Bayesian estimation of Weibull mixture in heavily censored data setting

Florence Ducros, Patrick Pamphile

Bayesian estimation of Weibull mixture in heavily censored data setting

Florence Ducros[a,b,c,*], Patrick Pamphile[a,b]

[a]*Laboratoire de Mathématiques d'Orsay, Université Paris-Sud, 91405 Orsay Cedex 9, France*
[b]*Select Inria, 91405 Orsay, France*
[c]*Nexter System, 11 Allée des Marronniers, 78022 Versailles Cedex, France*

Abstract

Lifetime data collected from a fleet of vehicles or, more broadly, park of systems are generally non-homogeneous and heavily censored. Indeed, system lifetime can be affected by the variability of production conditions and usage conditions. Most of the time, this variability is unobserved, but has to be taken into account for reliability or warranty cost analysis. This research proposes a two-component Weibull mixture model for modelling unobserved heterogeneity in in heavily censored lifetime data collection. Performance of classical estimation methods (maximum of likelihood, EM, full Bayes and MCMC) are significantly reduced due to the high number of parameters and the heavy censoring. Therefore, a Bayesian bootstrap method, called Bayesian Restauration Maximisation, is used. Sampling from the posterior distribution is obtained thanks to an importance sampling technique. Simulation results showed that, even with heavy censoring, BRM is effective both in term of estimates precision and computation times. The prior elicitation, sensibility analysis and comparaisons with EM are discussed. Finally, a real data set is analyzed to illustrate the application of the method.

Keywords: Reliability analysis, Warranty cost, Weibull mixture, Non-homogeneous data, Heavily censored data, Bayesian estimation,

---

[*]Corresponding author
Email addresses: `florence.ducros@nexter-group.fr` (Florence Ducros), `patrick.pamphile@u-psud.fr` (Patrick Pamphile)

1. Introduction

In reliability or warranty analysis, engineers must often deal with components which have more than one failure mode. Usually, this non-homogeneity coincides with presence of compliant/non-compliant components, early/random or wear out/random failures, and so forth. For example, in the automotive industry, components are produced by different suppliers. In this case, component reliability may depend on the supplier, but also on the car operating environment, (e.g., temperature, humidity, roads, etc.), the usage mode and the usage intensity. Generally, these various sources of non-homogeneity are not controlled or not observed. But neglect existing non-homogeneity, can lead to errors and misconceptions in the analysis. In manufacturing, reliability has a serious impact on the warranty servicing cost or maintenance cost. For safety studying, conclusions of diagnostic tests can be severely misguided, and so on. Then, mixture of distributions are usually used to modelling failure times from non-homogeneous data [1, 2].

Since more than five decades, Weibull distributions is extensively used for modelling life data in medical or industrial applications. Weibull distributions exhibit decreasing, constant or increasing hazard function which makes them suitable for modelling complex failure data. See for example [3, 4] and references therein. Hence, two-component Weibull mixture is a highly relevant model to capture latent heterogeneity for a large majority operating reliability analysis.

Furthermore, the other issue raised in reliability analysis is the large number of censored failure times. For example, in warranty data collections, during the time interval over which data are collected, components are put into service at different times : failure time is censored for the large number of components still operating. Percentage of censored failure times may be greater than 70% and lifetime inference is therefore a particularly difficult challenge.

Different methods are used for fitting a parametric model to data, but they have substantial difficulties for Weibull mixtures in the setting of heavily censored data. Graphical methods can provide a quick but crude estimation [5].

3

The maximum likelihood (ML) estimation is the most preferred method due to its desirable asymptotic properties. As the mixing cause is missing, ML is not suitable here. Expectation-maximization (EM) algorithm and is stochastic variant, S-EM, are iterative methods to obtain ML estimate when data are missing [6, 7]. But they can fail due to the heavy censoring. As noted in [8], full Bayes can address the issue of ML methods arising from limited data, but requires complexe integrations that make it computationally prohibitive in the framework of mixture models. The higher number of parameter combined with heavy censoring rate makes that Markov chain Monte Carlo algorithms appear to be deceptive in terms of computation time for Bayesian inference [9, 10].

In this paper, The Bayesian restoration maximization (BRM) algorithm is used to estimate the parameters of a two-component Weibull mixture in heavily censored data setting. BRM algorithm has been used for single Weibull inference in [11]. According to [12], it is a Bayesian bootstrap method which combines the prior information and importance sampling technique, to obtain a sample of the model parameters distributed similarly to the posterior distribution. In this paper, BRM was adapted to Weibull mixture inference. The particular issue of the prior elicitation was discussed. Simulation results showed that BRM is effective, both in term of estimates precision and computation times.

The paper is organized as follows. The Weibull mixture and notation are presented in Section 2. Classical estimation methods are set out in Section 3 In Section 4, BRM algorithm is extended to Weibull mixture. Calibration of the priors hyperparameters, sensibility analysis and performance comparisons with the S-EM algorithm are discussed in Section 5. A real data set is also analyzed to illustrate the application of the method. Finally, Section 6 gives a conclusion.

2. Mixture of Weibull distributions

Weibull distributions have been extensively used for modelling life data in medical or industrial applications. This is clearly illustrated by the large number of references on it. A review of the Weibull distribution, in many of its aspects,

4

can be found in [3] and [13]. In particular, Weibull distributions exhibit decreasing, constant or increasing hazard function which makes them suitable for modelling complex failure data.

2.1. Weibull distributions

A two-component mixture was used where both component belongs to the family of two-parameter Weibull distributions. For $x \geq 0$, the probability density function is

$$f_{MIX}(x|\theta) = \sum_{k=1}^{2} \alpha_k f_W(x|\beta_k, \eta_k), \tag{1}$$

where for $k = 1, 2$

- $f_W(x|\beta_k, \eta_k)$ is the Weibull probability density function with shape parameter $\beta_k > 0$ and scale parameter $\eta_k > 0$,

$$f_W(x|\beta_k, \eta_k) = \frac{\beta_k}{\eta_k} \left(\frac{x}{\eta_k}\right)^{\beta_k - 1} e^{-\left(\frac{x}{\eta_k}\right)^{\beta_k}}; \tag{2}$$

- $\alpha_k > 0$ is the mixing proportion of the k-th component, with $\alpha_1 + \alpha_2 = 1$;

Let's denote $\theta = (\alpha_1, \beta_1, \eta_1, \beta_2, \eta_2)$, the set of model parameters.

2.2. Data and notation

The available information in the data are :

$$
\begin{array}{ll}
X_i & \text{is a observed failure time, } i = 1, \ldots, n_o; \\
Y_i & \text{is a censored failure time, } i = 1, \ldots, n_c; \\
Z_{oik} & \text{indicates the component of failure time,} \\
& = \begin{cases} 1 & \text{if } X_i \text{ coming from the } k\text{-th component;} \\ 0 & \text{otherwise;} \end{cases} \\
Z_{cik} & \text{indicates the component of censored failure time,} \\
& = \begin{cases} 1 & \text{if } Y_i \text{ coming from the } k\text{-th component;} \\ 0 & \text{otherwise;} \end{cases}
\end{array}
$$

where n_o is the number of observed failures, n_c the number of censored failures and $n = n_o + n_c$, the sample size.

For a lifetime X, we get

$$X|z_{oik} \sim f_W(x|\beta_k, \eta_k), \tag{3}$$

and

$$X|(y_i, z_{cik}) \sim \left[\frac{f_W(x|\beta_k, \eta_k)}{R_W(y_i|\beta_k, \eta_k)} \mathbf{1}(x > y_i) \right]^{z_{cik}}. \tag{4}$$

Unfortunately, (Z_{oik}) and (Z_{cik}) are not observed : the component label, from which X_i or Y_i is coming, is unknown.

Both deterministic and random censoring are considered. In warranty data set, for example, during the time interval over which data are collected, components are put into service at different times. Then, components still under warranty have a failure time randomly right censored.

3. Fitting the model

3.1. Weibull quantile-quantile plot

Weibull quantile-quantile plot (QQ-plot) can provide a quick and simply confirmatory method for data non-homogeneity [14]:

- for standard single Weibull distribution, the Weibull QQ-plot has a straight line shape;

- for two-components Weibull mixture, the Weibull QQ-plot has a single inflection point (S-shaped) with parallel asymptotes.

According to [5], a roughly estimate of the two-component Weibull parameters, $\widetilde{\theta}_0$, can be obtained from the Weibull QQ-plot. These estimate are non precise, nor robust, but nonetheless can be used as starting values for alternative estimation algorithms.

3.2. Maximum likelihood estimation

The maximum likelihood (ML) estimate is

$$\widehat{\theta}^{MLE} = \arg\max_{\theta} \ln L((x,y,z)_1^n|\theta) \tag{5}$$

where

$$L((x,y,z)_1^n|\theta) = \prod_{i=1}^{n_o} \prod_{k=1}^{2} \left[\alpha_k \left(f_W(x_i\,|\beta_k,\eta_k)\right)\right]^{z_{oik}} \cdot \prod_{i=1}^{n_c} \prod_{k=1}^{2} \left[\alpha_k \left(R_W(y_i\,|\beta_k,\eta_k)\right)\right]^{z_{cik}}. \tag{6}$$

The likelihood contains all information in the data that is relevant to estimate parameters. ML estimation is suitable for complete data, but turn to be imprecise for small and incomplete data. Particulary for Weibull distribution which has not closed form of the shape parameter estimate [13].

3.3. EM and S-EM algorithms

Expectation-maximization (EM) algorithm is a ML estimation for data containing missing values [15]. Therefore, EM algorithm is suitable, not only for censored data, but also for mixture of distributions for which the mixing causes are missing [6, 7]. Here, the missing data are

- censored failure times, Y_i;

- component labels, Z_{ik}.

The EM algorithm is based on a completion of the missing data from the conditional distribution $\pi(z,y\,|\,x\,;\,\theta^{(r)})$, and followed by a maximization of the conditional expectation of the complete log-likelihood;

- **initialization:** choose $\theta^{(1)}$;

- **iteration:** for $r = 1,\dots,B-1$

    1. **expectation:** with $\theta^{(r)}$ compute the conditional expectation of the complete likelihood,

    $$\mathbf{Q}(\theta|\theta^{(r)}) = \sum_z \int_y \ln L(x,y,z)\,|\,\theta)\pi(z,y\,|\,x\,;\,\theta^{(r)})dy. \tag{7}$$

2. **maximization:** compute

$$\theta^{(r+1)} = \arg\max_{\theta} \mathbf{Q}(\theta|\theta^{(r)}).$$

- **estimation:** hence, from $(\theta^{(1)}, \ldots, \theta^{(B)})$ calculate

$$\widehat{\theta}^{SEM} = \frac{1}{B-M} \sum_{i=M+1}^{B} \theta^{(r)} \qquad \text{for } M \leq B. \tag{8}$$

At each step the log-likelihood $\ln L((x, y, z)_1^n|\theta)$ increases.

For highly censored or small data set, EM converges slowly and towards poor local maximizers. Stochastic EM (S-EM) resolves the main difficulty of EM that is the calculation of $\mathbf{Q}(\theta|\theta^{(r)})$, see [6]. The expected complete log-likelihood is estimated by simulating the missing data as follows:

- **simulations:**

    - **component labels:** $(\widetilde{z}_{oik})$ and $(\widetilde{z}_{cik})$ are sampled according to the conditionally distributions:

$$p_{oik}(\theta^r) = \mathbf{E}(Z_{oik}|x_i \, ; \, \theta^{(r)}) = \frac{\alpha_k^{(r)} \cdot f_W(x_i \, | \, \beta_k^{(r)}, \eta_k^{(r)})}{\displaystyle\sum_{j=1}^{\kappa} \alpha_j^{(r)} \cdot f_W(x_i \, | \, \beta_j^{(r)}, \eta_j^{(r)})}, \tag{9}$$

$$p_{cik}(\theta^r) = \mathbf{E}(Z_{cik}|y_i \, ; \, \theta^{(r)}) = \frac{\alpha_k^{(r)} \cdot R_W(y_i \, | \, \beta_k^{(r)}, \eta_k^{(r)})}{\displaystyle\sum_{j=1}^{\kappa} \alpha_j^{(r)} \cdot R_W(y_j \, | \, \beta_j^{(r)}, \eta_j^{(r)})}; \tag{10}$$

    - **censored failure times:** from the censored failure time y_i and label indice $\widetilde{z}_{cik}$, a failure time $\widetilde{x}_i$ is sampled according to (4).

From the completed sample of lifetimes $(x_1, \ldots, x_{n_o}, \widetilde{x}_1, \ldots, \widetilde{x}_{n_c})$ and labels $(\widetilde{z}_i)$, the maximization step is therefore a ML estimation with completed data for each Weibull component, see [13]. The ML estimates of mixing probabilities are

$$\widehat{\alpha}_k^{(r+1)} = \frac{\displaystyle\sum_{i=1}^{n} p_{oik}(\theta^{(r)}) + p_{cik}(\theta^{(r)})}{n_0 + n_c}. \tag{11}$$

In the context of heavy censoring and several parameters, S-EM algorithm converges very slowly too.

3.4. Bayesian estimation

Bayesian estimation is an alternative to ML estimation [16]. It consists of incorporating prior information, to produce, with the data, a posterior distribution of the parameter. This prior information, noted $\pi(\theta)$, may be data from previous studies or expert advices which quantifies uncertainty about the parameters. The posterior distribution $\pi(\theta \,|\, (x, y, z)_1^n)$, is obtained using Bayes' theorem,

$$\pi(\theta \,|\, (x, y, z)_1^n) \propto L((x, y, z)_1^n \,|\, \theta)\pi(\theta). \tag{12}$$

Therefore, a Bayesian estimate is

$$\widehat{\theta}^{Bayes} = \mathbf{E}[\theta \,|\, (x, y, z)_1^n)] = \int_\theta \theta \, \pi(\theta \,|(x, y, z)_1^n). \tag{13}$$

In case of Weibull mixture, due to the large number of parameter, direct Bayes estimation is quite difficult, see [8] for instance. The posterior distribution is not in closed form and numerical integration is very time consuming. Even Markov chain Monte Carlo methods appear to be deceptive in the context of Weibull mixture [10]. The Bayesian restoration maximization (BRM) algorithm has been used to estimate parameters of a basic Weibull distribution in [11]. Here, BRM was used to Weibull mixture inferences. The key feature of BRM is that the expectation (13) is computed using sampling importance resampling technique (SIR) [17].

4. Bayesian Restoration Maximization

4.1. BRM method

The integral (13) was calculated using importance sampling [18]. This technique exploits the identity

$$\widehat{\theta}^{Bayes} = \int_\theta \theta \, \pi(\theta \,|(x, y, z)_1^n) = \int_\theta \theta \, \frac{\pi(\theta \,|(x, y, z)_1^n)}{\rho(\theta)} \rho(\theta), \tag{14}$$

where the support of ρ includes the support of $\pi(\theta \,|(x, y, z)_1^n)$. Therefore (14) can be estimated by Monte Carlo :

1. $\widetilde{\theta}^{(1)}, \ldots \widetilde{\theta}^{(B)}$ are sampled according to a proposal distribution ρ;

2. hence, θ is estimated by

$$\widehat{\theta}^{BRM} = \frac{1}{B} \sum_{r=1}^{B} \widetilde{\theta}^{(r)} \, w_r, \tag{15}$$

with the weights

$$w_r = \frac{\pi(\widetilde{\theta}^{(r)} \mid (x, y, z)_1^n)}{\rho(\widetilde{\theta}^{(r)})}. \tag{16}$$

4.2. Proposal distribution

Now, we have to select the sampling distribution ρ. It must be a counterfeit of the posterior distribution easier to obtain. From Bayes' Theorem

$$\pi(\theta \mid (x, y, z)_1^n) \propto L((x, y, z)_1^n | \theta) \cdot \pi(\theta).$$

The likelihood, $L((x, y, z)_1^n | \theta)$, contains information on the data. For its part, the prior, $\pi(\theta)$, contains information on the parameter. Therefore, the distribution ρ was obtained as follows. Let B denotes the number of runs for BRM,

1. **(B) prior sampling :** $\widetilde{\theta}^{(1)}, \ldots, \widetilde{\theta}^{(B)}$ are sampled according to $\pi(\theta)$;

2. **(R) missing data restoration :** for $r = 1, \ldots, B-1$, using $\widetilde{\theta}^{(r)}$ as starting value, the missing variables Z_{oik}, Z_{cik} and Y_i are simulated according to (9-10) and (4).

3. **(M) maximization :** ML estimates $(\widehat{\theta}^{(r)})$ are obtained from the completed data;

4. **importance sampling :** the distribution of $\widehat{\theta}^{(1)}, \ldots, \widehat{\theta}^{(B)}$ is expected to be related to the posterior distribution. It cannot be obtained in closed form. Its kernel density estimate is

$$\widehat{\rho}(\theta) = \frac{1}{B \cdot h_\theta} \sum_{r=1}^{B} K\left(\frac{\theta - \widehat{\theta}^{(r)}}{h_\theta}\right), \tag{17}$$

where K is a Gaussian kernel.

The window size, h_θ, minimizing the mean integrated square error is given in [19].

10

4.3. Resampling

It is worth mentioning that the weights can be uneven distributed weights, and several may be nil or almost nil. This occurs when prior distribution and likelihood are separated. This issue can be resolved by a resampling step, see [17]. For $M < B$, $(\widetilde{\widetilde{\theta}})_1^M$ are resampled from $(\widetilde{\theta})_1^B$ according to $(w_i)_1^B$. Then, credibility interval can be obtain from $2, 5\%$-quantile and $95, 75\%$-quantile of $(\widetilde{\widetilde{\theta}}^{(r)})$. Resampling step eliminates $\widetilde{\theta}^{(r)}$ with low weight, and duplicates those with high weight. According to [17], very large ratio B/M may be require for suitable performance of resampling. Here the ratio was set at $M = B/30$.

The mixture model is invariant to permutation of the component labels. Hence, the posterior distribution can be symmetric if a weak prior is used, that involves bad inference. This issue, called label switching, is dealt with a k-means classification, post-resampling, on $\widetilde{\theta}^{(r)}$, see [20].

4.4. Prior elicitation

The choice of prior is the key problem in Bayes estimation [16]. One can distinguish informative/non-informative prior distribution reflecting the fact that prior information is more or less vague. The Bayes estimate (13) shrinks the ML estimate (5) toward the prior mean, according to how informative the prior is or how large the sample size is. Therefore, non-informative prior is useful when it is desired to let the data speak for themselves, without being overly influenced by the prior. The richer the family of prior distribution is, the more capture of information about the parameter is. On the other hand, the choice of prior is constrained by the tractability of the posterior distribution. For a given sampling distribution $f(x \,|\, \theta)$, a conjugate prior distribution $\pi(\theta)$ is one for which the prior distribution and the posterior are in the same distribution family. They are useful to provided tractable posterior distribution. Unfortunately, no natural conjugate prior distribution exists for Weibull distribution when both the shape and scale parameters are unknown [21]. Independent priors for (α_1), (β_1, η_1) and (β_2, η_2) was used.

Shape parameter prior.

Thanks to its shape parameter, β, Weibull distributions have the ability to assume the characteristics of many different types of lifetimes. This has made it extremely popular to fit data from various fields, see for instance [22, 23] :

- $\beta < 1$ indicates early failure, with L-shaped density;

- $\beta = 1$ indicates random failure, with exponential density;

- $1 < \beta \leq 3$ indicates wearout failure, with positively skewed density;

- $3 < \beta \leq 4$ indicates wearout failure, with bell-shaped density;

- for $\beta > 4$, the Weibull distribution is similar to an extreme value distribution.

A beta prior, symmetric on $[b^{min}; b^{max}]$, has been chosen for both (β_1, β_2)

$$\beta \sim (b^{max} - b^{min}) \, \mathcal{B}(b^{shape}, b^{shape}) \, + \, b^{min} \tag{18}$$

and

$$\pi(\beta) = (b^{max} - b^{min}) \frac{\Gamma(2b^{shape})}{\Gamma^2(b^{shape})} \beta^{b^{shape}-1}(1-\beta)^{b^{shape}-1} \, \mathbf{1}_{[0;1]} \left(\frac{\beta - b^{min}}{b^{max} - b^{min}} \right), \tag{19}$$

where

$$\Gamma(x) = \int_0^{+\infty} t^x \mathrm{e}^{-t} dt.$$

Hyperparameters (b^{min}, b^{max}) reflect engineer expertise on parameters (β_k). They shall be fixed in accordance to the field failure data. Hyperparameter b^{shape} reflects uncertainty about (β_k). Depending on b^{shape}, the beta distribution can be informative or non informative.

According to [24], (b^{min}, b^{max}) were chosen as follows:

$$b^{min} = 0.5 \quad , \quad b^{max} = 4. \tag{20}$$

A weakly informative prior has been chosen by setting $b^{shape} = 1.1$.

Scale parameter prior.

When β_k is known, the generalised inversed gamma distribution is the conjugate prior of the Weibull with parameter η_k, see [25]:

$$\pi(\eta \,|\, \beta_k, e^{shape}, e^{scale}) = \frac{\beta_k \left(e^{scale}\right)^{\beta_k\, e^{shape}}}{\Gamma(e^{shape})} \, \eta^{-e^{shape}\, \beta_k - 1} \, \mathrm{e}^{-\left(\frac{b^{shape}}{\eta}\right)^{\beta_k}}, \quad (21)$$

where $(e^{shape}, e_1^{scale}, e_2^{scale})$ are hyperparameters. A natural approach to compute $(e_1^{scale}, e_2^{scale})$ is to equate the expected value of the inverted gamma distribution to graphical estimation from the Weibull QQ-plot, $\widetilde{\eta}_k$, discussed in Section 3:

$$e_k^{scale} \frac{\Gamma\left(e^{shape} - \frac{1}{\beta_k}\right)}{\Gamma(e^{shape})} = \widetilde{\eta}_k. \quad (22)$$

This starting value not need to be sharp. The hyperparameter, e^{shape}, was chosen large to specify a weak informative prior, with $e^{shape} > \dfrac{1}{\beta_k}$ according to (22).

Mixing parameter prior.

The prior on the mixing parameter, α_1, represents the lake of information on α_1 on interval $[0; 1]$. A beta prior on $[a^{min}; a^{max}]$ was chosen:

$$\pi(\alpha) = (a^{max} - a^{min}) \frac{\Gamma(2a^{shape})}{\Gamma^2(a^{shape})} \, \alpha^{a^{shape} - 1} (1 - \alpha)^{a^{shape} - 1} \, \mathbf{1}_{[0;1]}\left(\frac{\alpha - a^{min}}{a^{max} - a^{min}}\right). \quad (23)$$

Hyperparameters (a^{min}, a^{max}) represent beliefs about α_1. They shall be fixed according to information on α_1. Hyperparameter a^{shape} reflect uncertainty about α_1. A weakly informative prior was chosen by setting $a^{shape} = 1.1$.

13

5. Simulations and results

Simulations were carried out, at first, to calibrate the hyperparameter, and next, to compared BRM and S-EM estimations. The data were simulated from the following Weibull mixture.

α_1	β_1	η_1	β_2	η_2
0.3	1.5	50	3	200

Table 1: Simulated data parameters

Various sample sizes and censoring rates were used.

First and foremost, the Weibull QQ-plot allowed to confirm graphicaly the presence of a mixture. According to [5], Weibull QQ-plot also allowed to obtain roughly estimate for α_1, η_1 and η_2, see Figure 1.

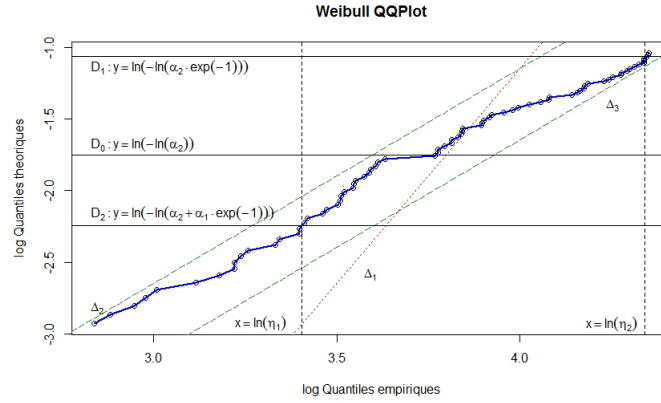Figure 1: Weibull QQ-plot. **Graphical test of mixture:** for two-component mixture of Weibull, the scatterplot has a S shape; as $x \to +\infty$ the Weibull QQ-plot asymptote is Δ_3 of slope $1/\beta_2$; as $x \to -\infty$ the Weibull QQ-plot asymptote is Δ_2 of slope of Δ_2; the line Δ_1 is of slope $1/\beta_1$. **Graphical estimation:** parameters α, η_1, η_2 can be roughly estimated, see [5]; Line D_0 highlights the inflexion point of the S curve, and allows to estimate α_2; Intersection of D_1 (respectively D_2) with the curve allows to estimate η_1 (respectively η_2).

The hyperparameters calibration was experimented from 500 simulated datasets with a sample size $n = 100$ and a right censoring level at 70%. The BRM estimation was applied to each dataset, with 1 000 runs. Hence, the average bias and the root mean square error (RMSE) of the mid-life reliability have been computed to assess the estimation performance:

$$\text{bias} \; = \frac{1}{500} \sum_{i=1}^{500} (\widehat{R} - R), \qquad \text{RMSE} \; = \sqrt{\frac{1}{500} \sum_{i=1}^{500} (\widehat{R} - R)^2}.$$

As expected, informative prior improved the performance: see Table 2 for the prior support of α, and Table3 for scale hyperparameter prior of η. Not sharply anticipated value of $(\widetilde{\eta}_k)$ provided relatively small bias and RMSE for $\eta^{shape} = 5$, see Table 4.

$[\mathbf{a_{min}} \, ; \, \mathbf{a_{max}}]$	$[0; 1.0]$	$[0; 0.5]$
average bias	-0.0212	-0.0016
RMSE	0.0725	0.0546

Table 2: α_1 **prior hyperparameters, $[\mathbf{a_{min}}, \mathbf{a_{max}}]$ calibration.** Mid-life reliability estimate for 500 datasets, with 1 000 runs and sample size $n = 100$; Diffuse prior $[0 \, ; \, 1]$ and more informative prior $[0 \, ; \, 0.5]$.

e^{shape}	3	5	10	20
average bias	-0.0113	-0.0016	0.006,8	0.0053
RMSE	0.0701	0.0546	0.0394	0.0336

Table 3: η **prior hyperparameters, e^{shape} calibration.** Mid-life reliability BRM estimate for 100 datasets, with 1 000 runs and sample size $n = 100$. Generalized inverse gamma prior with several values of the shape hyperparameter: the higher e^{shape} is, the more the prior is informative

210

15

	$\widetilde{\eta}$	$0.5\eta_0$	η_0	$1.5\eta_0$
$e^{shape} = 5$	average bias	-0.1721	-0.0016	0.0617
	RMSE	0.1407	0.0546	0.0523

	$\widetilde{\eta}$	$0.5\eta_0$	η_0	$1.5\eta_0$
$e^{shape} = 20$	average bias	-0.2244	0.0053	0.0644
	RMSE	0.1437	0.0336	0.0441

Table 4: η **prior hyperparameters, e_k^{scale} calibration.** Mid-life reliability BRM estimation for 500 datasets, with 1 000 runs and sample size $n = 100$. Generalized inverse gamma prior with several scale hyperparameter : $\widetilde{\eta} \ll \eta_0$, $\widetilde{\eta} \approx \eta_0$ or $\widetilde{\eta} \gg \eta_0$. To the left, a weak informative prior with $e^{shape} = 5$. To the right, more informative prior with $e^{shape} = 20$.

5.2. BRM number of runs

The number of runs is proportional to the computing time. The figure 5 shows that the RMSE decreases slightly with the number of runs.

BRM number of runs	500	1 000	6 000
average bias	-0.0030	-0.0016	0.0032
RMSE	0.0544	0.0546	0.0512

Table 5: BRM runs number calibration. Mid-life reliability BRM estimate for 500 datasets, with sample size $n = 100$ and censoring rate at 70%; runs number as 500, 1 000, or 6 000.

5.3. Sensitivity analysis

The sample size and the censoring rate are two inputs that influence the estimation. As can be seen from Table 6 and Table 7, BRM is still efficient even for small sample heavily censored.

sample size	100	200	600
average bias	-0.0016	0.0088	0.0125
RMSE	0.0546	0.0432	0.0631

Table 6: Sample size analysis. Mid-life reliability BRM estimation for 500 datasets, with 1 000 runs, 70% censoring and sample size $n = 100, 200$ and 600.

censoring	70%	80%	90%
average bias	0.0125	-0.0200	-0.0581
RMSE	0.0631	0.0828	0.0950

Table 7: Censoring. Mid-life reliability BRM estimation for 500 datasets, with 1 000 runs, sample size $n = 600$ and 70%, 80% and 90% censoring.

5.4. Comparisons

BRM and S-EM estimates have been compared from 500 datasets in several configuration : several sample size, deterministic or random right censoring and various censoring. As illustrated in figures 2 and 3, BRM estimation is more efficient than S-EM, in terms of biais and RMSE, and this for all configurations.

17

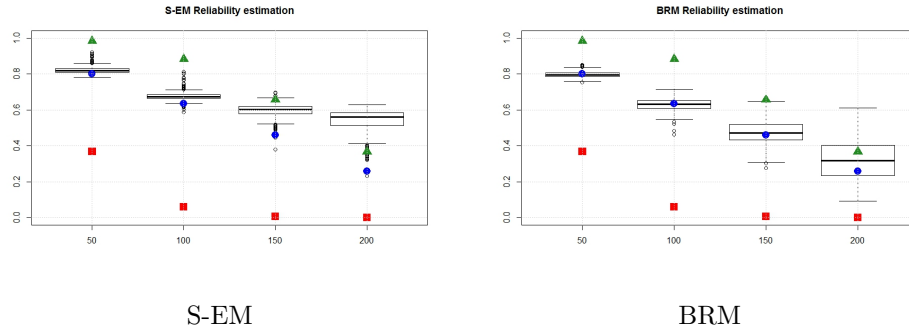S-EM   BRM

Figure 2: SEM-P and BRM estimations. Reliability for small, mid and long-term ; boxplots for 500 dataset, with $n = 600$ and 70% censoring ; green triangle denotes the component 1 reliability ; red box, the component 2 reliability ; blue circle the mixture reliability.
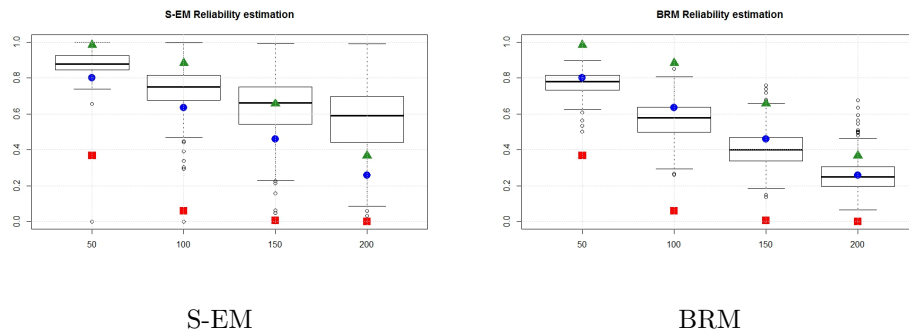

S-EM   BRM

Figure 3: SEM-P and BRM estimations. Reliability for small, mid and long-term ; boxplots for 500 dataset, with $n = 600$ and 90% censoring ; green triangle denotes the component 1 reliability ; red box, the component 2 reliability ; blue circle the mixture reliability.

In real applications, the mixture depicts the presence, in the used parts, of a small amount of substandard parts with short lifetime. This small proportion of "weak parts" plays a significant role in reliability for small-term, continues to impact the reliability for mid-term, and finally can be neglected for long-term. In this context, BRM still efficient for small and mid-term.

18

5.5. Real data

One real data set has been analyzed for illustration. It consist in 729 auto-
motive parts which have been put into service during several years. At the time
of this study, 74% were still operating. Therefore it is a random right censoring
context. In order to anticipate warranty costs, we need to estimate the reliabil-
ity in the short, medium and long term. Histogram and Weibull QQ-plot have
showed the presence a mixture. It can be costly not to take into account this
heterogeneity. A two-component Weibull mixture was used. The parameters
were estimated with BRM. A roughly estimate of parameters α, η_1 and η_2 was
obtained from the Weibull QQ-plot, see Figure 4.
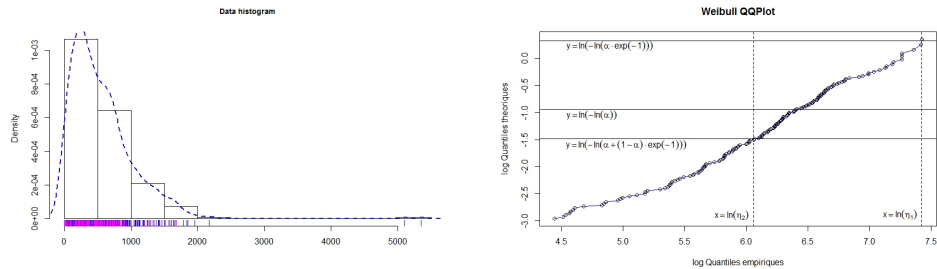
Figure 4: Real dataset, 729 automotive parts put into service during several years ; 74% of
them are still operating.

Using BRM, the data were fitted with a two-component Weibull mixture,
see Figure 5.

Credibility intervals and estimates was computed for all parameters and for
reliability to short, medium and long term, see Table 8. The component 2
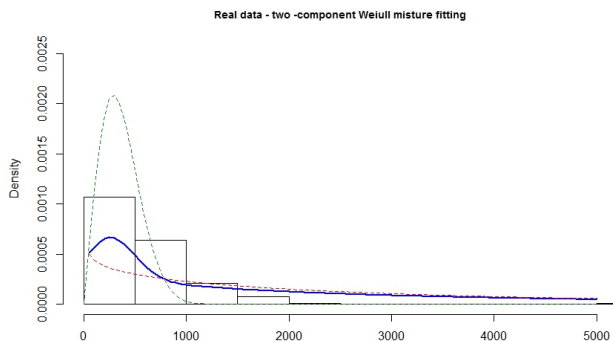contains weak parts, and component 1 strong ones, $\widehat{\beta}_2 > \widehat{\beta}_1$ and $\widehat{\eta}_2 < \widehat{\eta}_1$.

19

Figure 5: Real dataset ; two-component Weibull mixture fitting.

	credibility interval		
	2.5%	**97.5%**	**BRM estimation**
α	79.7%	89.5%	86.%
β_1	0.84	0.95	0.86
β_2	1.89	2.37	1.99
η_1	2 959.93	3 393.87	2966.0
η_2	388.31	534.92	409.77
$R(200)$	87.3%	90.4%	88.4%
$R(1\,000)$	53.5%	61.0%	55.4%
$R(2\,000)$	38.9%	45.9%	40.1%

Table 8: Credibility intervals and BRM estimations

5.6. Discussions

By incorporating prior information, BRM allows to overcome the issue of
highly censoring for inferences using Weibull mixture. In addition, since the
likelihood is easy to compute, BRM is less time-consuming than iterative al-
gorithms such as EM, S-EM or MCMC. It is a Bayesian bootstrap method :
parameters are firstly simulated from priors, then based on the weights, a new
sample is obtained. This sample is similarly distributed as the posterior distri-
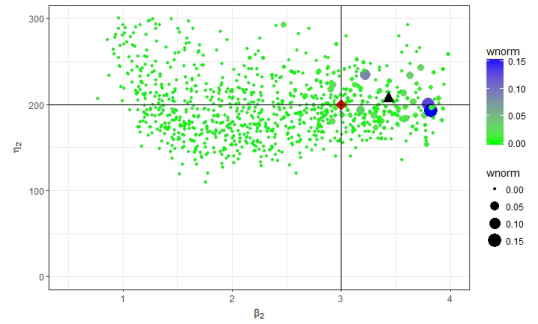bution, see Figure 6.
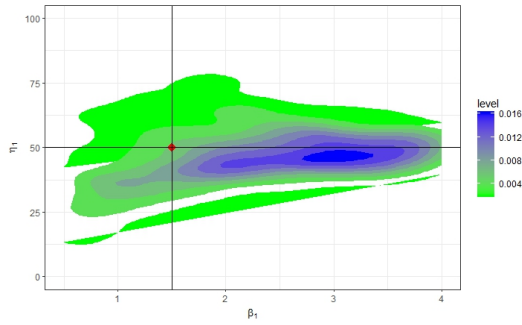
(β_1, η_1) joint prior pdf
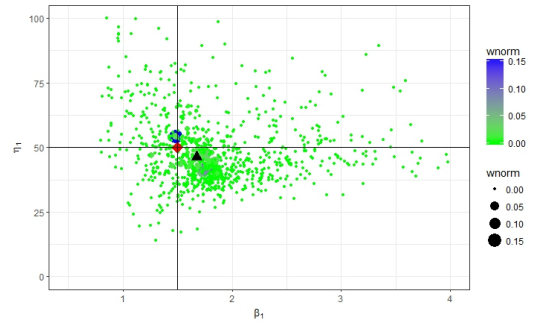
(β_1, η_1) with the weights
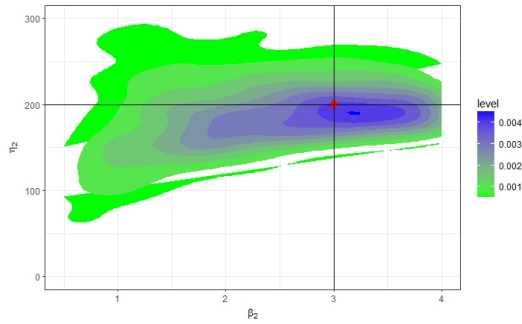
(β_2, η_2) joint prior pdf

(β_2, η_2) with the weights

Figure 6: BRM is a Bayesian bootstrap method [12]; parameters are simulated from prior, then, from importance sampling weights, a new sample is obtain which is similarly distributed as the posterior distribution; the true values of the parameter are highlighted and marked by a diamond ; the BRM estimates are marked by a triangle.

Simulations showed that there is no label-switching issue with a weak informative beta prior on $[0 \, ; \, 0.5]$ for the mixing proportion, α_1.

6. Conclusion

Systems and components lifetime depends on various internal and external factors. The failure rate could vary according to manufacturing quality, wearing intensity or environmental conditions. Frequently, these variations are very significant, and require specific analysis. Hence, Weibull mixture is a appropriated model to handle heterogeneity in lifetime data collection. Unfortunately, sources of this non-homogeneity are not controlled or not observed. Furthermore, in industrial context the data are heavily censored. In this framework, performance of classical estimation methods (maximum of likelihood, EM, full Bayes and MCMC) are significantly reduced. In this paper, inferences for a two-component Weibull mixture have been made using the Bayesian restoration method (BRM). It is Bayesian bootstrap method : firstly, model parameters are sampled from the joint prior. Next, a new sample is obtain using a importance sampling technique. This sample, thus obtained, is similarly distributed as the posterior distribution.

Priors elicitation is the main issue in Bayes estimations. Here, poor informative prior are used, and prior elicitation is very simple, making it operable in various industrial applications. By incorporating prior information, for the Weibull mixture model when the mixing cause is missing and in the context of heavily censoring, simulation shows that BRM provides more robust estimations than S-EM. In addition, since the likelihood is here accessible, BRM is less time-consuming than iterative method such as S-EM, or MCMC methods. To conclude, in an industrial context for which data heterogeneity must be handle, with however heavily censoring data, BRM is a simple, low-cost and quite precise estimation method for reliability or warranty statistical analysis.

References

[1] M. Finkelstein, J. H. Cha, Stochastic Modeling for Reliability, Springer, 2013.

[2] W. Wang, M. Jiang, Competing failure or mixed failure models, in: Reliability and Maintainability Symposium (RAMS), 2014 Annual, IEEE, 2014, pp. 1–6.

[3] D. N. P. Murthy, M. Xie, R. Jiang, Weibull Models, Jonh Wiley and Sons, 2004.

[4] E. E. Elmahdy, A new approach for Weibull modeling for reliability life data analysis, Applied Mathematics and Computation 250 (2015) 708–720.

[5] R. Jiang, D. N. P. Murthy, Modeling failure-data by mixture of 2 Weibull distributions: a graphical approach, IEEE Transactions on Reliability 44 (3) (1995) 477–488.

[6] D. Chauveau, A stochastic EM algorithm for mixtures with censored data, Journal of Statistical Planning and Inference 46 (1995) 1–25.

[7] J.-P. Baudry, G. Celeux, EM for mixtures, Statistics and computing 25 (4) (2015) 713–726.

[8] A. E. Touw, Bayesian estimation of mixed Weibull distributions, Reliability Engineering and System Safety 94 (2) (2009) 463–473.

[9] D. M. Titterington, A. F. Smith, U. E. Makov, Statistical analysis of finite mixture distributions, John Wiley and Sons, 1985.

[10] J.-M. Marin, K. Mengersen, C. P. Robert, Bayesian Modelling and Inference on Mixtures of Distributions, Handbook of Statistics 25 (2005) 459–507.

[11] M. Bacha, G. Celeux, Bayesian Estimation of a Weibull Distribution in a Highly Censored and Small Sample Setting, Tech. rep., INRIA (1996).

[12] D. B. Rubin, The Bayesian Bootstrap, The Annals of Statistics 9 (1) (1981) 130–134.

[13] H. Rinne, The Weibull distribution: a handbook, CRC Press, 2009.

[14] D. N. P. Murthy, M. Bulmer, J. A. Eccleston, Weibull model selection for reliability modelling, Reliability Engineering and System Safety 86 (3) (2004) 257–267.

[15] J. E. Gentle, G. J. McLachlan, T. Krishnan, The EM Algorithm and Extensions, Jonh Wiley and Sons, 1998.

[16] M. Hamada, A. G. Wilson, C. S. Reese, H. F. Martz, Bayesian Reliability, Springer, 2008.

[17] D. B. Rubin, Using the SIR algorithm to simulate posterior distributions, Bayesian statistics (1988) 395–402.

[18] J. M. Hammersley, D. C. Handscomb, Monte Carlo Methods, Springer, 2013.

[19] D. W. Scott, Multivariate density estimation: theory, practice, and visualization, John Wiley and Sons, 2015.

[20] G. Celeux, Bayesian inference for mixture: The label switching problem, in: Compstat, Springer, 1998, pp. 227–232.

[21] R. M. Soland, Bayesian analysis of the Weibull process with unknown scale and shape parameters, IEEE Transactions on Reliability 18 (4) (1969) 181–184.

[22] R. B. Abernethy, The New Weibull Handbook, RB Abernethy, 2004.

[23] R. Jiang, D. N. P. Murthy, A study of Weibull shape parameter: Properties and significance, Reliability Engineering and System Safety 96 (12) (2011) 1619–1626.

[24] P. H. Barringer, Weibull databases (2000).
URL http://www.barringer1.com/wdbase.htm

[25] P. Erto, M. Giorgio, Assessing high reliability via Bayesian approach and
accelerated tests, Reliability Engineering and System Safety 76 (3) (2002)
301–310.

335