© Universiti Tun Hussein Onn Malaysia Publisher's Office

JST

Journal of Science and Technology

Journal homepage: http://penerbit.uthm.edu.my/ojs/index.php/jst
ISSN : 2229-8460 e-ISSN : 2600-7924

A Bayesian Parametric Estimation of Beta Kumaraswamy Burr Type X (Beta Kum-BX) Distribution Based on Cure Models with Covariates

Umar Yusuf Madaki[1*], Mohd Rizam Abu Bakar[2]

[1]Deparment of Mathematics and Statistics,
Faculty of Science, Yobe State University, P.M.B 1144, Damaturu, NIGERIA

[2]Deparment of Mathematics and Statistics,
Faculty of Science, Universiti Putra Malaysia, 43400 Seri Kembangan, Selangor, MALAYSIA

*Corresponding Author

DOI: https://doi.org/10.30880/jst.2022.14.01.001
Received 4 May 2021; Accepted 21 April 2022; Available online 28 June 2022

Abstract: In statistical models for censored survival data which includes a proportion of individuals who are not subject to the event of interest under study are known as the long-term survival cured models. It has two most adopted and common models used in estimating the cure fraction namely: the mixture (standard cure) and the non-mixture models. In this research work, we introduce a Bayesian approach using the two models for survival data based on the Beta Kumaraswamy Burr Type *X* distribution with six parameters and compared with two existing models: beta-Weibull and beta-generalized exponential distributions in analyzing a real-life dataset. The proposed approach allows the inclusion of covariates in the model. The parameter estimation was obtained by maximum likelihood and Bayesian analysis methods. The win Bugs and MCMC pack library in R softwares were employed for the Gibbs sampling algorithm in other to obtain the posterior summaries of interest and also the trace plots by the applying of real data sets and a simulation study was done based on cure models to compare the performance of both models relating to actual sense of motivation and novelty which clarifies the usefulness of the proposed methodologies.

Keywords: Bayesian analysis, Beta Kumaraswamy Burr Type *X* distribution, cure fraction models, survival analysis, maximum likelihood estimation

1. Introduction

In a real-life data analysis nowadays the use of parametric models for survival data analysis has been increasing in the last few decades in response to more refined statistical tools to be able to analyze complex data structures and parameters. The most standard existing techniques employed methods for estimating and analyzing survival data include, the Cox proportional hazard models, the Kaplan-Meier method for estimating the survival and log-rank test for testing the comparison of survival function between two or more sample groups. It is also nonparametric test violating the normality assumptions for use when the data are right skewed and censored.

We were motivated to proposed a new model called Beta Kumaraswamy Burr type X (Beta Kum-BX) distribution with six parameters proposed by Madaki, Bakar and Chakraborty (2016), due the highly efficient flexibility of these three confound models property which provides an enticing model fittings at different level and kinds of large datasets. It is also a very flexible and versatile model having some special sub models

*Corresponding author: uymadaki@ysu.edu.ng

distributions properties where its density function can be transformed as the so-called sub-model adopting some features and properties of Kumaraswamy Marshal-Olkin family of distributions Alizadeh, Cordeiro, Mansoor and Zubair (2015) and Kumaraswamy Burr type X model from the family of generated Kumaraswamy-G proposed by Cordeiro and Castro (2011). As a special case and modification to Burr type X model by I. W. Burr (1942) and Merovci, Khaleel, Ibrahim and Shitan (2016) Although, the BX model density and cumulative distribution functions have a simple close form and it's also having a convenient and flexible feature in modeling censored (incomplete) data, unlike Gamma, Generalized Exponential and log-Gamma distributions respectively.

1.1 Burr Type X Distribution

The two-parameter BX has a monotonically increasing and decreasing hazard function features, which can be used for practical aspects in statistical distribution and modeling of applications by I. W. Burr (1942). The flexibility of its failure rate and the ease for estimation of its parameters, ever since it has been widely applied in modelling real-life data. One of the limitations in beta-Kumaraswamy Burr type X (Kum-BX) model is that its functions cannot be prove in a closed form, specifically when more covariates are considered, thus numerical approach that is the integration techniques are required to determine the parameter estimation of in the models with or without covariates by some authors like: Madaki and Bakar (2016) & Madaki, Bakar, Ibrahim, Arasan and Hussein (2016) respectively.

2. Cure Rate Models

The cure fraction model Achcar, Barros and Mazucheli (2012), is usually called an extension to the survival cure models who might probably not experience the event. It can also be called a long-term survival model according to the kind of event is specified. The two most common cure models or long-term survivors are the standard parametric cure (mixture) and non-mixture models. In both formulations, it is introduced in the model for a parameter related to the cure fraction. There are some instances, especially with the advancements in modern medicine, in which a proportion of the population of interest is "cured" and will therefore never experience the event of interest. This situation motivates the incorporation of a cure fraction in a statistical model in order to analyze the ability of a certain treatment to cure a disease of interest. Once that model is defined, the next step is to develop procedures to fit the model to study datasets by utilizing popular statistical software. In the literature, an extensive volume of articles on modelling survival data including the cure models can be found refer to some of the authors like: Achcar, Barros and Mazucheli (2012) & Martinez, Achcar, Jacome and Santos (2013). On the other hand, Berkson and Gage (1952), also Boag (1949) proposed this vital used cure fraction models in survival analysis. The mixture cure rate model assumes that the studied population is a mixture of susceptible individuals, who experience the $p\epsilon[0,1]$ which is the proportion of the uncured susceptible individuals in the cure population regarding the event of interest ($0 < p < 1$) and non-susceptible individuals that will never experience it ($1 - p$), The survival function for the entire population, denoted by $S(t)$ for this model is given by:

$$S(t) = p + (1 - p)So(t) = p + F_o(t),\ t > 0. \qquad (1)$$

where, $S_o(t)$ is the standard parametric survival curve function for the susceptible individuals also $F_o(t)$ is the improper cumulative distribution for the cured population. In estimating the" improper" survival function $S(t)$, it is quite easy in the sense that $S(\infty) > 0$ if $p < 1$. The mixture cure model in equation (1) has been discussed by other researchers in the literature, including: Madaki and Bakar (2016), proposed a comprehensive methodology using R-software and Winbugs which explains the frequentist methods in the inferential aspect of cure fraction models. Although, this model received well attention over the years but one of its shortcomings comes when the covariates were modelled using the proportion p which does not involve a proportional hazard (PH) model, as it is the property of survival model which corresponds to the models' involving covariates. On the other hand, Martinez, Achcar, Jacome and Santos (2013), makes some observations in equation (1), as it provides an improper posterior result which includes the non-informative prior (uniform). It is actually the main reason for its drawback in the recent years which makes its tedious to apply the Bayesian technique with non-informative priors which is very common nowadays.

It was recently Bayesian estimation based on the disadvantage using equation (1), was observe by Chen et al. (1999) based on the alternative model called the" non-mixture cure model" defines an asymptote for the cumulative hazard, and hence for the cure fraction Ibrahim, Taweab and Arasan (2014) explains about the cure of cancer study among patients. The survival function can be written as:

$$S(t) = p^{Fo(t)} = \exp(\ln(p)F_o(t)). \qquad (2)$$

where, $F_o(t)$ is the improper cumulative distribution. The non-mixture cure model, was also pointed out by Ibrahim et al. (2001) and Chen et al. (1999), making used of the Bayesian approach technique. They also observed that in equation (2), does not involve any proportional hazard (PH) models or structures, if the $S(t)$ is modelled by the PH structure.

The drawback of the mixture model gave by Maller and Zhou (1996) is the idea on how to overcome with a great solution to incorporate the non-parametric Cox proportional hazard models with cure probability which can be written as:

$$\Lambda(t) = \Lambda_o(t) \exp\left(w^T \sigma\right). \tag{3}$$

where, $\Lambda_o(t)$ is the baseline hazard function for the improper survival function and" ω" is the covariate to be incorporated. When the baseline $\Lambda_o(t)$ is considers, this implies $1-F_o(t)$, and the survival function for model in equation (3) gives,

$$S(t) = \left(1 - F_o(t)\right) \exp(w^T \sigma). \tag{4}$$

The improper PH model was investigated by Famoye et al. (2005), to solve the model based on the structure of model in equation (4).

3. Related Work

A cure rate model based on the beta-Weibull distribution was proposed by Cordeiro et al. (2011), the techniques of estimating the cure rates when there are partially observed the missing covariates. Based on the long-term survival models that the usefulness of beta-Weibull distribution for modeling censored survival data from a German breast cancer research was investigated by Madaki and Bakar (2016). It has been found out recently that the extension of the beta Weibull distribution was proposed by Cordeiro et al. (2013). A generalized modified Weibull distribution by Wahed et al. (2009), is also a generalization of the Weibull distribution by Wallodi Weibull (1951), due to its flexibility in accommodating different forms of the risk function seems to be an important family that can be used in a variety of problems in modeling survival data. A Bayesian formulation of the cure fraction model is given by several authors like: Achcar and Santos (2013) and Martinez, Achcar, Jacome and Santos (2013). We have the cure rate model which is the mixture or standard parametric cure model for the survival function of the cure and uncured patients contribute comprehensively by Martinez et al. (2013).

The cure rate model was proposed by Achcar and Coelho (2012), with some other researchers like: Maller and Zhou (1996), Wahed et al. (2009), Madaki and Bakar (2016). contributed their quotas in the literature of statistical modelling. The second one is the non-mixture model which was proposed by Berkson and Gage (1952), also Boag (1949) as an asymptote alternative to the long-term survival model for maintaining the proportional hazard form and also making it easy for interpretations of the effect of covariates on the cure probability by Achcar and Coelho (2012). These two models received a quite lot of attention in the recent years and gas been motivated by many health cancer researchers (medical mechanism).

4. Methods

4.1. Model Specification

Let T be a random variable representing the time until the occurrence of the event of interest, and let t,0 be an observation from T. We refer to the model in equation (1), the survival function in the standard (" mixture") cure model proposed by Berkson and Gage (1952), also Boag (1949, the probability that the time-to-event is larger than that some specified t and p the proportion of cure fraction (" long-term survivors"). Therefore, the cumulative distribution function of T is given as:

$$F(t) = P(T \geq t) = 1 - S(t) = (1-p)[1 - S_o(t)] = (1-p)F_o(t). \tag{5}$$

Therefore, the $lim_{t\to\infty} F_o(t) = 1$ implies that $lim_{t\to\infty} F_o(t) = 1 - p$. The prob- ability density function for T is:

$$f(t) = \frac{dF(t)}{dt} = (1-p) f_o(t), \tag{6}$$

where $f_o(t)$ is the baseline probability density function for the susceptible individuals. Let us assume right censored data and non-informative censoring. Considering a random sample (t_i, δ_i) of size n, i=1,….,n, the contribution of the i^{th} subject for the likelihood function given by:

$$L_i = [f(t_i)]^{\delta i}[S(t_i)]^{1-\delta i} = [(1-p)f_o(t_i)]^{\delta i}[p + (1-p)S_o(t_i)]^{1-\delta i}, \quad (7)$$

where δ_i is the censoring indicator variable, denoting $\delta_i = 1$ for an observed lifetime and $\delta_i = 0$ for a censored lifetime.

Alternatively, the non-mixture model formulation suggested by Berkson and Gage (1952), is defined as an asymptote for the cumulative hazard, and hence for the cure fraction. The survival function can be referred to equation (2), for the non-mixture model respectively.

where, $F_o(t) = 1-S_o(t)$, we should take note that $lim_{t\to\infty}F_o(t) = 1$ simply implies that $lim_{t\to\infty}S(t) = p$. Assuming this model, the contribution of the i^{th} subject for the likelihood function is given as:

$$L_i = [h(t_i)]^{\delta i}S(t_i) = [-(lnp)f_o(t_i)]^{\delta i}\exp[ln(p)F_o(t_i)], \quad (8)$$

where the hazard function $h(t) = \frac{f(t)}{S(t)}$ can be interpreted as the risk of an event immediate after time t conditional on surviving up until time t.

$$h(t) = \frac{(1-p)f_o(t)}{p + (1-p)S_o(t)} \quad (9)$$

On these, we denote Go(t) as the cumulative distribution function (cdf) of a random variable T, which has a generalized class of distribution defined by:

$$F_o(t) = I_{G_o(t)}(m,n) = \frac{B_{G_o(t)}(m,n)}{B(m,n)} \quad (10)$$

where, $B(m,n) = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m,n)}$, as $m>0, n>0$, is the beta function, while as $\Gamma(m) = \int_0^\infty z^{(m-1)}e^z dz$, $\Gamma(m)$ is a gamma function and BGo(t)(m,n) is the incomplete beta function. If Go(t) in (10) is taken to be a cdf of a normal distribution with mean μ and variance σ^2, we have beta-normal distribution by Eugene et al. (2002). A model based on the CDF of the Kumaraswamy Burr X distribution Madaki et al (2018) with shape parameters β,γ,ϑ and scale parameter λ assumes:

$$G_o(t) = 1-\left(1 - e-(\lambda t)^2\right)^{\vartheta\alpha i}, \; t > 0 \quad (11)$$

where, from (10), we have the CDF of Beta Kumaraswamy Burr X distribution with five shape parameters m, n, β,γ,ϑ and one scale parameter λ assumes:

$$F_o(t) = 1-\left\{\left(1-e-(\lambda t)^2\right)^{\vartheta\alpha i}\right\}^{(m,n)}. \quad (12)$$

In the context of survival analysis, the baseline survival function or standard parametric survival curve function for the susceptible individuals is given by:

$$S_o(t) = 1-F_o(t). \quad (13)$$

4.2. Beta Kumaraswamy Burr X Model Formulation

We observed that the function cannot be expressed in a closed form reference to the limitation propose by Cordeiro et al (2011). The baseline probability density function of the Beta Kumaraswamy Burr X distribution with six parameters by Madaki et al (2018), is written as follows:

$$f_0(t) = \frac{2}{B(m,n)}\alpha\beta\vartheta\lambda^2 te^{-(\lambda t)^2} D_i[R_i]^{\beta n-1}[1-[R_i]^{\beta}]^{m-1}, t>0 \tag{14}$$

We denote $Di = \left[1- e-(\lambda t)^2\right]^{(\alpha-1)}$ and $Ri = 1-\left[\left(1-e-(\lambda t)^2\right)^{\vartheta\alpha}\right]$ respectively.

where, m,n,α,β,ϑ,λ > 0. The corresponding hazard function is given by:

$$h_o(t) = \frac{2}{S_o(t)B(m,n)}\alpha\beta\vartheta\lambda^2 te^{-(\lambda t)^2} D_i[R_i]^{\beta n-1}[1-[R_i]^{\beta}]^{m-1} \tag{15}$$

.

4.3. Beta Kumaraswamy Burr X Model Without Cure Fraction

The likelihood function for Beta Kum-BX Model without cure fraction: θ = (m, n, α,β,ϑ,λ) is given by:

$$L_{BetaKum-BX}(\theta) = \prod_{i=1}^{r}\left[\frac{2(1-p)\alpha\beta\vartheta\lambda^2}{B(m,n)}te^{-(\lambda t)^2} D_i[R_i]^{\beta n-1}[1-[R_i]^{\beta}]^{(m-1)}\right]. \tag{16}$$

Their corresponding log-likelihood of Beta Kumaraswamy Burr X model distribution without cure fraction: is:

$$\begin{aligned} l_{BetaKum-BX}(\theta) =& r\ln\left[\frac{2(1-p)\alpha\beta\vartheta\lambda^2}{B(m,n)}\right] + r\ln(t) - r(\lambda t)^2 + \sum_{i=1}^{r} D_i + (\beta n-1) \\ &\times \sum_{i=1}^{r}[R_i] + (m-1)\sum_{i=1}^{r}[1-[R_i]^{\beta}]. \end{aligned} \tag{17}$$

4.4. A Bayesian Analysis for the Beta Kumaraswamy Burr X Model

We consider the Beta Kumaraswamy Burr X model with density function in equation (7) above and a non-informative joint prior distribution for the θ = (m, n, α,β,ϑ,λ) given by:

$$\pi_o(m,n,\alpha,\beta,\vartheta,\lambda) \propto \frac{1}{m,n,\alpha,\beta,\vartheta,\lambda} \tag{18}$$

where, m,n,α,β,ϑ,λ > 0. The joint posterior distribution for these parameters can be written as:

$$\begin{aligned} =&\pi(m,n,\alpha,\beta,\vartheta,\lambda|t) \propto \pi_o(m,n,\alpha,\beta,\vartheta,\lambda)\exp\left[r\ln(\gamma) + r\gamma\ln(\gamma) - r\ln(B(m,n))\right. \\ &\left. + \sum_{i=1}^{r}\ln t_i^r e^{-r(\lambda t)^2} D_i + (\beta n-1)\sum_{i=1}^{r}\ln[R_i] + (m-1)\sum_{i=1}^{r}\ln[1-[R_i]^{\beta}\right]. \end{aligned} \tag{19}$$

$$Where \quad B(m,n) = \ln\Gamma(m+n) - \ln\Gamma(m) - \ln\Gamma(n)$$

$Di = \left[1 -e-(\lambda ti)^2\right]^{(\alpha-1)}$ and $Ri = 1-\left[\left(1- e-(\lambda ti)^2\right)^{\vartheta\alpha}\right]$ respectively.

4.5. Beta Kumaraswamy Burr X model with Cure Fraction for the Mixture Model

Assuming the mixture model, the likelihood function for θ = (m,n,α,β,ϑ,λ,p) is given by:

$$L_{BetaKum-BX_{mixture}}(\theta) = \prod_{i=1}^{r}\left[\frac{2(1-p)\alpha\beta\vartheta\lambda^2}{B(m,n)}te^{-(\lambda t)^2}D_i[R_i]^{\beta n-1}[1-[R_i]^{\beta}]^{m-1}\right]^{\delta_i} \times \prod_{i=1}^{r}[p+(1-p)S_o(t_i)]^{1-\delta_i}. \quad (20)$$

Their corresponding log-likelihood of mixture model is:

$$Di = \left[1- e-\left(\lambda t\right)^2\right]^{(\alpha-1)}$$

$$Ri = 1-\left[\left(1-e-\left(\lambda t\right)^2\right)^{\vartheta\alpha}\right]$$

$$l_{BetaKum-BX_{mixture}}(\theta) = \ln\left[\frac{2(1-p)\alpha\beta\vartheta\lambda^2}{B(m,n)}\right]\sum_{i=1}^{r}\delta_i + \sum_{i=1}^{r}\ln(t_i) - \sum_{i=1}^{r}\delta_i(\lambda t_i)^2 + (\beta n-1)\sum_{i=1}^{r}\delta_i\ln[R_i] + (m-1)\sum_{i=1}^{n}\delta_i\ln[1-[R_i]^{\beta}] + \sum_{i=1}^{r}\delta_i\ln(D_i) + \sum_{i=1}^{r}(1-\delta_i)\ln[p+(1-p)S_o(t_i)] \quad (21)$$

.

4.6. Beta Kumaraswamy Burr X Model with Cure Fraction for the Non-mixture Model

Moreover, assuming the non-mixture model, the likelihood function for θ = (m,n,α,β,ϑ,λ,p) is given by:

$$L_{BetaKum-BX_{non-mixture}}(\theta) = \prod_{i=1}^{r}\left[-\frac{2(\ln(p))\alpha\beta\vartheta\lambda^2}{B(m,n)}t_ie^{-(\lambda t)^2}D_i[R_i]^{\beta n-1}[1-[R_i]^{\beta}]^{m-1}\right]^{\delta_i} \times \exp\left[\frac{\ln(p)}{B(m,n)}\sum_{i=1}^{r}\int_0^{[1-[e^{-(\lambda t_i)^2}]^{(\alpha-1)}]} w^{m-1}(1-w)^{n-1}dw\right]. \quad (22)$$

Their corresponding log-likelihood of non-mixture model is:

$$l_{BetaKum-BX_{non-mixture}}(\theta) = \ln\left[-\frac{2(\ln(p))\alpha\beta\vartheta\lambda^2}{B(m,n)}\right]\sum_{i=1}^{r}\delta_i + \sum_{i=1}^{r}\delta_i\ln(t_i) - \sum_{i=1}^{r}\delta_i(\lambda t)^2 + (\beta n-1) \times \sum_{i=1}^{r}\delta_i\ln(R_i) + (m-1)\sum_{i=1}^{r}\delta_i\ln[1-[R_i]^{\beta}] + \sum_{i=1}^{r}\delta_i\ln(D_i) + \frac{\ln(p)}{B(m,n)}\sum_{i=1}^{r}\int_0^{[1-[e^{-(\lambda t)^2}]^{(\alpha-1)}]} w^{m-1}(1-w)^{n-1}dw. \quad (23)$$

4.7. Beta-Weibull Model Formulation

Similarly, the survival and hazard functions cannot be expressed in a closed form. The baseline probability density function of the beta-Weibull model with four parameters is written as follows:

$$f_0(t) = \frac{\gamma\lambda^{\gamma}t^{\gamma-1}}{B(\alpha,\beta)}[1-\exp(-(\lambda t)^{\gamma})]^{\alpha-1}[\exp(-(\lambda t)^{\gamma})]^{\beta}\ , t>0 \quad (24)$$

.

Where, α,β,γ and λ > 0. The corresponding hazard function is given by:

$$h_o(t) = \frac{f_o(t)}{s_o(t)} = \frac{\gamma\lambda^{\gamma}t^{\gamma-1}\{1-\exp(-(\lambda t)^{\gamma})]\}^{\alpha-1}\exp-(\lambda t)^{\gamma}}{B(\alpha,\beta)\int_0^{1-\exp[-(\lambda t)^{\gamma}]} w^{\alpha-1}(1-w)^{\beta-1}dw}. \quad (25)$$

4.8. Beta-Weibull Model Without Cure Fraction

The likelihood function for beta-Weibull model without cure fraction: θ= (α,β,γ,λ) is given by:

$$L_{BW}(\theta) = \prod_{i=1}^{r}\left[\frac{\gamma\,\lambda^{\gamma}\,t_i^{\gamma-1}}{B(\alpha,\beta)}[1-\exp(-(\lambda t_i)^{\gamma})]^{\alpha-1}[\exp(-(\lambda t_i)^{\gamma})]^{\beta}\right]. \quad (26)$$

Their corresponding log-likelihood of beta-Weibull distribution without cure fraction is:

$$l_{BW}(\theta) = r\ln\left[\frac{\gamma\lambda^{\gamma}}{B(\alpha,\beta)}\right] + (\gamma-1)\sum_{i=1}^{r}\ln t_i + (\alpha-1) \times \sum_{i=1}^{r}\ln[1-\exp(-(\lambda t_i)^{\gamma})] + \beta\sum_{i=1}^{r}\ln[\exp(-(\lambda t_i)^{\gamma})]. \quad (27)$$

4.9. A Bayesian Analysis for the Beta-Weibull Model

We consider the BW model with density function in equation (17) above and a non-informative joint prior distribution for the θ = (α,β,γ,λ) given by:

$$\pi_o(\alpha,\beta,\gamma,\lambda) \propto \frac{1}{\alpha,\beta,\gamma,\lambda} \quad (28)$$

where, α,β,γ,λ>0. The joint posterior distribution for these parameters can be written as:

$$= \pi(\alpha,\beta,\gamma,\lambda|t) \propto \pi_o(\alpha,\beta,\gamma,\lambda)\exp\left[r\ln(\gamma) + r\gamma\ln(\lambda) - r\ln(B(\alpha,\beta)) + (\gamma-1)\sum_{i=1}^{r}\ln t_i + (\alpha-1)\sum_{i=1}^{r}\ln(1-\exp(-\lambda t_i)^{\gamma}) - \beta\lambda^{\gamma}\sum_{i=1}^{r}t_i^{\gamma}\right]. \quad (29)$$

$$Where \quad B(\alpha,\beta) = \ln\Gamma(\alpha+\beta) - \ln\Gamma(\alpha) - \ln\Gamma(\beta).$$

4.10. Beta-Weibull model with Cure Fraction for the Mixture Model

Assuming the mixture model, the likelihood function for θ = (α,β,γ,λ,p) is given by:

$$L_{BW_{mixture}}(\theta) = \prod_{i=1}^{r}\left[\frac{\gamma(1-p)\lambda^{\gamma}t_i^{\gamma-1}}{B(\alpha,\beta)}[1-\exp(-(\lambda t)^{\gamma})]^{\alpha-1}[\exp(-(\lambda t)^{\gamma})]^{\beta}\right]^{\delta_i} \times \Big]\prod_{i=1}^{r}[p+(1-p)S_o(t_i)]^{1-\delta_i}. \quad (30)$$

Their corresponding log-likelihood of mixture model is:

$$l_{BW_{mixture}}(\theta) = \ln\left[\frac{\gamma(1-p)\lambda^{\gamma}}{B(\alpha,\beta)}\right]\sum_{i=1}^{r}\delta_i + (\gamma-1)\sum_{i=1}^{r}\delta_i \ln(t_i) + (\alpha-1)\sum_{i=1}^{r}\delta_i \ln[1-\exp(-(\lambda t)^{\gamma})] + \beta\sum_{i=1}^{r}\delta_i \ln[\exp(-(\lambda t))^{\gamma})] + \sum_{i=1}^{r}(1-\delta_i)\ln[p+(1-p)S_o(t_i)]. \tag{31}$$

4.11. Beta-Weibull Model with Cure Fraction for the Non-mixture Model

The likelihood function for non-mixture model θ = (α,β,γ,λ,p) is given by:

$$L_{BW_{non-mixture}}(\theta) = \prod_{i=1}^{r}\left[-\frac{\gamma(\ln(p))\lambda^{\gamma}t_i^{\gamma-1}}{B(\alpha,\beta)}[1-\exp(-(\lambda t)^{\gamma})]^{\alpha-1}[\exp(-(\lambda t)^{\gamma})]^{\beta}\right]^{\delta_i} \times \exp\left[\frac{\ln(p)}{B(\alpha,\beta)}\sum_{i=1}^{r}\int_{0}^{1-\exp(-(\lambda t)^{\gamma})} w^{\alpha-1}(1-w)^{\beta-1}dw\right]. \tag{32}$$

and the corresponding log-likelihood of non-mixture model is:

$$l_{BW_{non-mixture}}(\theta) = \ln\left[\frac{-\gamma(\ln(p))\lambda^{\gamma}}{B(\alpha,\beta)}\right]\sum_{i=1}^{r}\delta_i + (\gamma-1)\sum_{i=1}^{r}\delta_i \ln(t_i) + (\alpha-1)\sum_{i=1}^{r}\delta_i \ln[1-\exp(-(\lambda t)^{\gamma})] + \beta\sum_{i=1}^{r}\delta_i \ln[\exp(-(\lambda t)^{\gamma})] + \frac{\ln(p)}{B(\alpha,\beta)}\sum_{i=1}^{r}\int_{0}^{1-\exp(-(\lambda t)^{\gamma})} w^{\alpha-1}(1-w)^{\beta-1}dw. \tag{33}$$

4.12. Beta Generalized Exponential Model Formulation

The baseline probability density function of the beta generalized exponential (BGE) model with four parameters by Barreto, Santos and Cordeiro (2010), is written as follows:

$$f_0(t) = \left(\frac{\alpha\lambda}{B(m,n)}\right)\exp(-(\lambda t))[1-\exp(-(\lambda t))]^{(m\alpha-1)}[1-(1-\exp(-(\lambda t))^{\alpha})^{(n-1)}]\ , t>0 \tag{34}$$

where m,n, α,λ > 0. The corresponding hazard function is given by:

$$h_o(t) = \frac{f_o(t)}{s_o(t)} = \frac{\alpha\lambda\exp(-(\lambda t))[1-\exp(-(\lambda t))]^{(m\alpha-1)}[1-(1-\exp(-(\lambda t))^{\alpha})]^{(n-1)}}{B(m,n)\int_{0}^{1-\exp[-(\lambda t)^{\alpha}]} w^{m-1}(1-w)^{n-1}dw} \tag{35}$$

4.13. Beta Generalized Exponential Model without Cure Fraction

The likelihood function for beta generalized exponential model without cure fraction:

θ = (m, n, α,λ) is given by:

$$L_{BGE}(\theta) = \prod_{i=1}^{r}\left\{\left(\frac{\alpha\lambda}{B(m,n)}\right)\left[\exp(-(\lambda t_i))[1-\exp(-(\lambda t_i))]^{(m\alpha-1)}[1-(1-\exp(-(\lambda t_i))^{\alpha})]^{(n-1)}\right]\right\}. \tag{36}$$

Their corresponding log-likelihood of beta generalized exponential distribution without cure fraction is:

$$l_{BGE}(\theta) = r\ln\left(\frac{\alpha\lambda}{B(m,n)}\right) - r\lambda\sum_{i=1}^{r} t_i + (m\alpha-1)\sum_{i=1}^{r}\ln[1-\exp(-(\lambda t_i))] + (n-1)\sum_{i=1}^{r}\ln[1-(1-\exp(-(\lambda t_i))^{\alpha})]. \quad (37)$$

4.14. A Bayesian Analysis for the Beta Generalized Exponential Model

We consider the BGE model with density function in equation (27) above and a non-informative joint prior distribution for the θ = (m,n,α,λ) given by:

$$\pi_o(m,n,\alpha,\lambda) \propto \frac{1}{m,n,\alpha,\lambda} \quad (38)$$

where, m,n,α,λ>0. The joint posterior distribution for these parameters can be written as

$$=\pi(m,\ n,\ \alpha,\lambda|t) \propto \pi_o(m,\ n,\ \alpha,\lambda)\exp\left[r\ln(\gamma) + r\gamma\ln(\gamma) - r\ln(B(m,n)) + \lambda\sum_{i=1}^{r} t_i + (m\alpha-1)\sum_{i=1}^{r}\ln[1-\exp(-(\lambda t_i))] + (n-1)\sum_{i=1}^{r}\ln[1-(1-\exp(-(\lambda t_i))^{\alpha})\right]. \quad (39)$$

4.15. Beta Generalized Exponential Model with Cure Fraction for the Mixture Model

Assuming the mixture model, the likelihood function for θ = (m, n, α,λ,p) is given by:

$$L_{BGE_{mixture}}(\theta) = \prod_{i=1}^{r}\left[\frac{\alpha(1-p)\lambda}{B(m,n)}\exp(-(\lambda t_i)[1-\exp(-(\lambda t_i))]^{(m\alpha-1)} \times [1-(1-\exp(-(\lambda t_i))^{\alpha})]^{(n-1)}\right]^{\delta_i} \times \prod_{i=1}^{r}\left[p+(1-p)S_o(t_i)\right]^{(1-\delta_i)} \quad (40)$$

.

Their corresponding log-likelihood of mixture model is:

$$l_{BGE_{mixture}}(\theta) = \ln\left(\frac{\alpha(1-p)\lambda}{B(m,n)}\right)\sum_{i=1}^{r}\delta_i - \lambda\sum_{i=1}^{r}\delta_i t_i + (m\alpha-1)\sum_{i=1}^{r}\delta_i\ln[1-\exp(-(\lambda t_i))] + (n-1)\sum_{i=1}^{r}\delta_i\ln[1-(1-\exp(-(\lambda t_i))^{\alpha})] + \sum_{i=1}^{r}(1-\delta_i)\ln[p+(1-p)S_o(t_i)] \quad (41)$$

4.16. Beta Generalized Exponential Model with Cure Fraction for the Non-mixture Model

The likelihood function for the non-mixture model θ = (m, n, α,λ,p) is given by:

$$L_{BGE_{non-mixture}}(\theta) = \prod_{i=1}^{r}\left[\left(-\frac{\alpha(\ln(p))\lambda}{B(m,n)}\right)\exp(-(\lambda t)[1-\exp(-(\lambda t))]^{(m\alpha-1)} \times [1-(1-\exp(-(\lambda t))^{\alpha})]^{(n-1)}\right]^{\delta_i}\exp\left[\frac{\ln(p)}{B(m,n)}\sum_{i=1}^{r}\int_{0}^{1-\exp[-(\lambda t)^{\alpha}]} w^{m-1}(1-w)^{n-1}dw\right] \quad (42)$$

and the corresponding log-likelihood of non-mixture model is:

$$l_{BGE_{non-mixture}}(\theta) = \ln\left[-\frac{\alpha(\ln(p))\lambda}{B(m,n)}\right]\sum_{i=1}^{r}\delta_i r - (\lambda t)\sum_{i=1}^{r}\delta_i \ln t_i + (m\alpha-1)\sum_{i=1}^{r}\delta_i \ln[1-\exp(-(\lambda t))] + \left[\frac{\ln(p)}{B(m,n)}\sum_{i=1}^{r}\int_{0}^{1-\exp[-(\lambda t)^{\alpha}]} w^{m-1}(1-w)^{n-1}dw\right] \quad (43)$$

4.17. Incorporating Covariates

Due to the intricacy and complexity case of all the joint log-likelihood functions,

a. $l_{Beta\ Kum-BX\ mixture}(\theta)$,
b. $l_{Beta\ Kum-BX\ non-mixture}(\theta)$,
c. $l_{BW\ \ mixture}(\theta)$,
d. $l_{BW\ \ non-mixture}(\theta)$,
e. $l_{BGE\ \ mixture}(\theta)$,
f. $l_{BGE\ \ non-mixture}(\theta)$.

the estimation of the parameter by maximization or direct method will be extremely a difficult task. In order to overcome in dealing with this type of problem, we consider the use of Bayesian inference based on (MCMC) methods. We can probably also take in the incorporate a vector of covariates yi that may be closely related and associated with the proportion p of cure rate fraction models by replacing p in the log-likelihood functions above by:

$$p_i(y) = \frac{exp(y_i'\eta)}{1+exp(y_i'\eta)} \quad (44)$$

5. Bayesian Analysis

For a Bayesian analysis of the long-term survival models without considering covariates Achcar et al 2013, on the other hand we also presume the beta prior for the given probability of proportion p of cure models which is denoted by p∼Beta(m,n) where m and n are known hyper parameters. We also assume a gamma prior distribution for the parameters θ =(m,n,α,β,θ,λ) by Achcar et al 2013. That is, m ∼ Gamma(cmdm), n ∼ Gamma(cndn), α ∼ Gamma(cαdα), β∼Gamma(cβdβ), θ∼Gamma(cϑdϑ), λ∼Gamma(cλdλ), where cm,dm,cn,dn,cα,dα,cβ,dβ,cϑ,dϑ ,cλ and dλ are known hyper parameters and Gamma(c,d) denotes a gamma distribution with mean $\frac{c}{d}$ $and\ variance$ $\frac{c}{d^2}$. In all cases the joint prior distribution is then establish by assuming prior independence between the parameters, or say,

$$\pi(\theta) = \pi(m), \pi(n), \pi(\alpha), \pi(\beta), \pi(\vartheta), \pi(\lambda) \propto m^{cm-1} n^{cn-1} \alpha^{c\alpha-1} \beta^{c\beta-1} \vartheta^{c\vartheta-1} \lambda^{c\lambda-1}$$
$$\times exp\Big(- \frac{m}{dm} - \frac{n}{dn} - \frac{\alpha}{d\alpha} - \frac{\beta}{d_\beta} - \frac{\vartheta}{d_\vartheta} - \frac{\lambda}{d_\lambda}\Big) p^{c-1}(1-p)^{d-1} \quad (45)$$

for models incorporating the following covariates, the prior distribution for the unknown parameters is assumed: m ~ Gamma(cmdm), n ~ Gamma(cndn),
α ~ Gamma(cαdα), β ~ Gamma(cβdβ), θ ~ Gamma(cϑdϑ), λ ~ Gamma(cλdλ),ζj ~ N(cζj,d2ζj), j= 0,1,...,J and ηk ~ N(cηk,dη2k), k = 0,1,...,K.

where N (c, d^2) denotes a normal distribution with mean c and variance d^2. are hyper parameters. In this situation we should focus on the independence between the prior distributions denotes a normal distribution with mean c and variance d^2 for the hyper parameters.

In all process we generated 3,500 Gibbs samples taking every 10th sample after a "burn-in-sample" of size 2,000 to eliminate the initial values used in the Gibbs sampling iterations. All the simulations were done using the Winbugs and R softwares to obtain the posterior summaries. The posterior convergence of the sampling algorithm was confirmed by trace plots in Figure 3, of the simulated results.

5.1. The Gibbs Sampler

We wish to obtain a sample from the multivariate distribution (1,...,d). We shall call this distribution the target distribution. In Bayesian statistics, the target distribution is the joint posterior distribution. The Gibbs sampler obtains a sample from (1,...,d) this simulation step is usually straightforward [34]. The algorithm is as follows:

- Initialize with. $\theta = (\theta_1^{(0)}, ..., \theta_d^{(0)})$
- Simulate $\theta_1^{(1)}$ from the conditional distribution. $\pi(\theta_1|\theta_2^{(0)}, \theta_3^{(0)}, ..., \theta_d^{(0)})$
- Simulate $\theta_2^{(1)}$ from the conditional distribution $\pi(\theta_2|\theta_1^{(1)}, \theta_3^{(0)}, ..., \theta_d^{(0)})$
- Continue sampling....
- Simulate$\theta_d^{(1)}$ from the conditional distribution$\pi(\theta_d|\theta_1^{(1)}, \theta_2^{(1)}, ..., \theta_{d-1}^{(1)})$.
- Iterate this procedure.

6. Model Selection

Several criterions have been employed to select the best model based on the type used to compare with the existing models. Here we consider the log Pseudo marginal likelihood measure (LPML) where it considers the large value as the best model selection.

6.1. Log Pseudo Maximum Likelihood

Comparison of the two cure models assuming different distributions will be accessed using log Pseudo marginal likelihood measure and the Pseudo factor. The LPML is derived from the conditional predictive ordinate (CPO). For the i^{th} observation, the CPO_i is given by:

$$f(D_i/y_{[i]}) = \int f(D_i/\Theta f\Big(\Theta|D_{[i]}\Big) d\Theta \quad (46)$$

where Θ is the incomplete vector of parameters, Di is each instance of the full data
D, D[i] is D without the current observation i and $f\Big(\Theta/D_{[i]}\Big)$ is the posterior density of Θ given D[i], i = 1,...,r. An MCMC approximation of CPOi is given by:

$$\widehat{CPO_i} = \Big[\frac{1}{B}\sum_{b=1}^{B} \frac{1}{f(D_i/\Theta_{(b)}}\Big]^{-1}, i = 1, ..., n. \quad (47)$$

where B is the number of iterations during the implementation of the MCMC procedure after the burn-in-period and Θb is vector of the samples that will be obtained at the 4th to 5th iterations Gelfand et al. (1992). For a given model, the LPML value is given by:

r

$$\text{LPMLd} = \text{XlogCPOd i.} \quad (48)$$
$$\text{i=1}$$

The larger the value of LPML, the better is fit of the model by Gelfand et al. (1992). Alternatively, the Pseudo Bayes factor (PMF) for comparing models b and b' is:

$$PBF_{bb'} = exp\left(\widehat{LPML_b} - \widehat{LPMLb'}\right. . \quad (49)$$

6.2. Highest Probability Density (HPD) Interval

We also obtained the highest probability density (HPD) intervals for parameters of interest Geweke et al (1991). A 100(1−ω)% HPD interval for a generic parameter θ is a subset of the parameter space CΘ given by C={θ:π(θ|D)≥k} where π(θ|D) is the posterior distribution for θ given the data D and k is the largest number such that Rπ(θ|D)≥k π(θ|D) =1−ω .

7. Application Censored Data

7.1. AIDS Clinical Trials Group Study Data

"AIDS Clinical Trials Group Study Data by Hosmer et al. (2008). This data contains the placebo-controlled trial that compared the three-drug regimen of indinavir (IDV), open label zidovudine (ZDV) or stavudine (d4T) and lamivudine (3TC) with the two-drug regimen of zidovudine or stavudine and lamivudine in HIV-infected patients". We consider the time-to-event of each patient denoted by "ti" and also the censoring indicator "δi", where "δi = 1" meaning death was observed or "δi=0" censored count of the patient probably lost to follow-up or missing .

7.2 Results

Table 1 - The posterior summaries for The Beta Kum-BX, BW and BGE models based on AIDS clinical trials group study data

Model	Parameter	MLE	Bayes	90%HPDa	95%HPDa	LPMLb	HWcp value	Geweke's p value
Beta Kum-BX	m	0.8008	0.8108	(0.5123,1.9603)	(1.0601,1.2922)		0.618	0.307
	n	0.9308	0.9505	(0.0871,2.8578)	(1.1189,2.3834)		0.099	0.223
	α	1.0910	1.0110	(0.4335,1.1979)	(1.3125,2.3629)	148	0.314	0.112
	β	1.0420	1.0520	(0.0117,1.0890)	(0.0186,0.2481)		0.392	0.251
	ϑ	0.4664	0.4564	(0.2922,0.4662)	(0.6625,1.2972)		0.348	0.277
	λ	0.7084	0.7184	(0.0855,1.0423)	(1.1376,1.3364)		0.099	0.223
BW	α	1.8016	1.8116	(0.4878,1.6794)	(1.5214,2.1259)		0.314	0.112
	β	0.6706	0.6806	(0.0089,0.0197)	(0.0156,0.2218)	145	0.372	0.251
	γ	0.9184	0.9084	(0.0132,1.0781)	(0.5411,1.2922)		0.614	0.307
	λ	2.3076	2.3126	(0.2747,0.4612)	(1.1845,1.3524)	0.012	0.223	
BGE	m	1.3529	1.3499	(0.2934,0.4921)	(1.0664,1.2017)		0.463	0.334
	n	0.8640	0.8590	(0.0661,0.2384)	(0.0250,0.1167)	141	0.214	0.772
	α	0.6606	0.6806	(0.0212,0.1024)	(0.0266,0.3657)		0.392	0.251
	λ	2.2778	2.2978	(0.2864,0.4601)	(1.0235,1.1308)		0.399	0.394

Table 2a - The posterior summaries for The Beta Kum-BX, BW and BGE models with mixture models based on AIDS clinical trials group study data

Model	Parameter	MLE	Bayes	95%*HPD*[a]	*LPML*[b]	*HW*[c] *p* value	Geweke's *p* value
	m	0.9022	0.9122	(0.5281,1.2922)		0.618	0.327
Beta	n	1.4225	1.4165	(0.1376,1.3364)		0.099	0.263
Kum-	α	1.8013	1.8123	(0.5334,1.7329)		0.314	0.112
BX	β	0.8236	0.8526	(0.0166,0.2681)	147	0.392	0.251
	ϑ	0.4424	0.4654	(0.6675,0.8072)		0.349	0.287
	λ	1.3213	1.3346	(1.1736,2.3743)		0.129	0.239
	p	0.8543	0.8786	(0.0134,0.2123)		0.472	0.351
	α	1.7235	1.7134	(1.5214,2.6545)		0.345	0.176
	β	0.9703	0.9806	(0.0145,0.2671)		0.374	0.265
BW	γ	0.9274	0.9134	(0.5370,1.2922)	143	0.623	0.367
	λ	2.4376	2.4526	(1.1086,1.1344)		0.659	0.243
	p	0.5176	0.5236	(0.0187,0.5234)		0.392	0.031
	m	1.3521	1.3499	(1.4324,2.2017)		0.463	0.334
	n	0.6480	0.6590	(0.0250,0.1167)		0.818	0.772
BGE	α	0.7775	0.7865	(0.2317,1.8654)	138	0.109	0.231
	λ	2.5083	2.4978	(1.2235,4.1378)		0.699	0.394
	p	0.8536	0.8811	(0.0236,0.2161)		0.392	0.435

Table 2b - The posterior summaries for The Beta Kum-BX, BW and BGE models with non-mixture models based on AIDS clinical trials group study data

Model	Parameter	MLE	Bayes	95%*HPD*[a]	*LPML*[b]	*HW*[c] *p* value	Geweke's *p* value
	m	0.9227	0.9012	(0.5511,1.2082)		0.618	0.347
Beta	n	1.3453	1.3546	(1.1076,2.3154)		0.069	0.232
Kum-	α	1.5356	1.5416	(1.5334,2.7675)		0.321	0.102
BX	β	0.9109	0.9123	(0.0376,0.2431)	147	0.392	0.251
	ϑ	0.9218	0.9344	(0.5431,1.2922)		0.618	0.307
	λ	0.5234	0.5426	(1.1566,2.1364)		0.057	0.167
	p	0.77846	0.7906	(0.0126,0.2761)		0.438	0.522
	α	1.8108	1.8116	(1.5334,1.7349)		0.314	0.112
	β	0.8217	0.8306	(0.0326,0.2321)		0.352	0.431
BW	γ	0.9217	0.9384	(0.2711,1.2982)	145	0.764	0.507
	λ	2.2178	2.3126	(1.1076,1.3364)		0.099	0.223
	p	0.6039	0.6082	(0.1536,0.2681)		0.392	0.241
	m	1.3532	1.3499	(1.9544,2.2117)		0.463	0.334
	n	0.9543	0.9590	(0.0220,0.1267)		0.818	0.772
BGE	α	0.5638	0.5434	(1.2235,2.1247)	138	0.187	0.134
	λ	2.3843	2.3978	(1.2435,1.4378)		0.699	0.394
	p	0.8386	0.8306	(0.0226,0.1351)		0.392	0.234

8.1. Remarks on Table 1 and Table 2a and Table 2b

From the Table 1, we observed that the convergence of the MCMC algorithm was partially obtained even when using a very large burn-in-period for the algorithm. These results were shown in Table 1, considering the Beta Kum BX, BW and BGE distributions respectively. Also, we noted that all p-values from (HW) Heidelberger and Welch convergence diagnostics criteria do not reject the null hypothesis of stationary of the chains, since they are all larger or equal than 0.05 level of significance. While the Geweke's *p*-values also suggest convergence, on the other hand these results shows that, among all the models considered BGE distribution has the least (LPML) Log pseudo marginal likelihood value, while Beta Kum-BX and BW distributions have similar LPML values, where strong evidence shows that these models are better fitted by the data than, From Table 2a and Table 2b, for each run, the convergence was assessed by monitoring the trace plots and using Heidelberger and Welch (HW) and Geweke's convergence diagnostics. The model based on the Beta Kum-BX and other existing models in the presence of cure fraction for the using the MCMC estimation as we demonstrate that the use of this Bayesian methodology is quite suitable to get the posterior inferences for the parameters of the model. We also showed that the model estimation in using cure fraction results in more precise inferences of the result based on the two models at all levels shows that the mixture models fit well better than the non-mixture models.

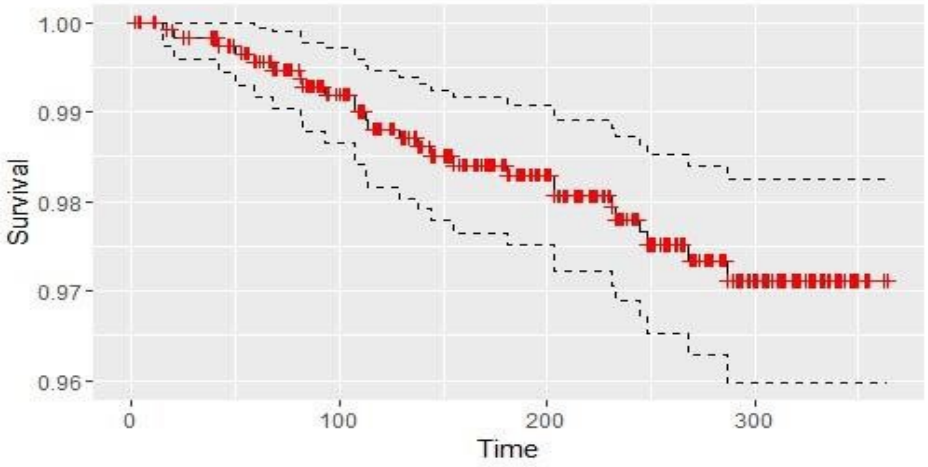

Fig. 1 - Survival functions estimated by Kaplan-Meier method

Fig. 1 shows the plots of the survival functions estimated by Kaplan-Meier method and from the models based on Time to AIDS diagnosis or death and the event indicator for AIDS defining diagnosis or death, based on AIDS clinical trials group study data.

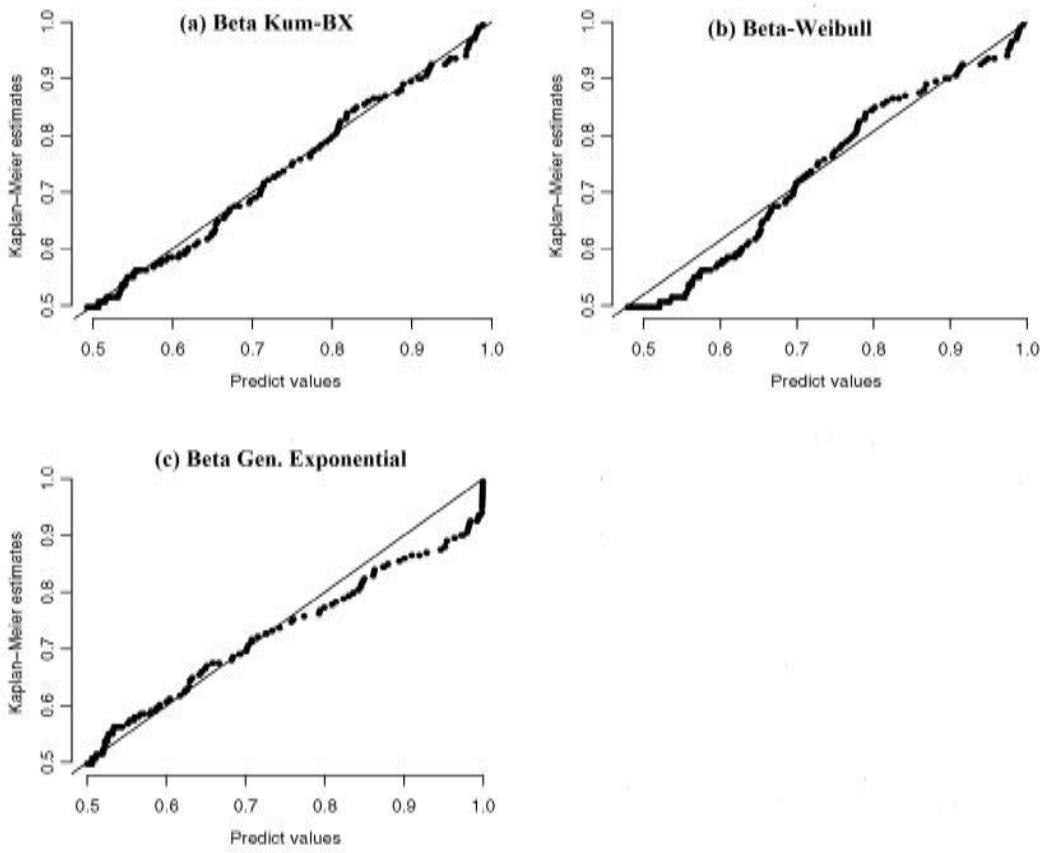

Fig. 2 - Kaplan-Meier estimates for the survival function

Fig. 2 shows the of the Kaplan-Meier estimates for the survival function versus the respective predict values obtained from the parametric mixture models for each probability distribution of interest: (a) Beta Kum-BX, (b) BW, (c) BGE distributions AIDS clinical trials group (ACTG) Study. The diagonal straight lines represent a perfect agreement between Kaplan-Meier estimates and predicted values as its clear shows how perfect our new model fits the AIDS clinical trials group study data.

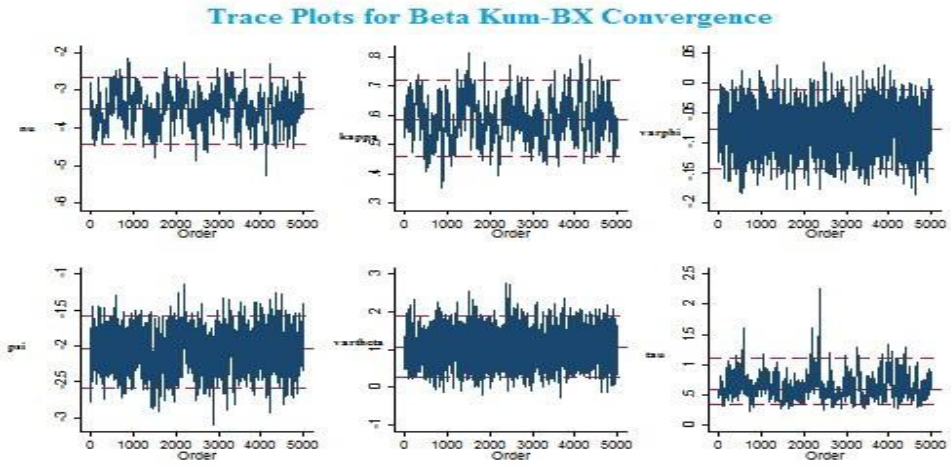

Fig. 3 - Plots of the convergence diagnostics and output analysis (coda)

Fig. 3 shows some of the trace plots of the new model Beta Kum-BX showing that the priori distribution is well calibrated which is indicated the parameters having sufficient state changes as the MCMC algorithm runs based on AIDS clinical trials group study data. Plots of the convergence diagnostics and output analysis (coda)

object were estimated with a bit close estimation but non-mixture models with covariates fits better than the mixture model having the highest (LPML) log pseudo marginal likelihood based on the weakness of the mixture cure model pointed out in the literature this actually validate the assumptions. Also, the 95% credible intervals for all the parameters and covariates. We can obtain Bayesian estimates for the cure fractions of each risk group considering the simulated samples for the covariates η0 = age at enrollment, η1= sex and η2= race. The Heidelberger and Welch (HW) convergence diagnostics calculates a test statistic based on stationary distribution and Geweke's convergence diagnostics, is based on the test for equality of the means of two nonoverlapping parts of a Markov chain process. It was considered in Bayesian survival analysis for the estimated results obtained for the cure fractions of the patients classified as Beta Kum-BX and BW fits better having highest probability intervals than the so-called BGE model with the least LPML values as shown above respectively.

Table 3a - The posterior summaries for The Beta Kum-BX, BW and BGE models with mixture models and covariates based on AIDS clinical trials group study data

Model	Parameter	MLE	Bayes	95%*HPD*[a]	*LPML*[b]	*HW*[c] p value	Geweke's p value
Beta Kum-BX	m	0.8946	0.9056	(0.5451,1.2922)	147	0.618	0.677
	n	1.3206	1.3426	(1.1376,2.3364)		0.099	0.223
	α	1.2676	1.2816	(1.5334,1.7089)		0.314	0.112
	β	0.9643	0.9706	(0.0166,0.2681)		0.392	0.251
	ϑ	0.8413	0.8674	(0.545,1.2922)		0.618	0.323
	λ	0.7712	0.7806	(0.0166,0.3961)		0.382	0.361
	η_0	0.8134	0.8337	(-1.1345,0.3965)		0.780	0.433
	η_1	0.5216	0.5433	(-0.3411,0.2922)		0.618	0.707
	η_2	0.6198	0.6040	(-1.7856,0.2374)		0.779	0.243
BW	A	1.4619	1.4729	(0.3904,2.3249)	145	0.457	0.659
	B	1.1512	1.1532	(0.1107,1.2681)		0.224	0.271
	γ	0.8211	0.8381	(0.5337,1.6854)		0.3648	0.727
	λ	2.1609	2.1816	(0.0266,0.2481)		0.392	0.251
	η_0	0.6337	0.6337	(-1.1376,0.3364)		0.799	0.643
	η_1	0.7431	0.7843	(-0.2201,0.2922)		0.038	0.508
	η_2	0.9021	0.9124	(-1.1334,0.3744)		0.624	0.243
BGE	m	1.0272	1.0649	(1.9664,2.2017)	140	0.483	0.147
	n	0.5213	0.5406	(0.0166,0.2761)		0.562	0.045
	α	0.5176	0.5236	(1.0972,2.7675)		0.216	0.654
	λ	2.2145	2.2324	(0.1296,0.2344)		0.238	0.231
	η_0	0.7218	0.7365	(-0.1547,1.9864)		0.079	0.245
	η_1	0.5010	0.5053	(-1.5731,1.2832)		0.068	0.357
	η_2	0.9098	0.9234	(-1.1376,0.0324)		0.099	0.223

Table 3b - The posterior summaries for The Beta Kum-BX, BW and BGE models with non-mixture models and covariates based on AIDS clinical trials group study data

Model	Parameter	MLE	Bayes	95%*HPD*[a]	*LPML*[b]	*HW*[c] p value	Geweke's p value
Beta Kum-BX	m	0.9014	0.9112	(0.5671,1.2752)	150	0.623	0.334
	n	1.3226	1.3215	(1.1376,1.3364)		0.099	0.223
	α	1.1087	1.1236	(1.5334,2.23189)		0.314	0.112
	β	0.6043	0.6176	(0.0446,0.2681)		0.392	0.251
	ϑ	0.9114	0.9084	(0.5611,1.2922)		0.618	0.307
	λ	0.7743	0.7806	(0.0166,0.2681)		0.392	0.251
	η_0	0.9421	0.9337	(-1.1376,0.3364)		0.799	0.343
	η_1	0.8653	0.8843	(-0.2611,0.2922)		0.618	0.707
	η_2	0.8125	0.8024	(-1.1976,0.3364)		0.779	0.243
BW	α	1.5012	1.5129	(0.3684,2.1799)	147	0.453	0.729
	β	0.5032	0.4932	(0.1507,1.2691)		0.244	0.261
	γ	1.0415	1.0121	(0.5337,1.6854)		0.748	0.617
	λ	2.0215	2.0706	(0.0166,0.2681)		0.352	0.261
	η_0	0.8216	0.8337	(-1.8976,0.3304)		0.799	0.343
	η_1	0.4732	0.4843	(-0.2611,0.2822)		0.618	0.707
	η_2	0.8321	0.8133	(0.1315,0.3764)		0.969	0.433
BGE	m	1.5479	1.5649	(1.8664,2.1217)	144	0.129	0.157
	n	0.3221	0.3133	(1.3015,1.3764)		0.969	0.465
	α	0.8242	0.8236	(0.2174,1.6134)		0.125	0.154
	λ	2.3632	2.3766	(0.0123,0.2651)		0.492	0.561
	η0	0.81765	0.8065	(-0.1376,1.9864)		0.099	0.223
	η1	0.4857	0.4753	(-1.4611,1.1922)		0.618	0.337
	η2	0.7219	0.7146	(-1.2376,-0.0231)		0.019	0.213

8.2. Remarks on Table 3a and Table 3b

From Table 3a and b, they are clearly shows that the inferences for the two models based on the of the Beta Kum-BX and other existing models. The two models were estimated with a bit close estimation but non-mixture models with covariates fits better than the mixture model having the highest (LPML) log pseudo marginal likelihood based on the weakness of the mixture cure model pointed out in the literature this actually validate the assumptions. Also, the 95% credible intervals for η_2= sex. We can obtain Bayesian estimates for the cure fractions of each risk group considering the simulated samples for the covariates η_0 = age at enrollment, η_1= sex and η_2= race. The Heidelberger and Welch (HW) convergence diagnostics calculates a test statistic based on stationary distribution and Geweke's convergence diagnostics, is based on the test for equality of the means of two non-overlapping parts of a Markov chain process. It was considered in Bayesian survival analysis for the estimated results obtained for the cure fractions of the patients classified as Beta Kum-BX and BW fits better having highest probability intervals than the so-called BGE model with the least LPML values as shown, respectively.

Table 4 - A simulation study comparison for the Beta Kum-BX and other existing models with long-term survival models

				n=30			n=50			n=100		
Model	**Censoring rate**	**Parameter**	**True value**	**Bias**	**MSE**	**CI**	**Bias**	**MSE**	**CI**	**Bias**	**MSE**	**CI**
				0.6734	0.4862	96.3	0.6421	0.4448		0.5453	0.2240	
				0.7550	0.6702	93.3	0.7252	0.6222		0.7386	0.5177	
mixture model		m	1.0	0.9316	0.1941	92.5	0.9137	0.1437	95.7	0.8334	0.0754	87.0
Beta	slight	n	1.4						90.6			85.8
Kum-	U[0,20]	α	1.8	0.1803	0.1224	92.5	0.1596	0.1137	91.6	0.1081	0.0879	84.2
BX	20%	β	0.8						90.7			86.0
		ϑ	0.5	0.2290	0.1807	93.8	0.2046	0.1656	91.8	0.1665	0.0222	84.6
		λ	1.3	0.4147	0.6743	96.5	0.3139	0.4251	94.3	0.2536	0.0432	83.4
		p	1.0	0.4644	0.1530	94.2	0.3470	0.1505	93.7	0.2135	0.1018	86.2
	moderate	m	1.0	0.6306	0.4646	92.7	0.6036	0.3194	91.8	0.5528	0.1918	87.4
	U[0,7]	n	1.4	0.6585	0.5830	96.8	0.6287	0.5377	94.2	0.5779	0.5172	89.0
	50%	α	1.8	0.9236	0.1872	95.4	0.8837	0.1874	94.4	0.8133	0.0695	90.0
		β	0.8	0.1562	0.1067	91.6	0.1281	0.1069	90.3	0.1373	0.1018	84.4
		ϑ	0.5	0.1977	0.0321	95.6	0.1840	0.1429	92.7	0.1531	0.0207	83.0
		λ	1.3	0.3742	0.5481	92.7	0.3106	0.4520	90.8	0.1632	0.0321	89.0
		p	1.0	0.3490	0.1392	95.0	0.2459	0.1493	92.8	0.2014	0.0901	85.1
				0.4613	0.3757	96.3	0.4052	0.2038		0.3417	0.2131	
				0.2856	0.1664	95.8	0.2502	0.1299		0.1740	0.1057	
	slight	α	1.7						95.8			83.2
	U[0,25]	β	1.0	0.1982	0.1328	94.4	0.1381	0.1028	92.3	0.0804	0.0655	83.4
BW	20%	γ	1.0						92.4			89.4
		λ	2.5	0.1655	0.1407	96.3	0.1318	0.0975	94.0	0.0855	0.0666	83.2
		p	0.5	0.1459	0.1385	95.8	0.1251	0.0975	93.4	0.0853	0.0430	88.2
	moderate	α	1.7	0.4238	0.1461	94.0	0.3839	0.1190	91.7	0.3349	0.5257	88.8
	U[0,8]	β	1.0	0.2903	0.1530	94.0	0.2429	0.1270	92.6	0.1414	0.0644	90.8
	50%	γ	1.0	0.3022	0.2905	94.0	0.2680	0.2429	92.8	0.1605	0.1764	89.4

		λ	2.5	0.1626	0.1284	95.7	0.1195	0.0879	93.8	0.0593	0.0312	86.4
		p	0.5	0.1338	0.1123	93.0	0.1023	0.0791	91.3	0.0556	0.0271	85.4
				0.1566	0.1384	95.4	0.1215	0.0773		0.0876	0.0475	
	slight		1.3	0.3092	0.1615	95.5	0.2635	0.1225	94.6	0.2081	0.0785	89.4
	U[0,23]	m	0.7	0.1714	0.2070	95.6	0.1216	0.1573	93.8	0.0818	0.0951	89.2
BGE	20%	n α	0.8						93.2			84.4
		λ	2.5	0.4513	0.3191	94.3	0.4125	0.2945	91.8	0.3235	0.2298	86.2
		p	0.9	0.2041	0.1667	94.7	0.1724	0.1360	93.6	0.0631	0.0531	86.3
	moderate	m	1.3	0.1265	0.1146	92.1	0.1063	0.0925	90.4	0.0527	0.0671	84.8
	U[0,6]	n	0.7	0.3209	0.1358	93.8	0.0067	0.0510	90.4	0.4970	0.4686	85.0
	50%	α	0.8	0.1430	0.1294	91.3	0.0993	0.0862	89.6	0.0514	0.0317	78.0
		λ	2.5	0.4662	0.2517	93.2	0.4246	0.2538	91.2	0.3070	0.2163	86.1
		p	0.9	0.2263	0.1807	91.6	0.1613	0.1283	90.6	0.0746	0.0593	83.2
				0.6528	0.3771	94.3	0.6106	0.4805		0.5417	0.2131	
non-mixture model				0.8148	0.9134	93.4	0.7837	0.6660		0.6740	0.3157	
		m	1.0	0.9149	0.3441	93.8	0.8822	0.2881	93.7	0.7404	0.1495	86.2
Beta	slight	n	1.4						90.8			85.4
Kum-	U[0,20]	α	1.8	0.2846	0.1489	94.2	0.1662	0.1059	90.5	0.1255	0.0866	86.4
BX	20%	β	0.8						91.8			86.2
		ϑ	0.5	0.3385	0.1937	93.9	0.2952	0.1309	90.3	0.1453	0.0730	83.2
		λ	1.3	0.1574	0.1849	94.7	0.1347	0.1263	91.4	0.1062	0.0735	85.2
		p	1.0	0.1385	0.0514	95.1	0.1077	0.0399	93.1	0.0642	0.0145	82.1
	moderate	m	1.0	0.6330	0.4461	95.6	0.6105	0.3857	91.7	0.5408	0.2788	90.4
	U[0,7]	n	1.4	0.6463	0.3511	94.3	0.5957	0.2535	92.7	0.4627	0.1044	87.6
	50%	α	1.8	0.9038	0.1876	94.3	0.2013	0.1025	90.5	0.1869	0.0582	86.4
		β	0.5	0.3063	0.1027	92.6	0.2390	0.2300	96.4	0.1748	0.1272	84.2
		ϑ	0.9	0.2331	0.1709	93.7	0.2187	0.1274	92.8	0.1651	0.0839	87.2
		λ	1.3	0.1646	0.1076	92.6	0.1289	0.0803	90.9	0.0751	0.0454	83.8

		p				1.0	0.1238	0.1461	94.5	0.0932	0.1190	92.9	0.0439	0.0632	84.4
							0.4305	0.1090	93.2	0.3852	0.0454		0.3186	0.0269	
	slight	*α*				1.7	0.6557	0.4608	92.4	0.6211	0.1095	90.3	0.5107	0.3113	86.2
	U[0,25]	*β*				1.0	0.1916	0.1272	94.2	0.1604	0.1180	93.2	0.4280	0.0665	85.8
BW	20%	*γ*				1.0	0.2057	0.1711	94.3	0.1655	0.1269	96.1	0.1277	0.0662	85.8
		λ				2.5	0.1720	0.3389	95.0	0.1355	0.2569	92.8	0.0783	0.2040	84.6
		p				0.5	0.4801	1.1891	92.4	0.1434	0.4802	94.3	0.1085	0.3203	82.4
	moderate	*α*				1.7	0.6600	0.5588	90.4	0.6647	0.5360	90.3	0.0761	0.4412	80.0
	U[0,8]	*β*				1.0	0.1093	0.2394	93.2	0.1590	0.1054	89.3	0.0938	0.0797	91.8
	50%	*γ*				1.0	0.1977	0.1754	95.2	0.1373	0.1115	94.2	0.1086	0.0573	84.6
		λ				2.5									83.8
			P	0.5	0.1373		0.2816	95.0	0.0984	0.1956	92.9	0.0461	0.1084	87.8	
	slight	m				1.3	0.1496			0.1080	0.1145	91.5	0.0646	0.0763	90.6
	U[0,23]	n				0.7	0.1778		93.38		0.1528	94.3		0.0893	82.8
	20%	*α*				0.8	0.1277			0.0980	0.0546	95.6	0.0456	0.0123	95.0
							0.1993		94.32						
							0.1243			0.073			0.0312		
BGE							0.1018		96.07						
		λ				2.5	0.3454			0.2890	0.1172	90.2	0.2372	0.0829	83.4
							0.1430		91.36						
		p				0.9	0.2089			0.1580	0.1268	92.1	0.1132	0.0932	84.2
							0.1687		93.26						
	moderate	m				1.3	0.1647			0.1200	0.1149	92.0	0.0952	0.0575	89.4
							0.1529		93.23						
	U[0,6]	n				0.7	0.1409		93.8	0.1060	0.0810	90.4	0.0670	0.0686	85.0
							0.1158			0.1067					
	50%	*α*				0.8	0.5051		93.2	0.4550	0.2970	90.7	0.3704	0.2320	84.0
		λ				2.5	0.1136		90.38		0.5357	89.7		0.2380	73.0
		p				0.9	0.2131		90.3	0.2450	0.0820	82.7	0.2061	0.0616	74.6
							0.3281			0.1377			0.0851		
							0.3024								
							0.1443								

9. Simulation Study

The joint Beta Kum-BX, BW and BGE models, simulations with cure fraction were done, where the two-parameter Kumaraswamy-model was used to generate the probability of cure. A uniform distribution, U[a,b], was used to generate censoring times, with constants a and b defined in order to give chosen censoring rates. The error distribution for the failure times of the uncured patients followed the two-parameter Weibull distribution. The Kumaraswamy Weibull data was generated in samples sizes of n = 30, 50 and 100 with 1000 replications. Also, 200 bootstrap samples were chosen for the simulation process due to the complexity of the models.

9.1. Discussion on (Simulation Study) Table 4

The results from the mixture model simulations show some similar outputs and trends for the non-mixture model of the Beta Kum-BX and BW. Unlike the BGE at all levels of the comparison either the slight or moderate at different censoring rates using the uniform distribution to generate the censored data for the simulation as explain above value of show relatively small biases overall. Again, mean square error decreased with increasing sample size and increased with higher censoring rates. The confidence interval values were considered for the Bayesian methods for comparison.

10. Conclusion

In the life time data analysis, we presented the cure fraction "*p"* and covariates "η0 = age at enrollment, η1= sex and η2= race". A data with this structure can be appropriately analyzed using different parametric formulations, as a mixture and non-mixture models. In this paper, we showed that parametric models based on the Beta Kum-BX and other existing models can be useful to analyze medical data sets. We showed that the use of Bayesian methodology using MCMC methods is a suitable way to get the inferences for the parameters of the model as the ACTG censored study data was well fitted and accommodate by Beta Kum-BX which is three confounded models with many advantages and properties in modeling a very large population in medical and engineering aspects. These approaches allow the inclusion of covariates in the model. The parameter estimation was obtained by the classical frequentist approach known as the maximum likelihood method and Bayesian approach using the Markov chain Monte Carlo simulation methods using ACTG censored study data. The Winbugs software and MCMCpack library in R was applied for the Gibbs sampling algorithm to obtain the posterior summaries of interest. Finally, a simulation study was employed based on the cure models to compare the performance of both models relating to actual sense of motivation to clarify the usefulness of the proposed methodology. An advantage of Bayesian approach over the other methods is that it explicitly incorporates the expert prior opinion for the parameters.

Acknowledgments

The authors are very grateful to the Yobe State University, Damaturu and Universiti Putra Malaysia for sponsoring the fund together with the referee for suggestions and comments which significantly improved the quality of the manuscript.

References

[1] Madaki, U. Y., Bakar, M. R. A., & Handique, L. (2018). Beta Kumaraswamy Burr type X distribution and its properties., MHPD Journals, 2, 23-32

[2] M. Alizadeh, M. Tahir, G. M. Cordeiro, M. Mansoor, M. Zubair, G. Hamedani, The kumaraswamy marshal-olkin family of distributions, Journal of the Egyptian Mathematical Society 23 (3) (2015) 546–557.

[3] G. M. Cordeiro, M. de Castro, A new family of generalized distributions, Journal of statistical computation and simulation 81 (7) (2011) 883–898.

[4] I. W. Burr, Cumulative frequency functions, The Annals of Mathematical Statistics 13 (2) (1942) 215–232.

[5] F. Merovci, M. A. Khaleel, N. A. Ibrahim, M. Shitan (2016), The beta burr type x distribution properties with application, SpringerPlus. 1–18.

[6] M. U. Yusuf, M. R. B. A. Bakar (2016), A bayesian estimation on right censored survival data with mixture and non-mixture cured fraction model based on beta-weibull distribution, in: AIP Conference Proceedings, Vol. 20079, 1739.

[7] M. U. Yusuf, M. R. A. Bakar, N. A. Ibrahim, J. Arasan, Q. N. Husain, A bayesian via laplace approximation on log-gamma model with censored data, Modern Applied Science 11 (1) (2016) 14.

[8] J. A. Achcar, E. A. Coelho-Barros, J. Mazucheli, Cure fraction models using mixture and non-mixture models, Tatra Mountains Mathematical Publications 51 (1) (2012) 1–9.

[9] J. Berkson, R. P. Gage, Survival curve for cancer patients following treatment, Journal of the American Statistical Association 47 (259) (1952) 501– 515.

[10] J. W. Boag, Maximum likelihood estimates of the proportion of patients cured by cancer therapy, Journal of the Royal Statistical Society. Series B (Methodological) 11 (1) (1949) 15–53.

[11] E. Z. Martinez, J. A. Achcar, A. A. Jacome, J. S. Santos, Mixture and non-mixture cure fraction models based on the generalized modified Weibull distribution with an application to gastric cancer data, Computer methods and programs in biomedicine 112 (3) (2013) 343–355.
[12] M.-H. Chen, J. G. Ibrahim, D. Sinha, A new bayesian model for survival data with a surviving fraction, Journal of the American Statistical Association 94 (447) (1999) 909–919.
[13] J. G. Ibrahim, M.-H. Chen, D. Sinha, Bayesian semiparametric models for survival data with a cure fraction, Biometrics 57 (2) (2001) 383–388.
[14] R. A. Maller, X. Zhou, Survival analysis with long-term survivors, Wiley New York, 1996.
[15] F. Famoye, C. Lee, O. Olumolade, The beta-weibull distribution, Journal of Statistical Theory and Applications 4 (2) (2005) 121–136.
[16] G. M. Cordeiro, A. B. Simas, B. D. Stoˇsic, Closed form expressions for moments of the beta weibull distribution, Anais da Academia Brasileira de Ciˆencias 83 (2) (2011) 357–373.
[17] G. M. Cordeiro, S. Nadarajah, E. M. Ortega, General results for the beta weibull distribution, Journal of Statistical Computation and Simulation 83 (6) (2013) 1082–1114.
[18] A. S. Wahed, T. M. Luong, J.-H. Jeong, A new generalization of weibull distribution with application to a breast cancer data set, Statistics in medicine 28 (16) (2009) 2077–2094.
[19] W. Weibull, Wide applicability, Journal of applied mechanics 103 (1951) 293–297.
[20] N. A. Ibrahim, F. Taweab, J. Arasan, A parametric non-mixture cure survival model with censored data, in: Computational Problems in Engineering, Springer, 2014, pp. 231–238.
[21] C. A. Santos, J. Alberto Achcar, A bayesian analysis in the presence of covariates for multivariate survival data: An example of application, Revist Colombiana de Estad´ıstica 34 (1) (2011) 111–131.
[22] Eugene, N., Lee, C., & Famoye, F. (2002). Beta-normal distribution and its applications. Communications in Statistics-Theory and methods, 31(4), 497-512.
[23] W. Barreto-Souza, A. H. Santos, G. M. Cordeiro, The beta generalized exponential distribution, Journal of Statistical Computation and Simulation 80 (2) (2010) 159–172.
[24] A. E. Gelfand, D. K. Dey, H. Chang, Model determination using predictive distributions with implementation via sampling-based methods, Tech. rep., DTIC Document (1992).
[25] D. W. Hosmer, S. Lemeshow, S. May, Model development, Applied Survival Analysis: Regression Modeling of Time-to-Event Data, Second Edition (2008) 132–168.
[26] J. Geweke, et al., Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments, Vol. 196, Federal Reserve Bank of Minneapolis, Research Department Minneapolis, MN, USA, 1991.