

A protocol for integrating FED and expert data in a study of durability using the Weibull distribution.

Nicolas BOUSQUET, Gilles CELEUX
INRIA Futurs, Equipe SELECT, Université Paris-Sud Orsay, France.

Emmanuel REMY
EDF R&D, site de Chatou, France.

ABSTRACT : The main issues raised by the estimation of lifetime parametric models used in industrial modelling of reliability are censoring and FED (Feedback Experience Data) sample size. Many studies are facing homogeneous, small-sized, censored failure times which have to be integrated into Bayesian procedures with informative prior parameter. This way of dealing with statistical inference has been especially followed by EDF for predicting failures on nuclear material. The example of the Weibull distribution will be here thoroughly treated. Firstly experts have to be asked about the durability of a material with precise and simple questions. According to the choice of the considered model, prior point estimations and confidence intervals about parameters must be given, directly or indirectly, by experts. Secondly efficient modelling has to be chosen for informative prior distributions. Once, it must produce posterior distributions easily estimated by classical methods. But computation complexity is often a limiting factor of Bayesian inference. The impact of prior choices on posterior results must be simple to derive. Then, hyperparameters of these prior distributions must be evaluated linking the intrinsic properties of the prior densities (mean, mode, variance, etc.) with expert information on parameters.

1 INTRODUCTION

The context of our study is typically the estimation of durability parametric models using high-censored small samples of failure times (called FED). Classical methods of likelihood maximization failed for these reasons and the use of Bayesian tools are highly recommended (Bacha *et al.*, 1998). For many reasons (cost, loss of know-how, etc.) but mainly because very few real failure data are available on nuclear materials, EDF R&D needs integrating expertise from engineers, technicians, etc. into efficient informative prior distributions for durability models. The Weibull distribution $\mathcal{W}(\eta, \beta)$ (described in §2) being the most used model in all reliability/durability problems, studying the integration of FED and expert data in Weibull prior distributions is of paramount importances.

2 NOTATION

Time random variable : T .

Weibull pdf : $f_W(t) = \frac{\beta}{\eta^\beta} t^{\beta-1} e^{-\left(\frac{t}{\eta}\right)^\beta} \mathbb{I}_{[0,+\infty)}(t)$.

Weibull hazard rate : $h_W(t) = \frac{\beta}{\eta} \left(\frac{t}{\eta}\right)^{\beta-1}$.

Weibull parameters : $\lambda = 1/\eta$, $\mu = \lambda^\beta = 1/\eta^\beta$.

Prior distribution on θ : $\pi(\theta)$.

Posterior distribution on θ knowing x : $\pi(\theta|x)$.

Gamma (a,b) distribution : $\mathcal{G}(a, b)$.

FED : $\mathbf{x}_n = (x_1, \dots, x_n)$ including

- r uncensored data $\mathbf{y}_r = (y_1, \dots, y_r)$
- $n - r$ censored data $\mathbf{c}_{n-r} = (c_1, \dots, c_{n-r})$.

3 WEIBULL INFERENCE

3.1 Prior conditioning

The scale parameter η is directly linked to lifetime T , by the mean, median and mode of the Weibull distribution. Moreover it takes the 63rd percentile value of the distribution. The Weibull likelihood does not admit any continuous conjugate distribution for the parameters (Soland, 1966) when β is unknown. However conjugate prior distribution on η can be defined conditionally to β .

Definition 1 : Generalized Inverse Gamma distribution.

$\forall (a, \beta, b) \in (\mathbb{R}^{*+} \times \mathbb{R}^{*+} \times \mathbb{R}^{*+}), \forall x \in \mathbb{R}^+$,

$$f_{\mathcal{G}\mathcal{I}\mathcal{G}(a,\beta,b)}(x) = \frac{b^\beta}{\Gamma(a)} \frac{1}{x^{\beta+1}} e^{-\frac{b}{x^\beta}} \mathbb{1}_{[0,+\infty)}(x).$$

Suppose $\pi(\eta|\beta) \rightsquigarrow \mathcal{G}\mathcal{I}\mathcal{G}(a,\beta,b)$. Denote $\mu = \lambda^\beta$. Then $\pi(\mu|\beta) = \mathcal{G}(a, b)$ and, using the Bayes' rule,

$$\pi(\mu|\beta, \mathbf{x}_n) = \mathcal{G}\left(a+r, b + \sum_{i=1}^n x_i^\beta\right),$$

$$\pi(\eta|\beta, \mathbf{x}_n) = \mathcal{G}\mathcal{I}\mathcal{G}\left(a+r, \beta, b + \sum_{i=1}^n x_i^\beta\right).$$

Using conjugate prior Gamma distributions for μ knowing β allows to simplify prior modelling and posterior estimation correcting. Posterior estimations will be derived from full conditional sampling of β sampling (by importance sampling or MCMC algorithms). We note that the prior control on a is conditional to r while the prior control on b is conditional to the FED values and β (see Billy *et al.*, 2005). Thus the μ parameter will be introduced into inference and the Bayesian model can be described by

$$T \rightsquigarrow \mathcal{W}(\mu, \beta) \text{ with} \\ \pi(\beta, \mu) = \pi(\beta) \pi(\mu|\beta) = \pi(\beta) \mathcal{G}(a, b(\beta)) \quad (1)$$

3.2 Questioning experts for building prior distributions

The choice of $\pi(\beta)$ and the hyperparameters a and b must be related to the answers given by one or several experts. Usually, they give quantities linked to lifetime T and reliability. An independent analyst must translate these expert opinions into quantities regarding η and β . By the fact that β is a shape parameter of the model it is more difficult to link T and β than T and η .

3.3.1 Choosing $\pi(\beta)$ from expert opinions

The shape parameter β gives a measure of the aging behavior of the system. However, an expert is not supposed to give direct values of β . If the system is submitted to aging, the prior variation domain for β , named Ω_β , is included into $[1, +\infty)$. Some precisions can be provided indirectly by one or many experts. Notably five understandable questions around the aging behavior have been proposed in Bacha *et al.* (1998).

One of the most interesting is questioning around the possibility of aging acceleration. For instance, the acceleration of aging is translated by $h_W(t) \geq 0 \forall t$ which means $\beta \geq 2$.

Another way is the proposed approach in Biernacki *et al.* (1998, p.15). A value of β is obtained from a question on the ratio r between two measures $h_W(t_0)$ and $h_W(t_1)$ of the hazard rate of the system at different times t_0 and t_1 . If an expert can give $r = h_W(t_1)/h_W(t_0)$, then an analyst obtains a value

$$\beta = 1 + \frac{\log r}{\log t_1 - \log t_0}.$$

However no reliability expert is expected to answer to this question. Hoping to be more understandable for experts, we propose another question. We ask to an expert the probability p_{t_0} and $p_{\alpha t_0}$ for the studied component to break down before time t_0 and before time αt_0 with $\alpha > 1$. Then $p_{t_0} = 1 - e^{-\mu t_0^\beta}$ et $p_{\alpha t_0} = 1 - e^{-\mu \alpha^\beta t_0^\beta}$. So

$$\frac{1 - p_{\alpha t_0}}{1 - p_{t_0}} = e^{-\mu t_0^\beta (\alpha^\beta - 1)} = (1 - p_{t_0})^{\alpha^\beta - 1} \\ \Rightarrow \beta = \log \left\{ \frac{\log(1 - p_{t_1})}{\log(1 - p_{t_0})} \right\} \log^{-1}(\alpha).$$

This way of asking an expert is more realistic. In reliability an answer under the form of a probability value seems less sensitive to uncertainty (and more honest) than direct values on quantities (already notified in IsdF (2000) and Clarotti (1998)).

Finally all expert can only propose indirectly some pointwise values for β or interval bounds for Ω_β . With no particular preference for one of them, a simple way of building $\pi(\beta)$ is to start from a large domain Ω_β including all possibilities, and choosing a prior uniform model

$$\pi(\beta) = \mathcal{U}_{\Omega_\beta}.$$

We denote $\Omega_\beta = [\beta_l, \beta_r]$. Note that a non-informative prior distribution on β is obtained when $\beta_l \rightarrow 0$ and $\beta_r \rightarrow \infty$. Very often, $\Omega_\beta = [1, 5]$ is chosen as a handy informative prior domain.

3.3.2 Choosing hyperparameters of $\pi(\mu|\beta)$

We consider two kinds of expert opinions. An expert is supposed to give

- a single value t_m which represents (according to him) a mean/median/mode of T .
- or two values (quantiles) (t_l, t_r) which surround the mean of T with confidence $1 - \alpha$.

The value of a must be independent of β and fixed conditionally to r , according to the will of the analyst to confront the prior modelling to the Weibull likelihood of FED. As an object of critical analysis of expert opinions, it will be especially studied in §4. According to the formula (1), the b values must be chosen conditionally to β .

Note that conditionally to β , a measure t of T can be converted into a value for η , using explicit transformation formulas from the Weibull characteristics

(mean/median/mode/quantiles). Typically the value t is right-shifted because the value of η is always above the Weibull mean/median/mode.

Case A : one prior value t_m available.

Conditionally to β , t_m can be converted into an estimation of the prior mean of η using the following relations (mean/median/mode)

$$\begin{aligned}\tilde{E}[\eta|t_m, \beta] &= \frac{t_m}{\Gamma(1+1/\beta)}, \\ \tilde{E}[\eta|t_m, \beta] &= \frac{t_m}{(1-1/\beta)^{1/\beta}}, \\ \tilde{E}[\eta|t_m, \beta] &= \frac{t_m}{(\ln 2)^{1/\beta}}\end{aligned}$$

With $\eta \rightsquigarrow \mathcal{G}\mathcal{I}\mathcal{G}(a, \beta, a)$ conditionally to β , we choose

$$\begin{aligned}E[\eta|\beta] &= \frac{\Gamma(a-1/\beta) b^{1/\beta}}{\Gamma(a)} = \tilde{E}[\eta|t_m, \beta] \\ \Rightarrow b(\beta) &= \left(\frac{\Gamma(a)}{\Gamma(a-1/\beta)} \tilde{E}[\eta|t_m, \beta] \right)^\beta.\end{aligned}$$

Case B : two prior values (t_l, t_r) available with confidence $1 - \alpha$.

Conditionally to β , (t_l, t_r) can be converted into $(\eta_l(\beta), \eta_r(\beta))$ including $E[\eta|\beta]$ with confidence $1 - \alpha$.

$$\eta_l(\beta, \alpha) = \frac{t_l}{\Gamma(1+1/\beta)}, \quad \eta_r(\beta, \alpha) = \frac{t_r}{\Gamma(1+1/\beta)}.$$

Moreover, with $\eta \rightsquigarrow \mathcal{G}\mathcal{I}\mathcal{G}(a, \beta, a)$ conditionally to β , an interval for $E[\eta|\beta]$ with confidence $1 - \alpha$ is known :

$$\left(\frac{2b}{\chi_{2a}^2(1-\alpha/2)} \right)^{1/\beta} \leq E[\eta|\beta] \leq \left(\frac{2b}{\chi_{2a}^2(\alpha/2)} \right)^{1/\beta}.$$

This confidence interval leads to define a interval $\Omega_{b(\beta)}$ for $b(\beta)$ whose bounds are

$$\begin{aligned}b_l(\beta) &= \left(\frac{\Gamma(a)}{\Gamma(a-1/\beta)} \eta_l(\beta, \alpha) \right)^\beta, \\ b_r(\beta) &= \left(\frac{\Gamma(a)}{\Gamma(a-1/\beta)} \eta_r(\beta, \alpha) \right)^\beta.\end{aligned}$$

Definition 2 : Integration of expert opinions by least square method.

In order to decide one $b(\beta)$ from both information (η_l, η_r) coming indirectly from expert opinion, we define the integration criterion of expert opinion as

$$\begin{aligned}Int_E(a, b, \beta, \alpha) &= \frac{\sqrt{\left[\eta_l(\beta) - \left(\frac{2b}{\chi_{2a}^2(1-\alpha/2)} \right)^{1/\beta} \right]^2}}{\eta_l(\beta)} \\ &+ \frac{\sqrt{\left[\eta_r(\beta) - \left(\frac{2b}{\chi_{2a}^2(\alpha/2)} \right)^{1/\beta} \right]^2}}{\eta_r(\beta)}.\end{aligned}$$

The minimization of $Int_E(a, b, \beta, \alpha)$ criterion on $\Omega_{b(\beta)}$ is realized with

$$b^*(\beta) = \frac{1}{2} \frac{\left[\frac{(\chi_{2a}^2(1-\alpha_E/2))^{1/\beta}}{\eta_l} + \frac{(\chi_{2a}^2(\alpha_E/2))^{1/\beta}}{\eta_r} \right]^\beta}{\left[\left(\frac{\chi_{2a}^2(1-\alpha_E/2)}{\eta_l} \right)^2 + \left(\frac{\chi_{2a}^2(\alpha_E/2)}{\eta_r} \right)^2 \right]^\beta}.$$

The graph (η, β) of joint prior sampling generally respects the classical form which reflects the antagonist behavior of both parameters (Fig. 1). It is well known that estimations of η are often underestimated while β is overestimated (and conversely).

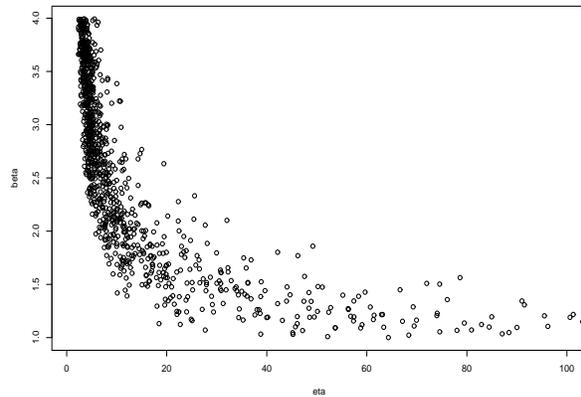


Figure 1. Typical form of the joint prior confidence area for the Weibull parameters

3.4 Posterior estimation

Such a conditioning way of building our prior distributions makes easy the use of Bayesian sampling

algorithms. Moreover, the missing data context makes data augmentation methods desirable (Robert, 1994). Each step of algorithms corresponds to a simulation of the conditional posterior pdf of both parameters (μ, β) . The most difficult part, at each step, in these stochastic algorithm is sampling from the conditional posterior of β whose pdf is not explicit. Many ways exist to circumvent this difficulty, for instance acceptance-rejection methods.

The nature of missing data and the will of getting rapidly a good approximation of the posterior pdf have led to the choice of importance sampling algorithms proposed in Douc *et al* (2005) and Celeux *et al* (2005), called Population Monte Carlo algorithms (PMC). The stabilization of conditional posterior sampling is reached in very few iterations.

4 CRITICAL ANALYSIS AND ADDITION OF EXPERTNESS

In Billy *et al.* (2005), the critical analysis of an expert corresponds to the choice of a size for a fictitious sample which is supposed to represent the expert opinion. In our case, the hyperparameter a represents this trust in the expert opinion and must be directly compared to the characteristics of the REX data.

Remark : If $a \in \mathbb{N}$, denote $(y_i(\beta))_{i \in \{1, a\}}$ such an unknown Weibull fictitious sample. Then $b(\beta) = \sum_{i=1}^a y_i^\beta$ (see Billy *et al.* (2005)).

Thus, if we suppose that β is known (independently from expertness), the posterior pdf of μ and η are entirely known. There is no need of Bayesian sampling and a can directly be compared to r (the number of uncensored failure times).

However, if β is considered as a random variable, the use of data augmentation algorithms in Bayesian procedure makes preferable the comparison between a and n (the size of the sample).

Note that in case of a two-values (t_l, t_r) opinion, the note a and the confidence $1 - \alpha$ can evolve in parallel for decreasing the trust given in an expert.

Considering an expert as a "provider" of fictitious data allows to combine expert opinions in concatenating all fictitious data. Indeed, the prior domain Ω_β has been created using all the available experts and

stays the same for each modelling. Suppose we have M experts $(E_i)_{i \in \{1, \dots, M\}}$ give time information (t_{E_i}) and that they lead to the modelling on μ defined by

$$\pi(\mu | \beta, t_{E_i}) = \mathcal{G}(a_i, b_i(\beta)).$$

Then the general prior modelling get from combining experts opinions will be

$$\begin{aligned} \pi(\mu, \beta) &= \pi(\mu | \beta, \cup_{i \in \{1, \dots, M\}} t_{E_i}) \pi(\beta) \\ &= \mathcal{G}\left(\sum_{i=1}^M a_i, \sum_{i=1}^M b_i(\beta)\right) \mathcal{U}_{\Omega_\beta}(\beta). \end{aligned}$$

Note that this combination must be realized with care. The total size of fictitious data $\sum_{i=1}^M a_i$ must be put in comparison with information quantities from real data, especially the Fisher information term \tilde{n} of the data (see Billy *et al.*, 2005). For instance, if the same importance is given to the real data and the addition of all experts, a multinomial choice of a_i in $]0, \tilde{n}]$ or a constant choice $a_i = \tilde{n}/M$ (if there is no preference between the experts) can be made. Because we work generally with high-censored small-sized samples, \tilde{n} is often fewer than 20. This implies to choose small values of a_i , which can oppose to an exaggerated self-confidence of the experts.

5 REFERENCES

- Bacha M., Celeux G., Idée E., Lannoy A. & Vasseur D. (1998) '*Estimation de modèles de durée de vie fortement censurées*', Eyrolles.
- Biernacki C., Celeux G., Villain B. & Vérité B. (1998) '*Utilisation des opinions d'experts pour l'analyse et la dégradation des structures passives*', Rapport de Recherche INRIA-EDF, Projet IS2, Montbonnot.
- Billy F., Bousquet, N. & Celeux G. (2005) '*Modeling and eliciting expert knowledge with fictitious data*'. Proceedings of the Workshop on the Use of Expert Judgement in Decision-Making, 21-23 June.
- Celeux, G., Marin J.-M. & Robert C.P. (2005) '*Iterated importance sampling in missing data problems*', Rapport de recherche INRIA RR-5534.
- Clarotti C.A. (1998). '*Les techniques fréquentielles et bayésiennes au service de l'ingénieur de sûreté de fonctionnement*'. Rapport final du projet IsdF 8/96, June.
- Douc, R., Guillin A., Marin J.-M. & Robert C.P. (2005) '*Convergence of adaptive sampling schemes*'. Rapport de recherche INRIA RR-5485.
- IsdF (2000). '*Méthodes de collecte et de traitement d'avis d'experts et guide de mise en oeuvre en sû-*

reté de fonctionnement'. Rapport de projet IsdF
n°6/98, February.

Meeker, W. Q. & Escobar, L. A. (1998) '*Statistical
Methods for Reliability Data*', Wiley.

Robert, C. P. (1994) '*The Bayesian Choice, a
Decision-Theoretic Motivation*', Springer-Verlag.

Soland, R. (1966). 'Use of the Weibull Distribution
in Bayesian Decision Theory', Report No. RAC-
TP-225, Research Analysis Corporation,
McLean, VA.