Europe, 38330 MONTBONNOT ST MARTIN
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
ISSN 0249-6399