Revista Colombiana de Estadstica
July 2019, Volume 42, Issue 2, pp. 225 to 243
DOI: http://dx.doi.org/10.15446/rce.v42n2.76815

Bayesian Inference for the Segmented Weibull Distribution

Inferencia bayesiana para distribuciones Weibull segmentadas

Emílio A. Coelho-Barros[1,a], Jorge A. Achcar[2,b], Edson Z. Martinez[2,c], Nasser Davarzani[3,d], Heike I. Grabsch[4,e]

[1]Department of Mathematics, Federal University of Technology, Cornélio Procópio, Brazil

[2]Department of Social Medicine, Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, Brazil

[3]Department of Pathology, GROW School for Oncology and Developmental Biology, Maastricht University Medical Center, Maastricht, The Netherlands

[4]Section of Pathology & Tumour Biology, Leeds Institute of Cancer and Pathology, University of Leeds, Leeds, United Kingdom

Abstract

In this paper, we introduce a Bayesian approach for segmented Weibull distributions which could be a good alternative to analyze medical survival data in the presence of censored observations and covariates. With the obtained Bayesian estimated change-points we could get an excellent fit of the proposed model to any data sets. With the proposed methodology, it is also possible to identify survival times intervals where a covariate could have significantly different effects when compared to other lifetime intervals, an important point under a clinical view. The obtained Bayesian estimates are obtained using standard Markov Chain Monte Carlo methods. Some examples with real data sets illustrate the proposed methodology and its potential clinical value.

***Key words*:** Bayesian methods; Censored data; Change-points; Covariates; Segmented Weibull distribution.

[a] PhD. E-mail: eabarros@utfpr.edu.br
[b] PhD. E-mail: achcar@fmrp.usp.br
[c] PhD. E-mail: edson@fmrp.usp.br
[d] PhD. E-mail: n.davarzani@maastrichtuniversity.nl
[e] PhD. E-mail: h.i.grabsch@leeds.ac.uk

Resumen

En este artículo introducimos un nuevo modelo Bayesiano para distribuciones Weibull segmentadas, que puede ser una buena alternativa en el análisis de datos aplicados a la investigación en salud, con la presencia de censuras y covariables. Con este método basado en la estimación de puntos de cambio, hemos obtenido un excelente ajuste a los datos utilizados como ejemplos. De acuerdo con el modelo propuesto, fue posible identificar rangos de valores en las series temporales en que una variable independiente podría tener diferentes efectos. Este es un resultado importante desde el punto de vista clínico. Los estimados bayesianos fueron obtenidos usando métodos de Monte Carlo en Cadenas de Markov. Ejemplos basados en conjuntos de datos reales fueran usados para ilustrar el uso de los modelos propuestos y sus potenciales aplicaciones en investigaciones clínicas.

***Palabras clave*:** Covariables; Datos censurados; Distribución Weibull segmentada; Métodos bayesianos; Puntos de cambio.

1. Introduction

In medical research, survival analysis techniques are often used to study time to an event such as death or disease recurrence. A widely used model is based on the Weibull probability distribution function with two parameters for a survival time T. The Weibull distribution was first described in detail in 1951 by the Swedish mathematician Waloddi Weibull (Weibull 1951). Among other advantages of the Weibull distribution, we note that:

- It assumes different shapes due to the flexibility of its hazard rate function (increasing, decreasing or constant depending on the value of its shape parameter);
- It can be seen as a family incorporating other different survival curve shapes, such as the exponential and Rayleigh distributions;
- It can be easily be adjusted for covariates and applied to the regression models;
- Their parameters can be easily estimated (see for example, Pak, Parham & Saraj, 2013; or Kizilaslan & Nadar, 2015).

In clinical research studies it is usual to have the presence of one or more changes in the failure rate (Desmond, Weiss, Arani, Soong, Wood, Fiddian, Gnann & Whitley 2002, Sertkaya & Sözer 2003, Jandhyala, Fotopoulos & Evaggelopoulos 1999, Noura & Read 1990, Whiteley, Andrieu & Doucet 2011). These change-points could be results of treatment effect. Chen and Baron (2014) point out that change-points could occur in many medical applications as in zoster pain resolution trials (Desmond et al. 2002), where the treatment lightens pain from acute to subacute and then to chronic, resulting in three different failure rates; in another application, Zucker & Lakatos (1990), describe the effect of beta-carotene on cancer incidence where new tumors need time to become detectable while the treatment does not affect pre-existing tumors where there is an approximately two-year waiting period before the effect of the treatment is noticeable. Survival times in these examples have a higher initial failure rate and a lower failure rate afterwards (see other applications in, Goodman, Li & Tiwari 2011; He, Kong & Su 2013; Müller & Wang 1990).

In this case, the literature presents many papers with classical or Bayesian approaches to get inferences for a change-point assuming the exponential distribution which is a special case of the Weibull distribution. In this direction, Matthews & Farewell (1982) considered the problem of testing the hypothesis of the change-point be equal to zero based on the likelihood ratio test statistics and used simulations to find the distribution of the statistics of this model with application to medical data related to the treatment of leukemia patients.

Assuming the special case of an exponential distribution, the hazard function in presence of a change-point is given by,

$$\lambda\left(t\right)=\begin{cases}\lambda & \text{if } t<\zeta\\ \lambda\rho & \text{if } t\geq\zeta\end{cases}\tag{1}$$

where $T>0$ denotes the lifetime of an individual, λ and $\lambda\rho$ denote the rates before and after the change-point ζ and the parameter $\rho>0$ denotes the change in the hazard function (a discontinuous change-point).

The probability density function (pdf) for the lifetime T is given by,

$$f\left(t\right)=\begin{cases}\lambda\exp\left(-\lambda t\right) & \text{if } t<\zeta\\ \lambda\rho\exp\left(-\lambda t-\lambda\rho\left(t-\zeta\right)\right) & \text{if } t\geq\zeta\end{cases}\tag{2}$$

This model approch also could be generalized for situations in presence of a covariate. As a special situation, let us consider a covariate X related to two treatments, that is, each different treatment could lead to different change-points ($X=0$ for treatment 1 and $X=1$ for treatment 2) with hazard function,

$$\lambda\left(t\right)=\begin{cases}\lambda & \text{if } t<\zeta\\ \lambda\rho\exp\left(\beta x\right) & \text{if } t\geq\zeta\end{cases}\tag{3}$$

where β is a regression parameter (Achcar & Bolfarine 1989, Achcar & Loibel 1998).

Similarly, assuming a Weibull distribution, the hazard function in presence of a change-point is given by,

$$\lambda\left(t\right)=\begin{cases}\lambda\gamma t^{\gamma-1} & \text{if } t<\zeta\\ \lambda\rho\gamma t^{\rho\gamma-1} & \text{if } t\geq\zeta\end{cases}\tag{4}$$

and the pdf is given by,

$$f\left(t\right)=\begin{cases}\lambda\gamma t^{\gamma-1}\exp\left(-\lambda t^{\gamma}\right) & \text{if } t<\zeta\\ \lambda\rho\gamma t^{\rho\gamma-1}\exp\left[-\lambda\left(\zeta^{\gamma}+t^{\rho\gamma}-\zeta^{\rho\gamma}\right)\right] & \text{if } t\geq\zeta\end{cases}\tag{5}$$

Similarly, it is possible to generalize these expressions for the case of two or more change-points (Achcar, Rodrigues & Tzintzun 2011*a*, Achcar, Rodrigues & Tzintzun 2011*b*).

Also with a classical inference approach, Matthews, Farewell & Pyke (1985) considered an asymptotic score statistic process to test for constant hazard against a change-point alternative. In another paper, Nguyen, Rogers & Walker (1984) obtained a consistent estimator for the change-point by examining the properties of the density represented as a mixture. Yao (1986) proposed a maximum likelihood estimator for the change-point subject to a natural constraint and Worsley (1986) also used maximum likelihood methods to test for a change-point and found the exact null and alternative distributions of the test statistics. Loader (1991) discussed inference based on the likelihood ratio process for a hazard rate change-point and derived approximate confidence regions for the change-point.

An approach derived from the Kaplan-Meier estimation of the survival function followed by the least-squares estimation for the change-point was introduced by Chen & Baron (2014). Zhao, Wu & Zhou (2009) proposed a change-point model for survival data accounting for long-term survivors with application to the leukemia data analyzed by Matthews & Farewell (1982). Achcar & Bolfarine (1989) presented a Bayesian analysis of the exponential model assuming a change-point as either known or unknown. In another paper, Achcar & Loibel (1998) considered Bayesian inferences for the exponential model assuming change-point using different prior densities. Karasoy & Kadilar (2007) introduced another Bayesian approach for constant hazard functions applied to data of breast cancer patients and some lymphoma data. Assuming a Weibull distribution in presence of a single change-point, Jiwani (2005) introduced a single change-point parametric Weibull model, considering the case where the survival function is subject to a change from a given instant. Yiannoutsos (2009) used the model proposed by Jiwani (2005) to estimate survival among HIV-infected patients who are initiating antiretroviral therapy in sub-Saharan Africa.

In this paper, we introduce a Bayesian inference approach for the segmented Weibull model assuming one or more unknown change-points. For the Bayesian analysis of the model, we use standard existing Markov Chain Monte Carlo (MCMC) methods to simulate samples of the joint posterior distribution of interest. Applications are considered using clinical data sets in presence of change-points with continuous survival function.

2. Methods

Let T denoting a continuous non-negative random variable denoting a lifetime with probability density function (pdf) $f\left(t\right)$ and cumulative distribution function (cdf) $F\left(t\right) = P(T \leq t)$. Under the assumption of a Weibull distribution, the pdf is given by,

$$f\left(t\right) = \frac{\alpha}{\mu^{\alpha}} t^{\alpha-1} \exp\left[-\left(\frac{t}{\mu}\right)^{\alpha}\right], \tag{6}$$

where $t > 0$, $\mu > 0$ and $\alpha > 0$. The Weibull distribution is characterized by two parameters μ and α, where μ is a scale parameter and α is a shape parameter.

The survival function is given by,

$$S\left(t\right) = 1 - F\left(t\right) = \exp\left[-\left(\frac{t}{\mu}\right)^{\alpha}\right]. \tag{7}$$

The hazard function or the instantaneous rate of occurrence is given by,

$$h\left(t\right) = \frac{f\left(t\right)}{S\left(t\right)} = \frac{\alpha}{\mu^{\alpha}}t^{\alpha-1}. \tag{8}$$

The hazard function $h\left(t\right)$ is increasing if $\alpha > 1$, decreasing if $\alpha < 1$ and constant if $\alpha = 1$ (an exponential distribution).

The mean and variance of the Weibull distribution with density (6) are given respectively by,

$$E\left(T\right) = \mu\Gamma\left(1 + \frac{1}{\alpha}\right) \tag{9}$$

and,

$$Var\left(T\right) = \mu^2\left[\Gamma\left(1 + \frac{2}{\alpha}\right) - \Gamma^2\left(1 + \frac{1}{\alpha}\right)\right]. \tag{10}$$

The cumulative hazard function is given by,

$$\Lambda\left(t\right) = \int_0^t h\left(x\right)dx = \left(\frac{t}{\mu}\right)^{\alpha}, \tag{11}$$

that is,

$$g\left(t\right) = \ln\left[\Lambda\left(t\right)\right] = \alpha\ln\left(t\right) - \alpha\ln\left(\mu\right), \tag{12}$$

In this way the survival function is,

$$S\left(t\right) = P\left(T > t\right) = \exp\left[-\Lambda\left(t\right)\right]. \tag{13}$$

2.1. Presence of One Change-Point

In presence of a change-point a_1, we define for the i^{th} observation $(i = 1, 2, \dots, n)$,

$$g_1\left(t_i\right) = c_i\left[\alpha_1\ln\left(t_i\right) - \alpha_1\ln\left(\mu_1\right)\right] + \left(1 - c_i\right)\left[\alpha_2\ln\left(t_i\right) - \alpha_2\ln\left(\mu_2\right)\right], \tag{14}$$

where, $c_i = 1 - step\left(t_i - a_1\right)$, and $step\left(t_i - a_1\right) = 1$ if $t_i > a_1$; $step\left(t_i - a_1\right) = 0$ if $0 < t_i \leq a_1$.

In this way we have five parameters to be estimated: a_1, α_1, α_2, μ_1 and μ_2. The cumulative hazard function in presence of the change-point a_1, is given by,

$$\Lambda_1\left(t_i\right) = \exp\left[g_1\left(t_i\right)\right], \tag{15}$$

and the hazard function is given by,

$$h_1\left(t_i\right)=\frac{d\Lambda_1\left(t_i\right)}{dt_i}=\frac{dg_1\left(t_i\right)}{dt_i}\exp\left[g_1\left(t_i\right)\right], \tag{16}$$

where, $\frac{dg_1(t_i)}{dt_i}=\frac{1}{t_i}\left[c_i\alpha_1+\left(1-c_i\right)\alpha_2\right]$, that is,

$$h_1\left(t_i\right)=\frac{1}{t_i}\left[c_i\alpha_1+\left(1-c_i\right)\alpha_2\right]\exp\left[g_1\left(t_i\right)\right], \tag{17}$$

and,

$$S_1\left(t_i\right)=\exp\left[-\Lambda_1\left(t_i\right)\right]=\exp\left\{-\exp\left[g_1\left(t_i\right)\right]\right\}. \tag{18}$$

The density function is given, from $f_1\left(t_i\right)=h_1\left(t_i\right)S_1\left(t_1\right)$, by,

$$f_1\left(t_i\right)=\frac{1}{t_i}\left[c_i\alpha_1+\left(1-c_i\right)\alpha_2\right]\exp\left\{g_1\left(t_i\right)-\exp\left[g_1\left(t_i\right)\right]\right\}. \tag{19}$$

2.2. Likelihood Function in Presence of Censored Observations

Let $T_1,T_2,\ldots,T_n$ be a random sample of size n of lifetimes in presence of censored observations, where the observed times are given by $t_i=\min\left(T_i,C_i\right)$ (type I censoring) where C_i are the censoring times and $i=1,2,\ldots,n$; thus we can define the indicator variable,

$$\delta_i=\left\{\begin{array}{l}1\text{ (complete observation)}\\0\text{ (censoring observation).}\end{array}\right. \tag{20}$$

For the i^{th} individual, the contribution for the likelihood function is given by,

$$L_i=\left[h_1\left(t_i\right)\right]^{\delta_i}\exp\left[-\Lambda_1\left(t_i\right)\right]. \tag{21}$$

The log-likelihood function assuming only one change-point is given by,

$$l\left(\boldsymbol{\theta}\right)=\sum_{i=1}^{n}\delta_i\ln\left[h_1\left(t_i\right)\right]-\sum_{i=1}^{n}\exp\left[g_1\left(t_i\right)\right] \tag{22}$$

where, $\boldsymbol{\theta}=\left(a_1,\alpha_1,\alpha_2,\mu_1,\mu_2\right)$, $h_1\left(t_i\right)$ is given in (17) and $g_1\left(t_i\right)$ is given in (14), that is,

$$\begin{aligned}l\left(\boldsymbol{\theta}\right)=\sum_{i=1}^{n}\delta_i\ln\left(t_i\right)+\sum_{i=1}^{n}\delta_i\ln\left[c_i\alpha_1+\left(1-c_i\right)\alpha_2\right]&\\+\sum_{i=1}^{n}\delta_i g_1\left(t_i\right)-\sum_{i=1}^{n}\exp\left[g_1\left(t_i\right)\right]&\end{aligned} \tag{23}$$

***Note* 1.** For the case of one change-point, when we have the continuity for the survival function in the change-point $t = a_1$, we have: $\alpha_1 \ln(a_1) - \alpha_1 \ln(\mu_1) = \alpha_2 \ln(a_1) - \alpha_2 \ln(\mu_2)$, that is,

$$\alpha_2 \ln(\mu_2) = \alpha_2 \ln(a_1) - \alpha_1 \ln(a_1) + \alpha_1 \ln(\mu_1), \tag{24}$$

or,

$$\mu_2 = \exp\left[\frac{(\alpha_2 - \alpha_1)\ln(a_1) + \alpha_1 \ln(\mu_1)}{\alpha_2}\right]. \tag{25}$$

2.3. A Bayesian Analysis of the Model Assuming One Change-Point

For a Bayesian analysis of the Weibull distribution in presence of change-points, we assume $Gamma\,(b_1, b_2)$ prior distributions for the shape parameters α_1 and α_2 and for the scale parameter μ_1, where $Gamma\,(b_1, b_2)$ denotes a gamma distribution with mean $\frac{b_1}{b_2}$ and variance $\frac{b_1}{b_2^2}$; the scale parameter μ_2 was estimated using (25). For the change-point a_1, we assume a uniform prior distribution on the interval $(0, T_m)$, where T_m is the maximum value observed in the lifetime data. Further let us assume prior independence among the parameters.

Combining the joint prior distribution for $\boldsymbol{\theta} = (a_1, \alpha_1, \alpha_2, \mu_1, \mu_2)$ with the likelihood function $L(\boldsymbol{\theta})$, the posterior distribution for $\boldsymbol{\theta}$ is determined from the Bayes formula (Box & Tiao 1973). The posterior summaries of interest are obtained using Markov Chain Monte Carlo (MCMC) methods (Gelfand & Smith 1990, Chib & Greenberg 1995). A great simplification in the generation of samples from the posterior distribution for $\boldsymbol{\theta}$ is obtained by using the procedure MCMC (SAS Institute Inc 2016) from the software SAS (University Edition), which only requires the specification of the distribution for the data and a prior distribution for the parameters of the model.

Under our proposed model approach, it is observed that it is not possible to get explicit forms for the marginal posterior distributions for each parameter. In this way, we could use some approximation method to solve integrals as the Laplace method (Tierney, Kass & Kadane 1989) or some numerical method (Naylor & Smith 1982). An alternative is to use simulation methods like the Markov Chain Monte Carlo methodology (Gelfand & Smith 1990, Hastings 1970) or acceptation-rejection algorithms such as the Adaptive Rejection Sampling (ARS) or the Adaptive Rejection Metropolis Sampling (ARMS) (Devroye 1986).

Monte Carlo Markov chains are becoming a standard way to simulate posterior summaries of interest that allows us to solve a wide range of problems (Tierney 1994). To simulate samples of the joint posterior distribution of interest, we need the full conditional posterior distribution for each parameter, from where it is used the Gibbs sampling algorithm (see, for example, Gelfand and Smith, 1990) when these conditional distribution are simple to simulate samples.

In this way, we follow the algorithm,

Step 1 Choose initial estimates $a_1^{(0)}$, $\alpha_1^{(0)}$, $\alpha_2^{(0)}$, $\mu_1^{(0)}$ and $\mu_2^{(0)}$.

Step 2 Given current estimates $a_1^{(i)}$, $\alpha_1^{(i)}$, $\alpha_2^{(i)}$, $\mu_1^{(i)}$ and $\mu_2^{(i)}$ simulate new values:

- $a_1^{(i+1)}$ from $\pi\left(a_1|\alpha_1^{(i)},\alpha_2^{(i)},\mu_1^{(i)},\mu_2^{(i)},\mathbf{t},\delta\right)$.
- $\alpha_1^{(i+1)}$ from $\pi\left(\alpha_1|a_1^{(i+1)},\alpha_2^{(i)},\mu_1^{(i)},\mu_2^{(i)},\mathbf{t},\delta\right)$.
- $\alpha_2^{(i+1)}$ from $\pi\left(\alpha_2|a_1^{(i+1)},\alpha_1^{(i+1)},\mu_1^{(i)},\mu_2^{(i)},\mathbf{t},\delta\right)$.
- $\mu_1^{(i+1)}$ from $\pi\left(\mu_1|a_1^{(i+1)},\alpha_1^{(i+1)},\alpha_2^{(i+1)},\mu_2^{(i)},\mathbf{t},\delta\right)$.
- $\mu_2^{(i+1)}$ from $\pi\left(\mu_2|a_1^{(i+1)},\alpha_1^{(i+1)},\alpha_2^{(i+1)},\mu_1^{(i+1)},\mathbf{t},\delta\right)$.

Step 3 Return to step 2.

The sequence $\left(a_1^{(i)},\alpha_1^{(i)},\alpha_2^{(i)},\mu_1^{(i)},\mu_2^{(i)}\right)$ $i=1,\ldots,L$ is a realization of a Markov chain which, under mild regular conditions, has an equilibrium distribution $\pi\left(a_1,\alpha_1,\alpha_2,\mu_1,\mu_2 \mid \mathbf{t},\delta\right)$, the joint posterior distribution of a_1, α_1, α_2, μ_1 and μ_2.

However, in the case that the conditional posterior densities for the parameters show that standard sampling schemes are not feasible since the conditional distributions are not in a known form, Bayesian inference for the parameters can be obtained using the Metropolis-Hastings algorithm (Chib & Greenberg 1995) considering the conditional distributions as the target densities.

2.4. Presence of More Than One Change-Point

Once a first change-point a_1 was estimated (denoted as stage 1), we could use the Bayesian approach to search for a second change-point a_2 based on the information that the first change-point is a known quantity a_1 (the Bayesian estimate based on a square error loss function of the first change-point a_1):

(1) In this way, we assume a uniform prior distribution $U\left(a_1,T_m\right)$ for the second change-point where T_m is the maximum value observed in the lifetime data. We also assume that for $a_1 < t < a_2$ the values of α_1 and μ_1 are assumed to be known and equal to the Bayesian estimates for α_2 and μ_2 obtained in the stage 1. This guarantee the continuity of the Weibull segmented survival function.

(2) Once the second change-point is estimated (denoted as stage 2), we assume a uniform $U\left(a_2,T_m\right)$ prior for the third change-point. We also assume that for $a_2 < t < a_3$ the values of α_1 and μ_1 are assumed to be known and equal to the Bayesian estimates for α_2 and μ_2 obtained in the stage 2. We continue this procedure until it is not possible to estimate more change-points.

(3) Observe that the procedure described by (1) and (2) assumes that the first change-point is smaller than the second change-point and the second change point is smaller than the third change-point. It is important to point out that the method also could be used in other situations. Assuming the second

change-point smaller than the first change-point we could assume an uniform $U(0, a_1)$ prior in place of an uniform prior $U(a_1, T_m)$ in (1).

3. Applications With Real Data

In this section we present two applications with real data sets. First we consider a data set presented in Chapter 2 of the book of Hosmer, Lemeshow & May (2008). The second data was obtained from a trial conducted by the Leeds Teaching Hospitals NHS Trust, England.

3.1. BPD Data Set

As a first application, we consider a data set (BPD Data) introduced in Chapter 2 of the book of Hosmer et al. (2008). This data set have 78 observations and we consider two variables; Days on Oxygen (time to event) and the Censoring Indicator: 1 =Off Oxygen and 0 =Still on Oxygen. In Figure 1, we have the plot of the (Kaplan & Meier 1958) nonparametric estimate of the survival function (use of the R software). From this graph we observe the indication of a possible first change-point close to time $t = 100$ and a possible second change-point close to $t = 500$.

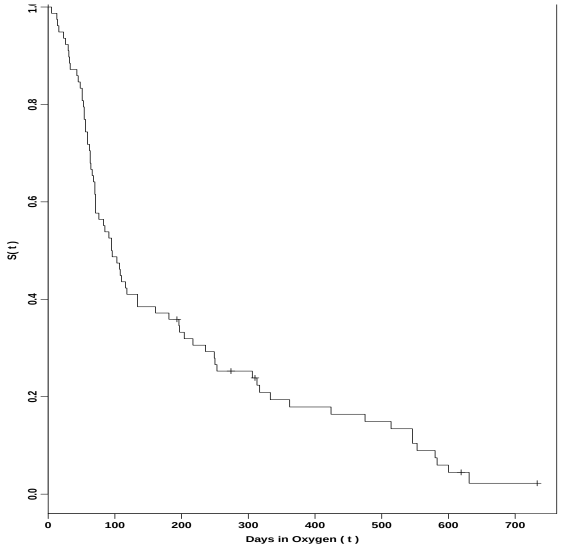


Figure 1: Kaplan Meier estimate for the survival function (BPD data).

For a Bayesian analysis of the model in a first stage, let us assume a Weibull distribution with density (6) in the presence of a change-point and assuming continuity of the survival function at the change-point, see (25). Assuming *Gamma*(0.01, 0.01) prior distributions for the shape parameters α_1 and α_2 and for the scale parameter μ_1; we consider an uniform $U(0, 733)$ prior distribution for the change-point a_1 (first stage of the Bayesian analysis). We have used the procedure MCMC (SAS Institute Inc 2016) from the software SAS (University Edition), a single chain has been used in the simulation of samples for the parameters considering a "burn-in-sample" of size $10,000$ to eliminate the possible effect of the initial values. After this "burn-in" period, we simulated other $500,000$ Gibbs samples taking every $200th$ sample, to get approximated uncorrelated values which result in a final chain of size $2,500$. Convergence of the algorithm was verified from trace plots of the simulated samples for each parameter and usual existing convergence diagnostics available in the literature for a single chain using the SAS/MCMC procedure indicated convergence for all parameters. In Table 1, we have the posterior summaries of interest (first segment of the Weibull distribution).

Table 1: Posterior summaries (Weibull with a change-point).

Parameter	Mean	Standard Deviation	95% Credible Interval
a_1	84.60	17.816	$(68.40; 125.0)$
α_1	1.846	0.3669	$(1.219; 2.643)$
α_2	0.739	0.1035	$(0.542; 0.941)$
μ_1	111.2	17.407	$(85.65; 151.1)$
μ_2	166.8	30.332	$(111.2; 230.6)$

For a second change-point a_2 (second stage of the analysis) we assume a uniform $U\,(84.6, 733)$ prior for a_2, $\alpha_1 = 0.739$, $\mu_1 = 166.8$ obtained from the first stage, and the same prior distributions for the other parameters α_2 and μ_2 considered in the first stage. Using the same simulation steps used in the estimation of the parameters in the first stage (results in Table 1), we have in Table 2, the posterior summaries of interest (second segment of the Weibull distribution).

Table 2: Posterior summaries (Weibull second change-point).

Parameter	Mean	Standard Deviation	95% Credible Interval
a_2	534.5	75.812	$(364.8; 715.0)$
α_2	2.302	1.1970	$(0.743; 5.432)$
μ_2	348.2	89.793	$(166.9; 572.6)$

Since the second change-point ($a_2 = 534.5$) is a large value close to the lifetime $t = 733$ (maximum observed lifetime), we stop the search method looking for new change-points. In this way, the survival function could be splited in three segmented Weibull pieces:

$$S\,(t) = \begin{cases} \exp\left[-\left(\frac{t}{111.2}\right)^{1.846}\right] \text{ if } 0 < t < 84.6 \\ \\ \exp\left[-\left(\frac{t}{166.8}\right)^{0.739}\right] \text{ if } 84.6 \leq t < 534.5 \\ \\ \exp\left[-\left(\frac{t}{348.2}\right)^{2.302}\right] \text{ if } t \geq 534.5 \end{cases} \tag{26}$$

In Figure 2, we have the plots of the Kaplan-Meier survival curve; the estimated survival curve assuming a Weibull distribution in the presence of two change-points, and also assuming a Weibull distribution not considering the presence of change-points. We clearly observe better fit of the Weibull distribution in the presence of two change-point for the BPD data set.

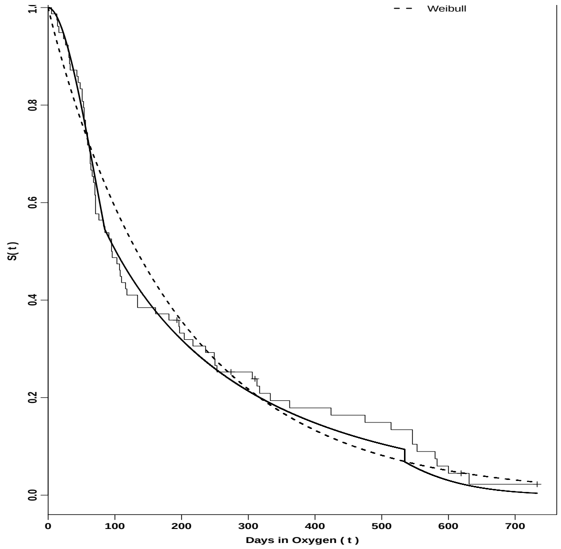


Figure 2: Bayesian estimates for the survival function (BPD data).

3.2. Leeds Gastric Cancer Series

In second application, we evalue a data set with 906 patients who were diagnosed with gastric cancer and underwent surgery between 1968 and 2009 at the Leeds Teaching Hospitals NHS Trust, Leeds, UK. The overall survival time of 861 patients is available as time to death or censoring time (in years). Lymph node metastasis status in gastric cancer (pN category) was measured for each patient, were $pN = 0$, no lymph node metastases; $pN = 1$, otherwise. The pN category is a very well established prognostic factor of gastric cancer (Deng & Liang 2014) and one might be interested to analyze pN category in terms of survival time of patients.

In this data set there are 861 patients where 222 lifetimes are right censored data. The median follow-up time was 1.67 years, ranging from 0.01 to 20.56 years.

First of all, we assume a lifetime regression model (Weibull regression model) for the response (overall survival) in the presence of covariate gender (1: male; 0: female) and pN (pN = 0, no lymph node metastases; pN = 1, otherwise) given by,

$$\ln(t_i) = \beta_0 + \beta_1 gender_i + \beta_2 pN_i + \varepsilon_i, \tag{27}$$

where, t_i denote the overall survival for the i^{th} patient, $i = 1, \dots, 861$; β_0, β_1 and β_2 are the regression parameters. Assume the error term ε_i in (27) is a random quantity with an extreme value distribution (Lawless 2003) with density,

$$f(\varepsilon) = \exp\left[\varepsilon - \exp(\varepsilon)\right], \quad -\infty < \varepsilon < \infty, \tag{28}$$

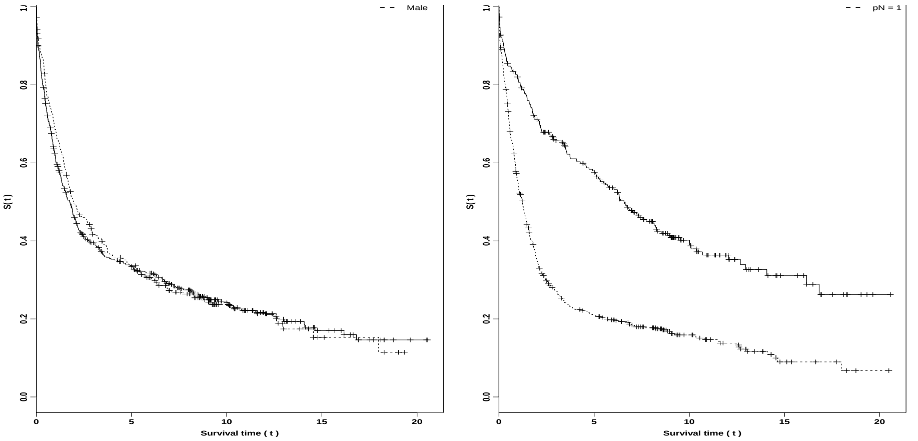
then we have a Weibull regression model. Other distributions also could be assumed for the error. If the error term ε_i in (27) has a standard normal distribution, we have a log-normal distribution for the lifetime T.

Using the procedure MCMC from the software SAS (University Edition), we have in Table 3, the Bayesian estimates of the regression parameters assuming a Weibull distribution with $n = 861$ observations (639 complete lifetimes and 222 censored lifetimes). We consider normal priors with mean $\mu = 0$ and variance $\sigma^2 = 10000$ for the regression parameters β_0, β_1 and β_2 and a $Gamma\,(0.01, 0.01)$ prior for the Weibull shape parameter α. We consider a single chain in the simulation of samples for the parameters considering a "burn-in-sample" of size $10,000$ to eliminate the possible effect of the initial values. After this "burn-in" period, we simulated other $500,000$ Gibbs samples taking every $200th$ sample, to get approximated uncorrelated values which result in a final chain of size $2,500$. Convergence of the algorithm was verified from trace plots of the simulated samples for each parameter and usual existing convergence diagnostics available in the literature for a single chain using the SAS/MCMC procedure indicated convergence for all parameters.

From the results of Table 3, we observe that the covariate gender has not significant effect in the overall survival times (zero is included in 95% credible interval), but the covariate pN presents significative effect on the overall lifetimes of the patients. In Figure 3, we have the plots of the Kaplan-Meier nonparametric estimate of the survival functions.

Table 3: Posterior summaries for parameters (Weibull regression model with covariates gender and pN).

Parameter	Mean	Standard Deviation	95% Credible Interval
β_0	2.4404	0.1375	$(2.1793; 2.7100)$
β_1	0.0651	0.1364	$(-0.217; 0.3300)$
β_2	-1.462	0.1523	$(-1.767; -1.168)$
α	0.6063	0.0197	$(0.5659; 0.6438)$



(a) Gender

(b) pN

FIGURE 3: Kaplan-Meier estimate for the survival functions (gender and pN).

From the plots of Kaplan-Meier presented in Figure 3, we observe that is an indication of a change-point close to the survival time $t = 2$. In this way, we will assume a segmented Weibull distribution with a change-point considering all data set.

3.2.1. Change-Point for the Survival Function Not Considering the Presence of Covariates

In Figure 4, we have the plot of the Kaplan-Meier nonparametric estimate of the survival functions not considering the presence of covariates. From the graph of Figure 4, we observe a possible change-point close to time $t = 2$. In this case, we have an indication of only one change-point. For a Bayesian analysis of the model, let us assume a Weibull distribution with density (6) in the presence of a change-point assuming continuity of the survival function at the change-point. Let us assume $Gamma\,(0.01, 0.01)$ prior distributions for the shape parameters α_1 and α_2 and for the scale parameter μ_1 and an uniform $U\,(0, 20.56)$ prior distribution for the change-point a_1 (first stage of the Bayesian analysis). Using the same simulation steps used in the estimation of the parameters in the Weibull regression model (results in Table 3), we have in Table 4, the posterior summaries of interest (a segmented Weibull distribution).

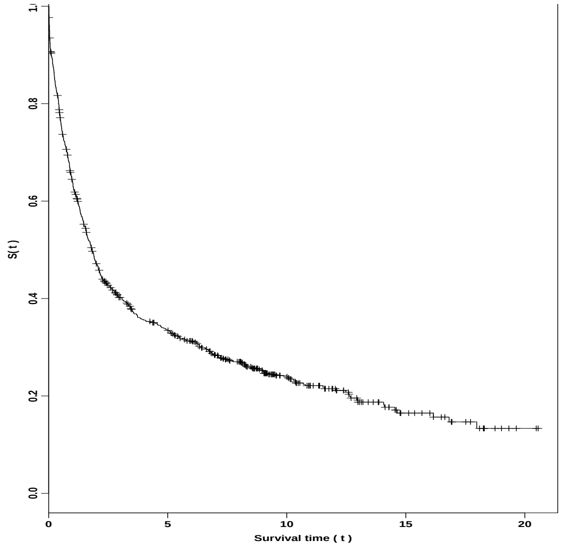


Figure 4: Kaplan Meier estimate for the survival function (not considering covariates)

Table 4: Posterior summaries (a change-point segmented Weibull).

Parameter	Mean	Standard Deviation	95% Credible Interval
a_1	2.2609	0.0866	$(2.0975; 2.4622)$
α_1	0.7265	0.0314	$(0.6678; 0.7881)$
α_2	0.3938	0.0267	$(0.3422; 0.4467)$
μ_1	3.0203	0.2096	$(2.6443; 3.4693)$
μ_2	3.8504	0.4119	$(3.0841; 4.6985)$

In this way, the survival function could be splited in two segmented Weibull pieces:

$$S\left(t\right)=\begin{cases}\exp\left[-\left(\frac{t}{3.0203}\right)^{0.7265}\right] \text{ if } 0<t<2.2609\\ \\ \exp\left[-\left(\frac{t}{3.8504}\right)^{0.3938}\right] \text{ if } t\geq 2.2609\end{cases} \tag{29}$$

In Figure 5, we have the plots of the Kaplan-Meier survival curve; the estimated survival curve assuming a Weibull distribution in the presence of a change-point, and also assuming a Weibull distribution not considering the presence of change-points. We observe an excellent fit of the segmented Weibull distribution considering a change-point for the data.

3.2.2. Effect of the Covariate pN On the Survival Probabilities Considering the Segmented Weibull Distribution

In the presence of a change-point, let us assume a Weibull segmented distribution with change-point equals to 2.2609 in presence of only the covariate pN. In

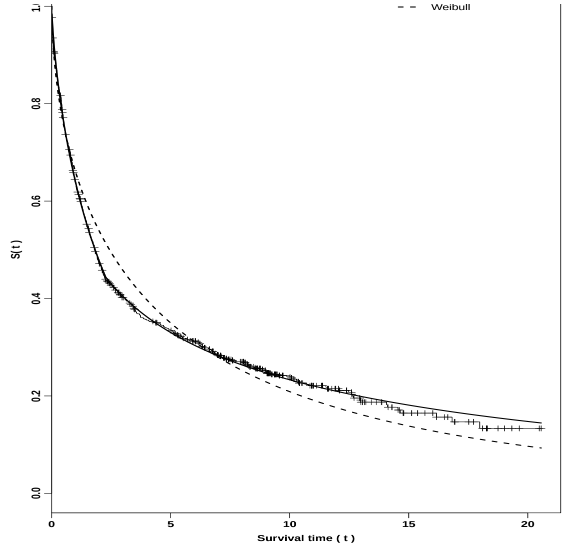


FIGURE 5: Bayesian estimates for the survival function (NHS data).

Tables 5 and 6, it is presented the Bayesian estimates of the regression parameters assuming a Weibull distribution in presence of the covariate pN and the change-point $a_1 = 2.2609$. We consider the same priors and simulation steps used in the estimation of the parameters in the Weibull regression model (results in Table 3).

TABLE 5: Posterior summaries for parameters (segmented Weibull regression model with covariate pN and survival < 2.2609).

Parameter	Mean	Standard Deviation	95% Credible Interval
β_0	-0.052	0.1071	$(-0.254; 0.1624)$
$\beta_{t<2.2609}$	-0.119	0.1188	$(-0.353; 0.1039)$
α	0.9898	0.0383	$(0.9173; 1.0664)$

TABLE 6: Posterior summaries for parameters (segmented Weibull regression model with covariate pN and survival ≥ 2.2609).

Parameter	Mean	Standard Deviation	95% Credible Interval
β_0	2.7931	0.0836	$(2.6384; 2.9630)$
$\beta_{t\geq 2.2609}$	-0.287	0.1072	$(-0.499; -0.078)$
α	1.4816	0.0962	$(1.2978; 1.6791)$

From Tables 5 and 6, we observe that for survival times less than 2.2609 years there was no statistical difference between survival of patients with lymph node metastases and patients with no lymph node metastases (zero is included in 95% credible interval). After 2.2609 years, there was a statistical difference between survival of patients with lymph node metastases and patients with no lymph node metastases (zero is not included in 95% credible interval) where it is observed that patients with lymph node metastases are 1.53 ($HR = exp\,(0.287 \times 1.4816)$) times more in the risk of death than patients with no lymph node metastases.

We can also observe the effect of the covariate pN for survival times < 2.2609 and survival times ≥ 2.2609 considering the uncensored observations by observing the sample means in each case (see Table 7).

Table 7: Sample means for uncensored lifetimes considering the covariate pN assuming a change-point 2.2609.

	$pN = 0$	$pN = 1$	% gain
	(no lymph node metastases)	(lymph node metastases)	in survival
$t < 2.2609$	0.85772	0.78979	8.6%
$t \geq 2.2609$	6.50545	5.51191	18.02%

4. Concluding Remarks

The use of segmented Weibull distributions could be a good alternative to analyze survival times, since with this methodology, we could fit different survival functions and correspondent hazard functions for lifetimes with any hazard function shape. The Bayesian approach introduced in this paper does not require sophisticated computational expertize specially using MCMC simulation methods and the free available software SAS (University Edition).

From the results of this study, we also could get useful interpretations for clinical medical survival data assuming segmented Weibull distributions. Some important points:

- With the use of standard Kaplan-Meier estimation for survival curves it is only possible to compare the survival times together; similarly, if we assume Weibull or other parametric lifetime distributions not considering the presence of change-points;

- The use of a segmented Weibull distributions in presence of change-points gives better estimates for the survival curves than using a Weibull distribution without considering the presence of change-points;

- By estimating the change-points, we can compare the survival curves in different intervals which might tell us that the survivals are different in one time interval and not different in another interval, as we have observed in the second application with real data set (Leeds Gastric Cancer Series).

These results can be of great interest in medical applications.

Acknowledgements

The authors are grateful for the suggestions of the editor and referees that led to a great improvement in the manuscript.

[Received: December 2018 — Accepted: May 2019]

References

Achcar, J. A. & Bolfarine, H. (1989), 'Constant hazard against a change-point alternative: a Bayesian approach with censored data', *Communications in Statistics-Theory and Methods* **18**(10), 3801–3819.

Achcar, J. A. & Loibel, S. (1998), 'Constant hazard function models with a change point: A Bayesian analysis using Markov chain Monte Carlo methods', *Biometrical journal* **40**(5), 543–555.

Achcar, J. A., Rodrigues, E. R. & Tzintzun, G. (2011*a*), 'Modelling interoccurrence times between ozone peaks in Mexico City in the presence of multiple change points', *Brazilian Journal of Probability and Statistics* **25**(2), 183–204.

Achcar, J. A., Rodrigues, E. R. & Tzintzun, G. (2011*b*), 'Using non-homogeneous poisson models with multiple change-points to estimate the number of ozone exceedances in Mexico City', *Environmetrics* **22**(1), 1–12.

Box, G. E. P. & Tiao, G. C. (1973), *Bayesian inference in statistical analysis*, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont. Addison-Wesley Series in Behavioral Science: Quantitative Methods.

Chen, X. & Baron, M. (2014), 'Change-point analysis of survival data with application in clinical trials', *Open Journal of Statistics* **4**(09), 663–677.

Chib, S. & Greenberg, E. (1995), 'Understanding the metropolis-hastings algorithm', *The American Statistician* **49**(4), 327–335.

Deng, J.-Y. & Liang, H. (2014), 'Clinical significance of lymph node metastasis in gastric cancer', *World Journal of Gastroenterology* **20**(14), 3967–3975.

Desmond, R. A., Weiss, H. L., Arani, R. B., Soong, S.-j., Wood, M. J., Fiddian, P. A., Gnann, J. W. & Whitley, R. J. (2002), 'Clinical applications for change-point analysis of herpes zoster pain', *Journal of pain and symptom management* **23**(6), 510–516.

Devroye, L. (1986), *Non-Uniform Random Variate Generation*, Springer-Verlag, New York.

Gelfand, A. E. & Smith, A. F. M. (1990), 'Sampling-based approaches to calculating marginal densities', *Journal of the American Statistical Association* **85**, 398–409.

Goodman, M. S., Li, Y. & Tiwari, R. C. (2011), 'Detecting multiple change points in piecewise constant hazard functions', *Journal of applied statistics* **38**(11), 2523–2532.

Hastings, W. K. (1970), 'Monte Carlo sampling methods using Markov chains and their applications', *Biometrics* **57**, 97–109.

He, P., Kong, G. & Su, Z. (2013), 'Estimating the survival functions for right-censored and interval-censored data with piecewise constant hazard functions', *Contemporary clinical trials* **35**(2), 122–127.

Hosmer, D. W., Lemeshow, S. & May, S. (2008), *Applied survival analysis: regression modeling of time to event data*, Wiley-Interscience.

Jandhyala, V., Fotopoulos, S. & Evaggelopoulos, N. (1999), 'Change-point methods for weibull models with applications to detection of trends in extreme temperatures', *Environmetrics: The official journal of the International Environmetrics Society* **10**(5), 547–564.

Jiwani, S. L. (2005), Parametric changepoint survival model with application to coronary artery bypass graft surgery data, PhD thesis, Statistics and Actuarial Science department, Simon Fraser University, Canada.

Kaplan, E. L. & Meier, P. (1958), 'Nonparametric estimation from incomplete observations', *Journal of the American Statistical Association* **53**, 457–481.

Karasoy, D. S. & Kadilar, C. (2007), 'A new Bayes estimate of the change point in the hazard function', *Computational statistics & data analysis* **51**(6), 2993–3001.

Kizilaslan, F. & Nadar, M. (2015), 'Classical and bayesian estimation of reliability inmulticomponent stress-strength model based on weibull distribution', *Revista Colombiana de Estadística* **38**(2), 467–484.

Lawless, J. F. (2003), *Statistical models and methods for lifetime data*, Wiley Series in Probability and Statistics, second edn, John Wiley & Sons, Hoboken, NJ.

Loader, C. R. (1991), 'Inference for a hazard rate change point', *Biometrika* **78**(4), 749–757.

Matthews, D. E. & Farewell, V. T. (1982), 'On testing for a constant hazard against a change-point alternative', *Biometrics* **38**(2), 463–468.

Matthews, D., Farewell, V. & Pyke, R. (1985), 'Asymptotic score-statistic processes and tests for constant hazard against a change-point alternative', *The Annals of Statistics* **13**(2), 583–591.

Müller, H.-G. & Wang, J.-L. (1990), 'Nonparametric analysis of changes in hazard rates for censored survival data: An alternative to change-point models', *Biometrika* **77**(2), 305–314.

Naylor, J. C. & Smith, A. F. M. (1982), 'Applications of a method for the efficient computation of posterior distributions', *Journal of the Royal Statistical Society, C* **31**, 214–225.

Nguyen, H., Rogers, G. & Walker, E. (1984), 'Estimation in change-point hazard rate models', *Biometrika* **71**(2), 299–304.

Noura, A. & Read, K. (1990), 'Proportional hazards changepoint models in survival analysis', *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **39**(2), 241–253.

Pak, A., Parham, G. A. & Saraj, M. (2013), 'Inference for the weibull distribution based on fuzzy data', *Revista Colombiana de Estadística* **36**(2), 337–356.

SAS Institute Inc (2016), *SAS/STAT® 14.2 User's Guide, The MCMC Procedure*, Cary, NC: SAS Institute Inc.

Sertkaya, D. & Sözer, M. T. (2003), 'A bayesian approach to the constant hazard model with a change point and an application to breast cancer data', *Hacettepe Journal of Mathematics and Statistics* **32**, 33–41.

Tierney, L. (1994), 'Markov chains of exploring posterior distributions', *Annals of Statistics* **22**, 1701–1762.

Tierney, L., Kass, R. E. & Kadane, J. B. (1989), 'Fully exponential Laplace approximations to expectations and variances of nonpositive functions', *Journal of the American Statistical Association* **84**(407), 710–716.

Weibull, W. (1951), 'A statistical distribution function of wide applicability', *Journal of Applied Mechanics* **18**, 293–297.

Whiteley, N., Andrieu, C. & Doucet, A. (2011), 'Bayesian computational methods for inference in multiple change-points models'. Discussion paper, University of Bristol, UK.

Worsley, K. (1986), 'Confidence regions and tests for a change-point in a sequence of exponential family random variables', *Biometrika* **73**(1), 91–104.

Yao, Y. C. (1986), 'Maximum likelihood estimation in hazard rate models with a change-point', *Communications in Statistics-Theory and Methods* **15**(8), 2455–2466.

Yiannoutsos, C. T. (2009), 'Modeling aids survival after initiation of antiretroviral treatment by Weibull models with changepoints', *Journal of the International AIDS Society* **12**(1), 1–10.

Zhao, X., Wu, X. & Zhou, X. (2009), 'A change-point model for survival data with long-term survivors', *Statistica Sinica* **19**, 377–390.

Zucker, D. M. & Lakatos, E. (1990), 'Weighted log rank type statistics for comparing survival curves when there is a time lag in the effectiveness of treatment', *Biometrika* **77**(4), 853–864.