Received: 28 April 2023 | Revised: 8 May 2023 | Accepted: 11 May 2023

DOI: 10.1002/asmb.2781

COMMENTARY

WILEY

Discussion of "Specifying prior distributions in reliability applications"

Gianpaolo Pulcini

Istituto di Scienze e Tecnologie per Energia e Mobilità Sostenibili (STEMS), CNR, Napoli, Italy

Correspondence
Gianpaolo Pulcini, Istituto di Scienze e Tecnologie per Energia e Mobilità Sostenibili (STEMS), CNR, Napoli, Italy.
Email: gianpaolo.pulcini@stems.cnr.it

1 | INTRODUCTION

The article of Tian et al.[1] presents a very extensive review of Bayesian inferential procedures for the analysis of censored data, providing useful guidances for setting prior distributions for log-location-scale distributions commonly used in reliability applications, such as the lognormal, the Weibull, the loglogistic, and the Fréchet distribution.

Several kinds of censoring that arise in practical applications are discussed, In particular: (a) the right censoring, which arises when the test ends before all units fail, (b) the interval censoring, which arises when failures are found only at inspection times, so that all that is known is that a failure occurred between the last and the current inspection, and (c) the left censoring, which arises when the unit is found to be failed at the first inspection time.

Large attention is devoted to the right censoring, distinguishing between time (Type 1) censoring, which occurs when the test time is fixed and the number m of observed failures is random (and can be zero), and failure (Type 2) censoring, which occurs when the test ends after a specified number m of units have failed, so that the test time is random. Clearly, a complete sample can be viewed as a special case of the Type 2 censored sample, when m is set equal to the sample size n.

For such kinds of censoring, noninformative and weakly informative prior distributions are discussed. In particular, the Jeffreys prior, the Independence Jeffreys prior, and the reference prior are proposed, under the very useful reparameterization of the log-location-scale distributions in terms of the p-th quantile of the distribution, say t_p, and of the scale parameter σ (or its reciprocal $\beta = 1/\sigma$).

Possible combinations of the above mentioned noninformative or weakly informative prior distributions with informative prior distributions for t_p and σ are also proposed and discussed. Among all these combinations, those based on proper prior distributions appear to be of great interest and practical application when a Bayesian approach is used for model selection, in particular by using the Bayes factor, as discussed in Section 2 of this short discussion.

2 | MODEL SELECTION AND BAYES FACTOR

Many of the proposed Bayesian methods for model comparison usually rely on the Bayes factor, originally developed by Harold Jeffreys,[2] that allows one to quantify the evidence in favor of one statistical model compared to another.

In particular, given two competing models, say M_0 and M_1, the Bayesian approach to model selection is based on the posterior probabilities:

$$\Pr\{M_0|\text{data}\} \quad \text{and} \quad \Pr\{M_1|\text{data}\} = 1 - \Pr\{M_0|\text{data}\}, \tag{1}$$

which measure the evidence in favor of models M_0 and M_1, respectively, given the observed data. Once the prior probabilities $\Pr\{M_0\}$ and $\Pr\{M_1\} = 1 - \Pr\{M_0\}$ that the model M_k $(k = 0, 1)$ is the correct one have been expressed, the posterior probability $\Pr\{M_k|\text{data}\}$ $(k = 0, 1)$ results in:

$$\Pr\{M_k|\text{data}\} = \frac{\Pr\{\text{data}|M_k\}\Pr\{M_k\}}{\Pr\{\text{data}|M_0\}\Pr\{M_0\} + \Pr\{\text{data}|M_1\}\Pr\{M_1\}}, k = 0, 1, \tag{2}$$

where $\Pr\{\text{data}|M_k\}$ denotes the marginal probability of the observed data under the model M_k ($k = 0, 1$), which is obtained by integrating over the parameters space the product of the likelihood function $\mathscr{L}(\text{data}|\,\boldsymbol{\theta}; M_k)$ and the joint prior distribution for the vector $\boldsymbol{\theta}$ of the model parameters. For example, if model M_0 is the Weibull distribution of Equation (3) in Tian et al.[1] reparameterized in terms of the p-th quantile t_p (see equation (7) in Tian et al.[1]), and the data arise from a Type 1 right censoring test, then:

$$\begin{aligned}\Pr\{\text{data}|M_0\} &= \int_{t_p}\int_{\beta} \mathscr{L}(\text{data}|\,t_p, \beta; M_0)\,\pi\left(t_p, \beta|M_0\right)\,\mathrm{d}t_p\mathrm{d}\beta \\ &= \int_{t_p}\int_{\beta}\left(\frac{\beta\,[-\ln(1-p)]}{t_p}\right)^m \left(\prod_i^m \left(\frac{t_i}{t_p}\right)^{\beta-1} \exp\left[\ln(1-p)\left(\frac{t_i}{t_p}\right)^{\beta}\right]\right) \\ &\quad \times \exp\left[(n-m)\ln(1-p)\left(\frac{T}{t_p}\right)^{\beta}\right]\,\pi\left(t_p, \beta|M_0\right)\,\mathrm{d}\beta\,\mathrm{d}t_p,\end{aligned} \tag{3}$$

where $\mathscr{L}(\text{data}|\,t_p, \beta; M_0)$ is the likelihood function of the observed data under the Weibull model, n is the sample size, T is the censoring time, m is the (random) number of observed failures, t_i ($t_i \le T$, $i = 1, \ldots, m$) are the observed failure times, and $\pi\left(t_p, \beta|M_0\right)$ denotes the joint prior distribution for the parameters t_p and β of the model M_0. The marginal probability $\Pr\{\text{data}|M_k\}$ is sometimes called the marginal likelihood or the integrated likelihood.

Once the posterior probabilities (2) have been computed, a summary of evidence provided by the observed data in favor of model M_0 is given by the Bayes factor B_{01} (see, e.g., Kass and Raftery[3] and Campbell and Gustafson[4]), which is defined as the ratio of the posterior odds of M_0 to its prior odds:

$$B_{01} = \frac{\Pr\{M_0|\text{data}\}\,/\,\Pr\{M_1|\text{data}\}}{\Pr\{M_0\}\,/\,\Pr\{M_1\}}\,. \tag{4}$$

When the models M_0 and M_1 are equally prior probable, so that $\Pr\{M_0\} = \Pr\{M_1\} = 0.5$, then the Bayes factor is equal to the posterior odds $\Pr\{M_0|\text{data}\}\,/\,\Pr\{M_1|\text{data}\}$. Moreover, from (2), we have that:

$$B_{01} = \frac{(\Pr\{\text{data}|M_0\}\Pr\{M_0\})\,/\,(\Pr\{\text{data}|M_1\}\Pr\{M_1\})}{\Pr\{M_0\}\,/\,\Pr\{M_1\}} = \frac{\Pr\{\text{data}|M_0\}}{\Pr\{\text{data}|M_1\}}, \tag{5}$$

so that the Bayes factor is equal to the ratio of the marginal probabilities (3) and is independent from the prior probabilities $\Pr\{M_k\}$. Thus, the Bayes factor is sometimes interpreted as the actual odds of the models implied by the data alone (see, Berger and Selike[5]). However, the presence of the prior distributions $\pi\left(t_p, \beta|M_k\right)$ for the models parameters in the marginal probability (3) prevents any interpretation from being in a non-Bayes context.

Large values of B_{01} provide evidence in favor of the model M_0, whereas large values of $B_{10} = 1/B_{01}$ provide evidence against M_0. Appropriate bounds for $2\ln(B_{10})$ able to measure the evidence against the model M_0 can be found in Kass and Raftery,[3] and are listed below:

$2\ln(B_{10})$	Evidence against M_0
0–2	Not worth more than a bare mention
2–6	Positive
6–10	Strong
> 10	Very strong

The above bounds are given on B_{10} (and not on B_{01}) because it is more familiar to speak in terms of evidence against the model M_0, rather than in favor of it. Alternative bounds, given in that case for B_{10}, can be found, for example, in Jeffreys[6] and in Lee and Wagenmakers.[7]

In order to compute the Bayes factor, the prior distributions for the parameters of each model have to be specified. This step is anything but trivial, first of all because the prior distributions must be proper in order to compute the marginal probability $\Pr\{\text{data}|M_k\}$ of the observed data under the model M_k $(k = 0, 1)$. Otherwise, the $\Pr\{\text{data}|M_k\}$, and consequently the Bayes factor, would be computed up to an undefined multiplicative constant. Thus, if no information for a model parameter is available, a proper noninformative or a weakly informative prior distribution must be used. To this end, if the competing models belong to the log-location-scale family, the weakly informative prior distributions discussed in Tian et al.[1] and provided in table 2 constitute a very useful summary of feasible prior distributions.

Moreover, the Bayes factor is quite sensitive to the choices of prior distributions. In particular, if the prior distribution for a parameter of the model M_0 is strong but wrong, then the Bayes factor B_{01} can provide evidence against the model M_0 even if this model is the most appropriate to describe the observed data. Thus, strong prior distributions are suggested to be used only if the prior distribution is expressed in terms of a parameter that indexes both the models. To this end, the reparameterization suggested in Tian et al.[1] for the log-location-scale distributions in terms of the p-th quantile t_p allows strong prior information for the failure probability to be elicited, without taking the risk of unintentionally penalizing one of the models. A useful choice of informative prior distributions for t_p, able to describe any prior information for this quantile, is provided in table 2 of Tian et al.[1]

Thus, a large number of the recommended prior distributions given in table 2 of Tian et al.[1] in order to make inference on the parameters, and function thereof, of log-location-scale distributions, can be used to compute the Bayes factor and then select the model that, within the family of log-location-scale distributions, best fits a given set of failure data. In particular, all the combinations of a strong or weakly informative prior distribution for the quantile t_p and of a weakly informative or proper noninformative prior distribution for $\beta = 1/\sigma$ can be used for model selection, with the certainty that: (a) the Bayes factor is not computed up to an undefined multiplicative constant, and (b) the use of a strong but wrong prior information for t_p does not cause the best model to be rejected.

3 | FURTHER DEVELOPMEMENTS

In some circumstances, the analyst possesses prior information on the reliability level at a given time τ, say $R_\tau = 1 - F(\tau; \mu, \beta)$, that can not be easily converted into prior information on a quantile of the distribution. Thus, it could be useful to reparameterize the log-location-scale distribution in terms of β and R_τ that, in case of the Weibull distribution, means to replace η with the expression $\eta = \tau/[-\ln(R_\tau)]^{1/\beta}$ in (3) of Tian et al.,[1] so that the cumulative distribution function becomes:

$$\mathrm{F}(t; R_\tau, \beta) = 1 - \exp\left[\ln(R_\tau)\left(\frac{t}{\tau}\right)^{\beta}\right], t > 0.$$

Once informative and weakly informative prior distributions on R_τ have been formulated, similarly to what was done in Tian et al.[1] for the quantile t_p, these prior distributions can be used together to the proper noninformative or weakly informative prior distributions for β (or for $\sigma = 1/\beta$) both for making inference and for model selection by using the Bayes factor.

4 | CONCLUSIONS

The article of Tian et al.[1] proposes and discusses a large number of prior distributions for log-location-scale distributions used in reliability applications. These prior distributions are formulated both in terms of the classical parameters, say μ and σ, of the log-location-scale distribution, and in terms of σ and the p-th quantile t_p. This last reparameterization, when accompanied by a weakly informative or proper noninformative prior distribution for σ (or its reciprocal $\beta = 1/\sigma$) and by an informative prior distribution on the quantile t_p, allows also to correctly perform a model selection (among the log-location-scale family of distributions) based on the Bayes factor, that is in compliance with the requirements for prior distributions: (a) all the prior distributions must be proper, (b) the prior distribution on a parameters indexing only one of the competing models must be noninformative or weakly informative, and (c) a strong prior distribution can be used only on a parameter indexing both the competing models, in their original or new parameterization, such as the p-th quantile t_p or the reliability level R_τ at a given time τ.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

ORCID

Gianpaolo Pulcini https://orcid.org/0000-0001-5474-7095

REFERENCES

1. Tian Q, Lewis-Beck C, Niemi JB, Meeker WQ. Specifying prior distributions in reliability applications. *Applied Stochastic Models in Business and Industry* (forthcoming). 2023. doi:10.1002/asmb.2752
2. Jeffreys H. Some test of significance, treated by the theory of probability. *Procedings of the Cambridge Philosophy Society*. 1935;31:203-222.
3. Kass RE, Raftery AE. Bayes factors. *J Am Stat Assoc*. 1995;90:773-795.
4. Campbell H, Gustafson P. Bayes factors and posterior estimations: two sides of the very same coin. *The American Statistician*. 2022. doi:10.1080/00031305.2022.2139293
5. Berger JO, Selike T. Testing a point null hypothesis: the irreconcilability of p values and evidence. *J Am Stat Assoc*. 1987;82:112-122.
6. Jeffreys H. *Theory of Probability*. Oxfortd University Press; 1961.
7. Lee MD, Wagenmakers EJ. *Bayesian Cognitive Modeling: A Practical Course*. Cambridge University Press; 2014.

How to cite this article: Pulcini G. Discussion of “Specifying prior distributions in reliability applications”. *Appl Stochastic Models Bus Ind*. 2023;1-4. doi: 10.1002/asmb.2781