Noname manuscript No.
(will be inserted by the editor)

Bayesian analysis of survival data under generalized extreme value distribution with application in cure rate model

Dooti Roy[1] · **Vivekananda Roy**[2] · **Dipak K. Dey**[1]

Received: date / Accepted: date

Abstract This paper introduces both maxima and minima generalized extreme value (GEV) distribution to analyze right-censored survival data with a cure fraction. Our proposed GEV model leads to extremely flexible hazard functions. Our proposed Bayesian model achieves proper posterior distribution under some weak conditions even when improper priors are used. We further provide theoretical and numerical results showing that our GEV models offer a richer class of models than the widely used Weibull models. Finally, a glioblastoma multiforme cancer data is analyzed to illustrate the proposed GEV model.

Keywords Cure fraction · Generalized extreme value distribution · Gibb's sampler · MCMC · Metropolis Hastings · Survival analysis · Real data analysis · Seer · Glioblastoma Multiforme · Melanoma of the skin

1 Introduction

Ever since 1936, when Richard Von Mises had studied Extreme Value Theory (EVT) for the very first time, it has become a popular tool to model risk as risky events by definition happen with low probability. Today Generalized Extreme Value (GEV) distribution finds its use across several disciplines like reliability, hydrology, meteorology and finance. However its usefulness in survival modeling domain remains unexplored. In [12], the authors show that GEV gives rise to extremely flexible hazard functions with slight variations of the scale and the shape parameters. This flexible hazard functions will be useful in several practical situations. For example, when studying the disease

Dooti Roy
[1]Department of Statistics,University of Connecticut, Storrs, Connecticut
E-mail: deydipak@gmail.com

Vivekananda Roy
[2]Department of Statistics and Statistical Laboratory, Iowa State University, Ames, IA

span of a cancer patient, it is quite possible that a three-phase behavior of the failure rate will be observed. The initial risk of survival before the cancer reacts to medications, retreats and then after a span spikes again. A model with a bathtub or "U" shaped failure rate would be appropriate to describe the population's survival capacity. Extending the results of [12] in this article, we build flexible survival models using both the maxima and minima GEV distribution for log T, where T denotes the failure time/ survival time of an individual. We are able to show that by changing the shape parameter of the GEV distribution, a variety of shape for the hazard function including upside-down and bathtub shapes are obtained.

During the last two decades, several cancer related studies have demonstrated that a significant fraction of the diseased population has been cured. This can be attributed to the development of the advanced medical techniques as well as timely diagnosis of the cancer. It is due to this fact, that survival models incorporating a cure fraction have gained significant attention and importance in the literature. We use the log GEV survival models mentioned above to construct flexible cure rate model.

The cure rate model is popularly used for modeling time-to-event survival data. Perhaps the most popular type of cure rate model is the mixture model introduced by [1] around 60 years ago. This binary mixture model has played an important role in the literature and has been discussed by multiple authors including [4], [8], [9], [11] and [16] among others.

In [2], it is shown that the cure rate model discussed above has several drawbacks and the authors proposed a different kind of cure rate model. This proposed model has a proportional hazards structure through the cure rate parameter, and thus has an appealing interpretation. Unlike the standard cure rate model, the model in [2] yields proper posterior distributions under a wide class of non-informative improper priors for the regression coefficients, including an improper uniform prior. Also by introducing latent variables, posterior samples for parameters can be efficiently obtained for this model.

In [2], the proposed model uses Gamma and Weibull distributions, which are quite popular with monotone hazard rates. However, it is well known that in practical situations for survival analysis, the hazard function is often not monotone and is either upside-down shaped or bathtub shaped or a combination of both shapes.

In this paper, we introduce cure rate models using GEV distributions. We use Bayesian methodology to do the inference for parameters. An important issue in Bayesian analysis is the specification of a prior distribution. This is especially true in survival analysis when one wants to assess the importance of certain prognostic factors such as age, gender, etc. It may be difficult to specify informative priors for all possible candidate models, especially if little prior information is available. When such prior elicitation is difficult or when little prior information is available, one may consider analyses with conventionally chosen non-informative priors, such as the improper uniform prior. We provide sufficient conditions under which the joint posterior distribution of the parameters of interest is proper when improper uniform prior is used for regression

parameters for covariates incorporated through the cure rate parameter. This prior is analytically and computationally attractive and facilitates a direct comparison with maximum likelihood.

The article is organized as follows. In Section 2 we provide a short introduction to the standard log generalized extreme value distribution for maxima and minima. In Section 3 we derive the likelihood function for cure rate models with covariates and consider a useful class of non informative prior distributions. We also derive several results regarding proprieties of the resulting posterior distributions. In Section 4 we consider some numerical simulation examples. In this section we compare both GEV maxima and minima models with the Weibull distribution model to show the superiority of our proposed models. We carry out model selection using Deviance Information Criterion (DIC) and Log Pseudo Marginal Likelihood (LPML) measures and also compute the model diagnostics for GEV minima (mGEV) model under perturbation. In Section 5 we present a real glioblastoma multiforme data set from SEER ([15]) to illustrate our proposed methodology. We conclude with a brief discussion in Section 6. Supplementary materials contains proofs of some of the theorems and another real data example using a melanoma cancer data set.

2 Generalized extreme value distribution

2.1 Introduction to the Generalized extreme value distribution

Suppose $Y_1, Y_2, ...$ is a sequence of independent and identically distributed random variables each having the distribution function $F(y)$. Let $M_n = \max\{Y_1, Y_2, ..., Y_n\}$ and $m_n = \min\{Y_1, Y_2, ..., Y_n\}$. If the distribution of Y_i is specified then the exact distributions of M_n and m_n are known. Besides, extreme value theory also considers the existence of the limiting distributions of M_n and m_n. If there exists a non-degenerate distribution function $G_\xi(x)$, and a pair of sequences a_n, b_n,with $a_n > 0$, such that

$$\lim_{n\to\infty} P\{a_n^{-1}(M_n - b_n) \leq x\} = G_\xi(x) \tag{1}$$

on all points in the continuity set of $G_\xi(x)$, we say that $G_\xi(x)$ is a generalized extreme value distribution for maxima. The possible forms of $G_\xi(x)$ are completely specified as follows:

$$G_\xi(x) = \begin{cases} \exp[-(1+\xi\frac{x-\mu}{\sigma})_+^{-\frac{1}{\xi}}] & \text{if } \xi > 0 \text{ or } \xi < 0 \\ \exp[-\exp(-\frac{x-\mu}{\sigma})] & \text{if } \xi = 0, \end{cases}$$

where $\mu \in R$, $\sigma \in R^+$, and $\xi \in R$ are the location, scale, and shape parameters respectively, and $x^+ = \max(x, 0)$. Similarly, we could get the limiting distribution of m_n, called the generalized extreme value distribution for minima. Cumulative distribution function of the GEV distribution for minima is given

by:

$$G_\xi(x) = \begin{cases} 1 - \exp[-(1+\xi\frac{x-\mu}{\sigma})_+^{\frac{1}{\xi}}] & \text{if } \xi > 0 \text{ or } \xi < 0 \\ 1 - \exp[-\exp(\frac{x-\mu}{\sigma})] & \text{if } \xi = 0, \end{cases}$$

where $\mu \in R$, $\sigma \in R^+$, and $\xi \in R$ are the location, scale, and shape parameters, respectively. In electrical engineering literature, the models for minima are often dealt with.

2.2 Shape of Hazard function for the logGEV distribution

Going forward, we will use the notations MGEV and mGEV for representing the GEV maxima model and the GEV minima model respectively.

If we assume a MGEV distribution for log T, i.e., $\log T \sim \text{MGEV}(\mu, \sigma, \xi)$, then the corresponding cdf and pdf for T, i.e. $\text{T} \sim \log MGEV(\mu, \sigma, \xi)$ are respectively

$$F_M(t|\mu,\sigma,\xi) = \Psi_{MGEV}\left(\frac{\log(t)-\mu}{\sigma}\right),$$

$$f_M(t|\mu,\sigma,\xi) = \frac{1}{\sigma t}\psi_{MGEV}\left(\frac{\log(t)-\mu}{\sigma}\right),$$

where Ψ_{MGEV} and ψ_{MGEV} are the cdf and pdf respectively for the standardized MGEV distribution.

In other words,

$$f_M(t|\mu,\sigma,\xi) = \begin{cases} \frac{\exp[-(1+\xi\frac{\log t-\mu}{\sigma})^{-\frac{1}{\xi}}]}{\sigma t(1+\xi\frac{\log t-\mu}{\sigma})^{\frac{1}{\xi}+1}} & t > \exp(\mu - \frac{\sigma}{\xi}) \text{ if } \xi > 0 \text{ or} \\ & t < \exp(\mu - \frac{\sigma}{\xi}) \text{ if } \xi < 0 \\ \frac{1}{\sigma t}\exp(-\frac{\log t-\mu}{\sigma})\exp[-\exp(-\frac{\log t-\mu}{\sigma})] & 0 < t < \infty \text{ if } \xi = 0 \end{cases}$$

and the survival function which is $(1 - F_M(t|\mu,\sigma,\xi))$ is given by,

$$S_M(t|\mu,\sigma,\xi) = \begin{cases} 1 - \exp[-(1+\xi\frac{\log t-\mu}{\sigma})_+^{-\frac{1}{\xi}}] & \text{if } \xi \neq 0 \\ 1 - \exp(-\exp(-\frac{\log t-\mu}{\sigma})) & \text{if } \xi = 0 \end{cases}$$

In the special case when $\mu = 0$ and $\sigma = 1$ as mentioned in the [12], the density of T is

$$f_M(t|\xi) = \begin{cases} \frac{\exp[-(1+\xi\log t)^{-\frac{1}{\xi}}]}{t(1+\xi\log t)^{\frac{1}{\xi}+1}} & t > \exp(-\frac{1}{\xi}) \text{ if } \xi > 0 \text{ or} \\ & t < \exp(-\frac{1}{\xi}) \text{ if } \xi < 0 \\ \frac{1}{t^2}\exp(-\frac{1}{t}) & 0 < t < \infty \text{ if } \xi = 0, \end{cases}$$

and the corresponding survival function is

$$S_M(t|\xi) = \begin{cases} 1 - \exp[-(1+\xi\log t)_+^{-\frac{1}{\xi}}] & \text{if } \xi \neq 0 \\ 1 - \exp(-\frac{1}{t}) & \text{if } \xi = 0. \end{cases}$$

Hence, the corresponding hazard function $\lambda_M(t|\xi) = f(t|\xi)/S(t|\xi)$ is given by

$$\lambda_M(t|\xi) = \begin{cases} \frac{1}{t(1+\xi\log t)_+^{\frac{1}{\xi}+1}[\exp(1+\xi\log t)_+^{-\frac{1}{\xi}}-1]} & \text{if } \xi \neq 0 \\ \frac{1}{t^2[\exp(\frac{1}{t})-1]} & \text{if } \xi = 0. \end{cases}$$

Figure 1(a) shows the plot of the hazard function of MGEV($\mu = 0, \sigma = 1, \xi$) for different values of ξ. [12] used this plot to show that the MGEV model is extremely flexible in modeling survival data.

On the other hand, if we assume a mGEV distribution for log T, i.e., log T $\sim$ mGEV(μ, σ, ξ) , the corresponding cdf and pdf for T are

$$F_m(t|\mu,\sigma,\xi) = \Psi_{mGEV}\left(\frac{\log(t)-\mu}{\sigma}\right),$$

$$f_m(t|\mu,\sigma,\xi) = \frac{1}{\sigma t}\psi_{mGEV}\left(\frac{\log(t)-\mu}{\sigma}\right),$$

where Ψ_{mGEV} and ψ_{mGEV} are the cdf and pdf respectively for the standardized mGEV distribution. When $\mu = 0$ and $\sigma = 1$, the density of T could be written as:

$$f_m(t|\xi) = \begin{cases} \frac{1}{t}(1+\xi\log t)^{\frac{1}{\xi}-1}\exp[-(1+\xi\log t)^{\frac{1}{\xi}}] & t > \exp(-\frac{1}{\xi}) \text{ if } \xi > 0 \text{ or} \\ & t < \exp(-\frac{1}{\xi}) \text{ if } \xi < 0 \\ \exp(-t) & 0 < t < \infty \text{ if } \xi = 0. \end{cases}$$

The corresponding survival and the hazard functions respectively are given as follows,

$$S_m(t|\xi) = \begin{cases} \exp[-(1+\xi\log t)_+^{\frac{1}{\xi}}] & \text{if } \xi \neq 0 \\ \exp(-t) & \text{if } \xi = 0, \end{cases}$$

$$\lambda_m(t|\xi) = \begin{cases} \frac{1}{t}(1+\xi\log t)^{\frac{1}{\xi}-1} & \text{if } \xi \neq 0 \\ 1 & \text{if } \xi = 0. \end{cases}$$

Figure 1(b) shows the plot of the above hazard function of mGEV($\mu = 0, \sigma = 1, \xi$) for three different values of $\xi = -0.03, 0, 0.03$. Note that when $\xi = 0$, this is the hazard from of Weibull$(1, 1)$ (See lemma 1 in Section 4.2). Although the value of ξ changes only slightly, the hazard functions change significantly, especially when the time-to-failure is very small. Figure 1(c) shows the plot of the hazard function of mGEV($\mu = 0, \sigma = 1.5, \xi$) with three different values of $\xi = -0.03, 0, 0.03$. Note that the change of the scale parameter σ results in totally different shapes of the plot for hazard functions from that in Figure 1(b). Thus the mGEV(μ,σ,ξ) distribution leads to extremely flexible hazard functions.

3 Modeling log survival data with a cure fraction with GEV distributions

Cure rate models have been used for modeling time-to-event data for various types of cancers. For these diseases, a significant proportion of patients are "cured". As mentioned in the introductory section, in the popular standard cure rate model ([1]) the survival function for the entire population, denoted by $S_1(t)$, is given by $S_1(t) = \pi + (1-\pi)S^*(t)$, where a fraction π of the population are considered "cured" and the remaining $1-\pi$ are "non-cured", and $S^*(t)$ denotes the survival function for the non-cured group in the population. This model, though attractive, still has several drawbacks. For example, when including covariates through π, we might get improper posterior distributions for many types of non-informative improper priors. This is a serious drawback, since a proper posterior is required if we want to obtain Bayesian inference. To overcome the drawbacks of the standard cure rate model, a different type of cure rate model is introduced in [2] and the specified distribution is the Weibull distribution. In Lemma 1, it is proved that the logarithm of the Weibull distribution is a special case of the generalized extreme value distribution. In this paper, we apply the logarithm of the generalized extreme value distribution to the model given in [2] to incorporate a larger class of models.

Suppose that we have n subjects, and let N_i denote the number of clonogenic cells for the ith subject. Further, assume that the N_i's are i.i.d. Poisson random variables with mean θ_i, $i = 1, \ldots, n$. We emphasize here that the N_i's are not observed and can be viewed as latent variables. The incubation times for the N_i clonogenic cells for the ith subject, which are unobserved, are assumed to be i.i.d. with common cdf $F(\cdot)$, $i = 1, \ldots, n$. Let t_i denote the failure time for subject i, where t_i is right-censored. Let c_i denote the censoring time, so that we observe $y_i = \min(t_i, c_i)$, where the censoring indicator $\delta_i = I(t_i < c_i)$ equals 1 if t_i is a failure time and 0 if it is right-censored. Represent the observed data by the vector $(n, \mathbf{y}, \delta)$, where $\mathbf{y} = (y_1, \cdots, y_n)$ and $\delta = (\delta_1, \cdots, \delta_n)$. Also, let $\mathbf{N} = (N_1, \cdots, N_n)$, $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_n)$. The complete data are then given by $\mathbf{D} = (n, \mathbf{y}, \delta, \mathbf{N})$, where $\mathbf{N}$ is the unobserved vector of latent variables. We assume the density for y_i is $f(y_i|\xi)$ and the corresponding survival function is $S(y_i|\xi)$. Here, ξ is the shape parameter in the standard MGEV or mGEV distribution which we will use later. The complete data likelihood function of the parameters $(\boldsymbol{\theta}, \xi)$ can be written as

$$L(\boldsymbol{\theta}, \xi|D) = \prod_{i=1}^{n} S(y_i|\xi)^{N_i - \delta_i} (N_i f(y_i|\xi))^{\delta_i} \times \exp\Big\{ \sum_{i=1}^{n} (N_i \log(\theta_i) - \log(N_i!) - \theta_i) \Big\}. \tag{2}$$

Now we incorporate the covariates through θ. For each subject, $i = 1, \cdots, n$, let $\mathbf{x}_i' = (x_{i1}, \cdots, x_{ik})$ denote the $k \times 1$ vector of covariates for the ith subject, and let $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_k)$ denote the corresponding vector of regression coefficients. We relate $\boldsymbol{\theta}$ to the covariates by $\theta_i = \exp(\mathbf{x}_i'\beta)$, $i = 1, \cdots, n$. so that the chance of cure for the i'th subject is given by $P(N_i = 0) = \exp(-\theta_i) =$

$\exp(-\exp(\mathbf{x}_i'\beta))$ Thus we could write the complete-data likelihood of (β, ξ) as

$$L(\boldsymbol{\beta}, \xi|\mathbf{D}) = \prod_{i=1}^{n} S(y_i|\xi)^{N_i-\delta_i}(N_i f(y_i|\xi))^{\delta_i} \times \exp\left\{N_i\mathbf{x}_i'\boldsymbol{\beta} - \log(N_i!) - \exp(\mathbf{x}_i'\boldsymbol{\beta})\right\} \quad (3)$$

Assume that the prior distribution for (β, ξ) is $\pi(\beta, \xi)$, then the posterior distribution $\pi(\beta, \xi|\mathbf{D}_{obs})$ satisfy this:

$$\pi(\beta, \xi|\mathbf{D}_{obs}) \propto \sum_{\mathbb{N}} L(\beta, \xi|\mathbf{D})\pi(\beta, \xi). \quad (4)$$

3.1 MGEV Distribution for Right-Censored Data

Let t_i be the failure time for the i'th subject. Let c_i be the censoring time. Then $y_i = \min(t_i, c_i); i = 1, .., n$. Assume that $\log y_i \sim \text{MGEV}(\mu = 0, \sigma = 1, \xi)$. We use an improper uniform prior on β, that is, $\pi(\beta) \propto 1$, and the prior on ξ is $\pi(\xi) = 1/(t-s)I_{[s,t]}(\xi)$,where $s < 0 < t$, are fixed numbers. We also assume that $\pi(\beta, \xi) = \pi(\beta) \cdot \pi(\xi)$.

Theorem 1 *Let $\mathbf{X}^*$ be an $n \times k$ matrix with rows $\delta_i\mathbf{x}_i'$. If the following two conditions hold:*

1. *$\mathbf{X}^*$ is of full rank,*
2. *For every i with $\delta_i = 1$,*

$$\exp(-1/t) < y_i < \exp(-1/s) \text{ ; or } y_i \geq \max(\exp(-1/s), e), \quad (5)$$

then the posterior distribution given in (4) is proper.

The proof of the Theorem 1 is given in the supplementary materials.

In the special case when $t = -s = 1$, that is, when $\pi(\xi) = (1/2)I_{[-1,1]}(\xi)$, we have the following corollary.

Corollary 1 *Let $\mathbf{X}^*$ be an $n \times k$ matrix with rows $\delta_i\mathbf{x}_i'$. If the following two conditions hold:*

1. *$\mathbf{X}^*$ is of full column rank,*
2. *For every i with $\delta_i = 1$, $y_i > 1/e$,*

then the posterior distribution given in (4) is proper.

Remark 1 If $s = -t$ & $0 < t \leq 1$, (5) changes to $y_i > \exp(-1/t)$. This is a relatively weaker condition. If $s = -t$ & $t > 1$, (5) changes to $\exp(-1/t) < y_i < \exp(-1/s)$ or $y_i \geq e$.

Now we use the following prior on ξ: $\pi(\xi) = c\exp(-|\xi|/2)$, $-a < \xi < a, a > 0$, along with the previously used uniform prior on β. Simple calculations show that $c = 4(1-\exp(-a/2))$. The following theorem provides sufficient conditions for posterior propriety in this case. Our goal is to show that the posterior attains propriety under different reasonable priors.

Theorem 2 *Let* $\mathbf{X}^*$ *be an* $n \times k$ *matrix with rows* $\delta_i \mathbf{x}_i'$. *If these two conditions hold:*

1. $\mathbf{X}^*$ *is of full column rank,*
2. *For every* i *with* $\delta_i = 1$,

$$\exp(-1/a) < y_i < \exp(1/a) \quad or\, y_i \geq \max(\exp(1/a), e), \tag{6}$$

then the posterior distribution given in (4) is proper.

The proof of the Theorem 2 is similar to that of Theorem 1 and hence omitted.

3.2 mGEV Distribution for Right-Censored Data

As in Section 3.1, we can establish sufficient conditions under which the posterior density (4) is proper when a mGEV distribution for $\log y_i$ is used. Consider the uniform prior for ξ on (s, t), that is, $\pi(\xi) = 1/(t-s) I_{[s,t]}(\xi),\ s < 0 < t$, and $\pi(\beta) \propto 1$.

Theorem 3 *Let* $\mathbf{X}^*$ *be an* $n \times k$ *matrix with rows* $\delta_i \mathbf{x}_i'$. *If the following two conditions hold:*

1. $\mathbf{X}^*$ *is of full column rank,*
2. *For every* i *with* $\delta_i = 1$, $\exp(-1/t) < y_i < \exp(-1/s)$,

then the posterior distribution given in (4) is proper.

The proof of the Theorem 3 is given in supplementary materials. Following the proof of Theorem 3, similar results can be established for other priors on ξ considered in Section 3.1.

4 Simulation Studies

In this section we conduct simulation studies to evaluate model fitting performance along with model comparison, selection and model diagnostics against competing models.S

4.1 Simulation of MGEV model

Here we simulate a right-censored data set as described in Section 3. The simulation process is given below: 1. Let the sample size be $n = 1000$. One covariate (age) and an intercept term are included in the analysis. Assume $\beta = (2, 0.6)'$. In the analysis below, we first generate age as a random number between 1 to 100 with replacement and then standardized the simulated age to stabilize the posterior computations. So we have the $n \times 2$ covariate matrix X with the first column as $1's$, the second column as the standardized age. Denote the i'th row of X by $\mathrm{x}_i^{'}$, so $\theta_i = \exp(\mathrm{x}_i'\boldsymbol{\beta}),\ i = 1, ..., n$. 2. For every i, $i = 1, ..., n$,

we draw a sample from Poisson(θ_i), denoted by N_i. Then obtain a sample of size N_i from $\log MGEV(\mu = 0, \sigma = 1, \xi = 0.3)$, denote them as $Z_{i1}, \ldots, Z_{iN_i}$. Set $t_i = \min(Z_{i1}, \ldots, Z_{iN_i})$. Notice here, if $N_i = 0$, then set $t_i = \infty$. 3. We take $y_i = \min(t_i, c_i)$ and the indicator $\delta_i = I(t_i < c_i)$ equals 1 if y_i is a failure time and 0 if it is right-censored. For every i, $i = 1, ..., n$, let c_i denote the censoring time. Here, we choose c_i so as to let the censoring percentage be close to 16% . The censoring percentage can be moderated by varying c_i 4. Next we estimate the parameters using $\log T \sim$ MGEV($\mu = 0, \sigma = 1, \xi$) (denoted as the fitting model).

The estimated Kaplan-Meier curve is shown in Figure 2(a), which displays a plateau in the survival curve, and consequently a cure rate model is perhaps suitable for the simulated data set. In Figure 2(a) we also obtain the survival function estimate at all simulated survival times using the proposed MGEV model without incorporating any covariate information. We see that the two plots are nearly identical.

We now include covariates in the model and perform a Bayesian analysis. We use a non-informative prior on β ($\pi(\beta) \propto 1$), and the proper prior on ξ is $\pi(\xi) = \frac{1}{2} I_{[-1,1]}(\xi)$, and assume that $\pi(\beta, \xi) = \pi(\beta) \cdot \pi(\xi)$. In this example, 10,000 MCMC iterations are used in dealing with the posterior estimation of parameters after a burn-in of 1,000 iterations. Convergence is checked by observing trace plots, autocorrelations and ergodic mean plots. ([7]). The posterior estimates are shown in Table 1. Figure 2(b) shows two box plots of the MLE's of the cure rates and the posterior estimates of the cure rates. These two estimates of the cure rates are very similar.

4.2 Model comparison between mGEV and the Weibull distribution

We begin with presenting three lemmas which show that the popular Weibull, Rayleigh and the Exponential distributions are special cases of mGEV distribution.

Lemma 1 *If* $T \sim$ *Weibull*(α, λ), *then* $\log T \sim mGEV(\mu = \log(\lambda), \sigma = \frac{1}{\alpha}, \xi = 0)$.

Lemma 2 *If* $T \sim$ *Rayleigh*(λ), *then* $\log T \sim mGEV(\mu = \log(\sqrt{2}\lambda), \sigma = \frac{1}{2}, \xi = 0)$.

Lemma 3 *If* $T \sim$ *Exponential*(λ), *then* $\log T \sim mGEV(\mu = \log(\lambda), \sigma = 1, \xi = 0)$.

The proof of Lemma 1 is given in supplementary materials. Lemma 2 and 3 can be proved similarly.

Next we use two numerical simulations to compare these two distributions. Both simulation processes are almost the same as the simulation in Section 4.1 except in the following aspects:

For the first simulation, the true model is Weibull$(1.03, 1)$, and both mGEV$(\mu = 0, \sigma = 1, \xi)$ and Weibull$(\alpha, \lambda = 1)$ are fitted to compare the goodness of fit,

and the value of all $c_i's$ are set as 0.55. Here the goal was to examine the goodness of fit of mGEV even when the location and the scale parameter were fixed, thus allowing flexibility through only the shape parameter ξ.

For the second simulation, the true model is mGEV($\mu = 0, \sigma = 1, \xi = 0.5$), and again both Weibull$(\alpha, 1)$ as well as mGEV($\mu = 0, \sigma = 1, \xi$) are fitted. The value of all $c_i's$ are set to 0.5. Here the censoring time is chosen to make the censoring percentage of the survival time to be around 10% to match real life scenarios. The values of the $c_i's$ are chosen so as to keep the censoring percentage similar.

From Figure 2(c), we see that when fitting the simulated data set of the Weibull distribution with our proposed model of the $mGEV$ distribution and the Weibull distribution, the two survival function estimates are both quite similar to the Kaplan-Meier estimation for the survival function. This indicates that the proposed model with the $mGEV$ may be quite a good fit to the simulated data set. From Table 2, the estimated posterior means for β_0 and β_1 are 1.989 and 0.628 respectively, which are very close to the true values of β_0 and β_1 (2 and 0.6 respectively). For the parameter ξ in the $mGEV$ distribution, the posterior mean estimate for ξ is 0.011. These results show that the proposed model with $mGEV$ is quite adequate for the simulated data set.

However, from Figure 2(d) and the estimates from Table 2, when fitting the simulated data set of the $mGEV$ distribution with the Weibull distribution, the Kaplan-Meier estimates and the parametric estimates for the survival function do not match very well. The posterior means for β_0 and β_1 are also quite different from the true value of β_0 and β_1. The HPD intervals do not even contain the true parameter values for β_0.

4.3 Model Selection Criteria

Several methodologies allow us to compare different competing models for a given dataset and select the best fitting one. In this paper we use one of the most widely used model selection criteria in applied Bayesian research, which is derived from the conditional predictive ordinate (CPO) statistic. [6] and [5] give a more detailed description of CPO statistics and their applications. Let $\mathcal{D}$ be the full data and $\mathcal{D}^{(-i)}$ denote the data with i'th observation deleted.

In our model for an uncensored time to event ($\delta_i = 1$) we have the data likelihood for the i'th observation, $g(y_i|\vartheta) = (\theta_i f(y_i|\xi)) \exp\{-\theta_i(1 - S(y_i|\xi))\}$, and for a censored time, $g(y_i|\vartheta) = \exp\{-\theta_i(1 - S(y_i|\xi))\}$. We denote the posterior density of $\vartheta = (\beta_0, \beta_1, \xi)$ given $\mathcal{D}^{(-i)}$ by $\pi(\vartheta|\mathcal{D}^{(-i)})$, $i = 1, \cdots, n$. For the i'th observation,

$$CPO_i = \left\{ \int_{\mathcal{H}} \frac{\pi(\vartheta|\mathcal{D}^{(-i)})}{g(y_i|\vartheta)} d\vartheta \right\}^{-1}$$

A Monte Carlo estimate of CPO_i can be obtained by using a single MCMC sample of the posterior $\pi(\vartheta|\mathcal{D})$. Let $\vartheta^{(1)}, \cdots, \vartheta^{(Q)}$ be a sample of size Q from

the posterior distribution after the burn-in. A Monte Carlo approximation of CPO_i ([3]) is given by,

$$\widehat{CPO_i} = \left\{ \frac{1}{Q} \sum_{q=1}^{Q} \frac{1}{g(y_i|\vartheta^q)} \right\}^{-1}.$$

For model comparison, we use the log pseudo marginal likelihood (LPML) defined by $\sum_{i=1}^{n} \log(\widehat{CPO_i})$. The model with larger LPML provides better fit to the data.

In this paper, we also use deviance information criterion (DIC) proposed by [14]. This criterion is based on the posterior mean of the deviance, which can be approximated by $\bar{d} = \frac{1}{Q} \sum_{q=1}^{Q} d(\vartheta_q)$, where, $d(\vartheta) = -2 \sum_{i=1}^{n} \log[g(y_i|\vartheta)]$. The DIC can also be estimated from the MCMC output by $\widehat{DIC} = \bar{d} + \hat{P}_d = 2\bar{d} - \hat{d}$, where P_D is the effective number of parameters, which is defined as E$[d(\vartheta)|D]$ - $d[E(\vartheta)|D]$, where $d[E(\vartheta)]$ is the deviance evaluated at the posterior mean and is estimated as,

$$\widehat{d[E(\vartheta)]} = d\left(\frac{1}{Q} \sum_{q=1}^{Q} \beta_0^{(q)}, \frac{1}{Q} \sum_{q=1}^{Q} \beta_1^{(q)}, \frac{1}{Q} \sum_{q=1}^{Q} \xi^{(q)} \right).$$

A smaller DIC implies better fit to the data.

Exponentiated Weibull Distribution:
This distribution was introduced in [10] which is a generalization of the Weibull distribution and adds flexibility to the usual Weibull distribution by introducing an additional shape parameter. The respective pdf and cdf of the Exponentiated Weibull is given by:

$$f_{EW}(t|k,\lambda,\alpha) = \alpha \frac{k}{\lambda} \left[\frac{t}{\lambda} \right]^{k-1} \left[1 - \exp\left(-\frac{t}{\lambda}\right)^k \right]^{\alpha-1} \exp\left(-\frac{t}{\lambda}\right)^k,$$

$$F_{EW}(t|k,\lambda,\alpha) = \left[1 - \exp\left(-\frac{t}{\lambda} \right)^k \right]^{\alpha},$$

where $t > 0$, $k > 0$ is the first shape parameter, $\alpha > 0$ is the second shape parameter and $\lambda > 0$ is the scale parameter of the distribution. For model selection 1000 datasets are generated from each of the four distributions: mGEV(0,1,0.5), MGEV(0,1,0.5), Weibull(1.03,1) and Exponentiated Weibull (1.03,1,1.5). For the mGEV and the MGEV distributions, the μ and σ parameters are standard while we know from practical experience that ξ parameter rarely goes beyond [-0.5, 0.5]. For the Weibull model, value for α was chosen so that the mGEV(0,1,ξ) would correspond to it (Refer Lemma 1). The Exponentiated Weibull has an additional parameter so that the data has more flexibility than the usual Weibull model.

Three competing models (mGEV, MGEV and Weibull) are fitted to each of the above datasets. The censoring percentage for each simulated data from mGEV, Weibull and Exponentiated Weibull distribution is approximately 10%. For data simulated from MGEV distribution censoring percentage is close to 17%. The censoring percentage was varied keeping in mind real time scenarios. Table 3 gives the average LPML and DIC values for each model under each scenario. From Table 3, we find that the mGEV model fitted on simulated Weibull data performs marginally better and gives slightly higher average LPML and slightly lower average DIC than the competing Weibull model. However, the difference between the two averages is not very large and in many data sets, the Weibull fit produced better fits in terms of LPML and DIC criterion. The better performance of mGEV is not surprising since the Weibull model is a special case of the mGEV model (See Lemma 1). The fitted MGEV model performs worst among the three fitted models.

The Weibull model fitted on simulated mGEV data has the lowest average LPML and highest average DIC amongst the three competing models. The data simulated from MGEV model does not fit any other model except MGEV itself, which performs the best among the three. When the data is simulated from Exponentiated Weibull, although the Weibull model performs the best among all three fitted models, but the MGEV model comes very close to the Weibull model in terms of average LPML and DIC. It shows the flexibility of our proposed models.

4.4 Comparison of Influence Diagnostics between mGEV model and comparison with Weibull model

To examine the performance of the proposed mGEV model, we consider simulated datasets with one or more cases perturbed. As a baseline we consider the same data used for simulation of mGEV model in section 4.1. Cases 1, 200 and 600 were selected for perturbation. For creation of influential observations in the dataset, one or two or all three of the selected were chosen and the response variable was perturbed as follows: $\tilde{y_i} = y_i + 4 * s_y, i = 1, 200, 600$, where s_y is the standard deviation of the $y_i^{'}s$. This is a valid perturbation since most of the data can be assumed to lie between $\pm 3 * s_y$ limits.

Similarly for better comparison a data set is simulated with Weibull as the true model and similar cases of perturbation were enforced. Table 4 shows posterior inference for parameters β_0, β_1 and ξ and Monte Carlo estimates of DIC and LPML for each perturbed version of the original dataset. The estimates of β_0 and β_1 fluctuate but estimates of ξ are pretty stable under perturbation. The original fitted model turns out to be the best one according to DIC and LPML criteria. Table 5 shows posterior inference for the Weibull fit. We find that only the estimates of β_0 show sensitivity towards perturbation. The unperturbed model gives the best fit in terms of DIC and LPML. This puts our proposed mGEV model in advantage over Weibull model as our model shows more sensitivity to outliers/influential observations etc.

5 Application: Glioblastoma Multiforme (GBM) Data

We consider a GBM data set from the National Cancer Institute SEER database [15]. This particular data was first discussed in details in [13]. Glioblastoma multiforme is one of deadliest form of cancers with an extremely small surviving fraction. Although there are reports that certain patients are known to survive for more than 10 years, it has been observed that these patients are typically younger than 40 years at the time of diagnosis. The data has a patient population of 1725 subjects who are diagnosed with only GBM cancer between the year 1970 and 2004. This implies we are looking at a follow up window of 10 years or more. A patient surviving for more than 10 years after diagnosis is considered cured ([13]). All the subjects are considered "young adults" with age at diagnosis between 16 years to 39 years. All the cases are followed annually and vital status is recorded. Subjects who died due to the cancer were considered failed and the rest (those who died due to other causes, dropped, or survived until the end of the study) were considered censored. In [13], the authors observed a cure rate of 12% among the young adults. The variable considered in this analysis is: lifetime in months since diagnosis. Subjects with survival time 0 were removed from the data set. The covariates included were: age at diagnosis,gender, radiation (treated with any type of radiation or not) and marital status (married or divorced/separated/widowed/single). The covariates were selected based on the authors findings in [13]. Table 6 lists some basic information for the data set.

5.1 Model fitting using the mGEV model

In [13], the authors used a non mixture cure fraction model with a Weibull survival function. Since the Weibull distribution is a special case of the mGEV distribution, we fit both mGEV and Weibull models to the data. The complete-data likelihood function of the parameters $(\mu, \sigma, \xi, \theta)$ can be written as

$$L(\mu,\sigma,\xi,\theta|\mathbf{D}) = \prod_{i=1}^{n} S(y_i|\mu,\sigma,\xi)^{N_i-\delta_i}(N_i f(y_i|\mu,\sigma,\xi))^{\delta_i} \times \exp\{\sum_{i=1}^{n}(N_i\log(\theta_i) - \log(N_i!) - \theta_i)\}. \quad (7)$$

If we incorporate covariates through θ, then for each subject, $i = 1, \cdots, n$, let $\mathbf{x}_i' = (x_{i1}, \cdots, x_{ik})$ denote the $k \times 1$ vector of covariates for the ith subject, and let $\beta = (\beta_1, \cdots, \beta_k)$ denote the corresponding vector of regression coefficients. We relate θ to the covariates by $\theta_i = \exp(\mathbf{x}_i'\beta), i = 1, \cdots, n$. Thus we could

write the complete-data likelihood of $(\beta, \mu, \sigma, \xi)$ as

$$L(\beta, \mu, \sigma, \xi|\mathbf{D}) = \prod_{i=1}^{n} S(y_i|\mu, \sigma, \xi)^{N_i - \delta_i} (N_i f(y_i|\mu, \sigma, \xi))^{\delta_i} \times \exp\{\sum_{i=1}^{n} N_i \mathbf{x}_i' \beta - \log(N_i!) - \exp(\mathbf{x}_i'\beta)\}. \quad (8)$$

Assume that the prior distribution for $(\beta, \mu, \sigma, \xi)$ is $\pi(\beta, \mu, \sigma, \xi)$, then the posterior distribution satisfies:

$$\pi(\beta, \mu, \sigma, \xi|\mathbf{D}_{obs}) \propto \sum_{\mathbb{N}} L(\beta, \xi|\mathbf{D}) \pi(\beta, \mu, \sigma, \xi) \quad (9)$$

For simplicity, we assume independent priors on μ, σ, ξ and β. Specifically, we take $\beta \sim N_k(0, \sigma_\beta^2 I_k)$, a normal distribution distribution with mean vector zero and covariance matrix $\sigma_\beta^2 I_k$, where I_k is the $k \times k$ identity matrix. Also, we take $\mu \sim N(0, \sigma_\mu^2)$, $\sigma^2 \sim IG(a_\sigma, b_\sigma)$ and $\xi \sim Uniform(-a, a)$. We pick $a = 1$, $a_\sigma = 0.01$, $b_\sigma = 2$, $\sigma_\beta^2 = 36$ and $\sigma_\mu^2 = 16$. The choices of the hyper parameters are such that the posterior distribution is proper but diffused. Here we use the Gibbs sampler with Metropolis-Hastings steps to get the posterior samples for parameters. For this data set, we have 10,000 MCMC iterations. Convergence was checked using the trace plots, ergodic mean plots and also the autocorrelation plots for all the parameters. And we find that 1,000 iterations are adequate as a burn-in. Further we computed all HPD intervals for all 7 parameters.

5.2 Model fitting using the Weibull model

We also fit the Weibull(λ, α) model with covariates as in the previous case. Again for simplicity, we assume that λ, α and β are independent and assume normal priors for all the β as before. Also, we take $\alpha \sim Gamma(a_\alpha, b_\alpha)$ and $\lambda^2 \sim IG(a_\lambda, b_\lambda)$. We suppose $a_\lambda = 0.1$, $b_\lambda = 2$, $a_\alpha = 0.01$, $b_\alpha = 0.01$ and σ_β^2 $= 25$.

As before we use the Gibbs sampler with Metropolis-Hastings steps to get the posterior samples for parameters. 10,000 MCMC iterations were run to achieve convergence. We find that 1,000 iterations are adequate as a burn-in. We compute the posterior estimates and HPD intervals for all 6 parameters.

5.3 Model Comparison

We plot the difference of logarithm of CPO at each data point of the two models fitted (mGEV and Weibull) to get an initial idea of which model fit is better. Mathematically, the difference d_i is given as

$$d_i = \log CPO_{imGEV} - \log CPO_{iWEI}. \quad (10)$$

Most of the d_i greater than zero will imply that our proposed mGEV model is favored over the Weibull model. We also compute LPML and DIC values for each model fit to facilitate better comparison.

5.4 Results

The marginal estimated Kaplan-Meier curve in Figure 3(a) shows a clear plateau in the survival curve, so a cure rate model seems to be appropriate for the cancer data set. From the plot, the empirical cure rate is around 12% which is consistent with the authors findings in [13]. Figure 3(a) also shows the survival function estimates at all simulated survival times without incorporating the covariate information using the proposed mGEV model, which matches the Kaplan-Meier curves, and hence the proposed model might be a good fit to the data set. Figure 3(a) also includes the simulated curve for the fitted Weibull model. Without covariates there seems to be slight advantage to fit the mGEV model over the Weibull one. We now consider Bayesian analysis with the covariates included. Tables 7 and 8 shows the posterior estimates for all parameters.

Table 7 shows the estimate of ξ is non-zero and the 95% HPD interval does not contain 0, implying there is a positive probability that ξ is non-zero, justifying the need of modeling the data as mGEV which can accommodate an additional shape parameter. The estimates of the β 's from both the fitted models match in sign and they are also close in values. Figure 3(b) shows the CPO plot of the difference. 59.4% of the points lie above zero (blue dots) giving us an indication that mGEV model is better suited to the data in hand. Table 9 shows the LPML and DIC values of the two model fits. We find that our proposed model mGEV has significantly lower DIC and higher LPML in comparison to Weibull fit.

6 Conclusion

In this paper we have implemented a new form of survival modeling for right-censored data with a cure fraction using the generalized extreme value distribution. We propose modeling the log survival time with censoring as a GEV distribution. We show through the hazard and survival plots, that the proposed model achieves a lot of flexibility and thus has an obvious advantage over the commonly used Weibull models. We also establish sufficient conditions for the propriety of the posterior distribution when an improper uniform prior is used for the regression coefficients through cure rates. Our results could be extended to the models for the generalized extreme value distribution in the presence of frailty parameters. We have also developed model comparison and influence diagnostics based on LPML & DIC which sufficiently demonstrates the superiority of our proposed model. Under perturbation our mGEV model shows more sensitivity than compared to the Weibull model. We can also include spatial and temporal covariate effects in our proposed model. Although

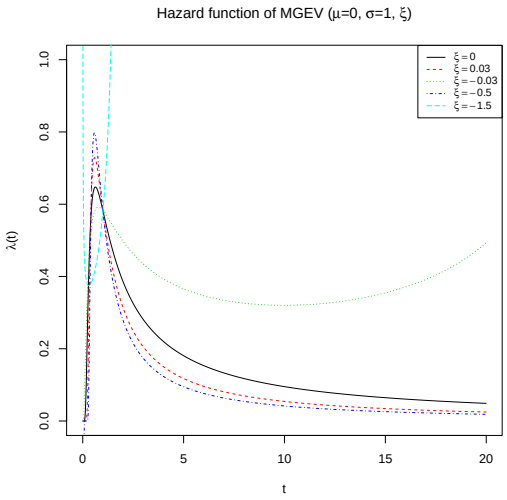
we only deal with right-censored data, the proposed methodology can be extended to other type of censoring like interval censored or left censored data. Another possible extension is to study the survival function with cure rate under proportional hazard structure with GEV as baseline hazard and further extend using semi parametric approach. The proposed methodology can also be extended to multiple cancer data. We leave that exercise for future work.

7 Supplementary Material

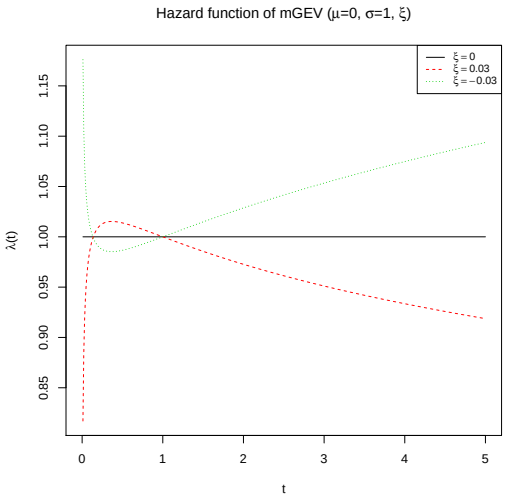
Supplementary material is available online. It includes the proofs of the theorems and lemmas as well as another real data application of the MGEV model.

References

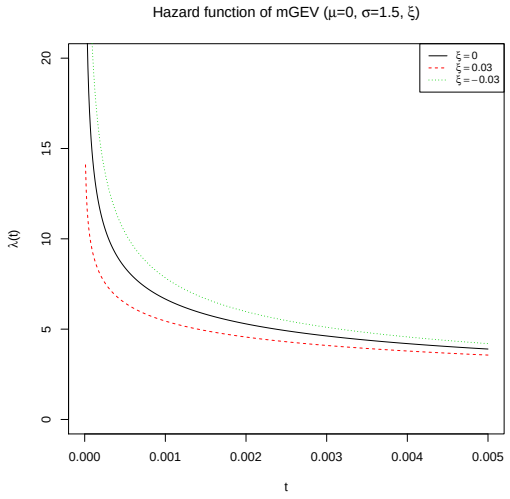
1. Berkson, J., Gage, R.P.: Survival curve for cancer patients following treatment. Journal of the American Statistical Association **47**, 501–515 (1952)
2. Chen, M., Ibrahim, J., Sinha, D.: A new bayesian model for survival data with a surviving fraction. Journal of the American Statistical Association **94**, 909–919 (1999)
3. Dey, D., Chen, M., Chang, H.: Bayesian approach for nonlinear random effects models. Biometrics **53**, 1239–1252 (1997)
4. Farewell, V.T.: The use of mixture models for the analysis of survival data with long-term survivors. Biometrics **38**, 1041–1046 (1982)
5. Geisser, S., Eddy, W.: A predictive approach to model selection. Journal of the American Statistical Association **74**, 153–160 (1979)
6. Gelfand, A., Dey, D., Chang, H.: Model Determination using predictive distributions with implementation via sampling based methods(with discussion). Oxford University Press
7. Junker, B.W., Patz, R.J., VanHoudnos, N.: Handbook of model item response theory. Chapman and Hall/CRC (2013)
8. Kuk, A.Y.C., Chen, M.H.: A mixture model combining logistic regression with proportional hazards regression. Biometrika **79**, 731–739 (1992)
9. Maller, R., Zhou, X.: Survival analysis with long-term survivors. John Wiley and Sons, New York (1996)
10. Mudholkar, G., Srivastava, D.: Exponentiated weibull family for analyzing bathtub failure-rate data. Reliability, IEEE Transactions **42**, 299–302 (1993)
11. Peng, Y., Dear, K.B.J.: A nonparametric mixture model for cure rate estimation. Biometrics **56**, 237–243 (2000)
12. Roy, V., Dey, D.: Propriety of posterior distributions arising in catgorical and survival models under generalized extreme value distribution. Statistica Sinica **2**, 699 – 722 (2014)
13. Smoll, N.R., Schaller, K., Gautschi, O.P.: The cure fraction of glioblastoma multiforme. Neuroepidemiology **39**, 63–69 (2012)
14. Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Linde, A.V.D.: Bayesian measures of model complexity and fit. Journal of Royal Statistical Society B **64**, 583–639 (2002)
15. Surveillance, E., End Results (SEER) Program, w.S.: Total u.s., 1969-2010 counties, national cancer institute, dccps,surveillance research program, surveillance systems branch. Incidence - SEER 17 Regs Research Data, Nov 2011 Sub (1973-2010), Katrina/Rita Population Adjustment, Linked To County Attributes (2013)
16. Sy, J.P., Taylor, J.M.G.: Estimation in a cox proportional hazards cure model. Biometrics **56**, 227–236 (2000)



(a)



(b)



(c)

Fig. 1:
1(a) Hazard functions of the generalized extreme value distribution for MGEV($\mu = 0, \sigma = 1, \xi$) for different values of ξ.
1(b) Hazard functions of the generalized extreme value distribution for mGEV($\mu = 0, \sigma = 1, \xi$) for different values of ξ.
1(c) Hazard functions of the generalized extreme value distribution for mGEV($\mu = 0, \sigma = 1.5, \xi$) for different values ofξ .

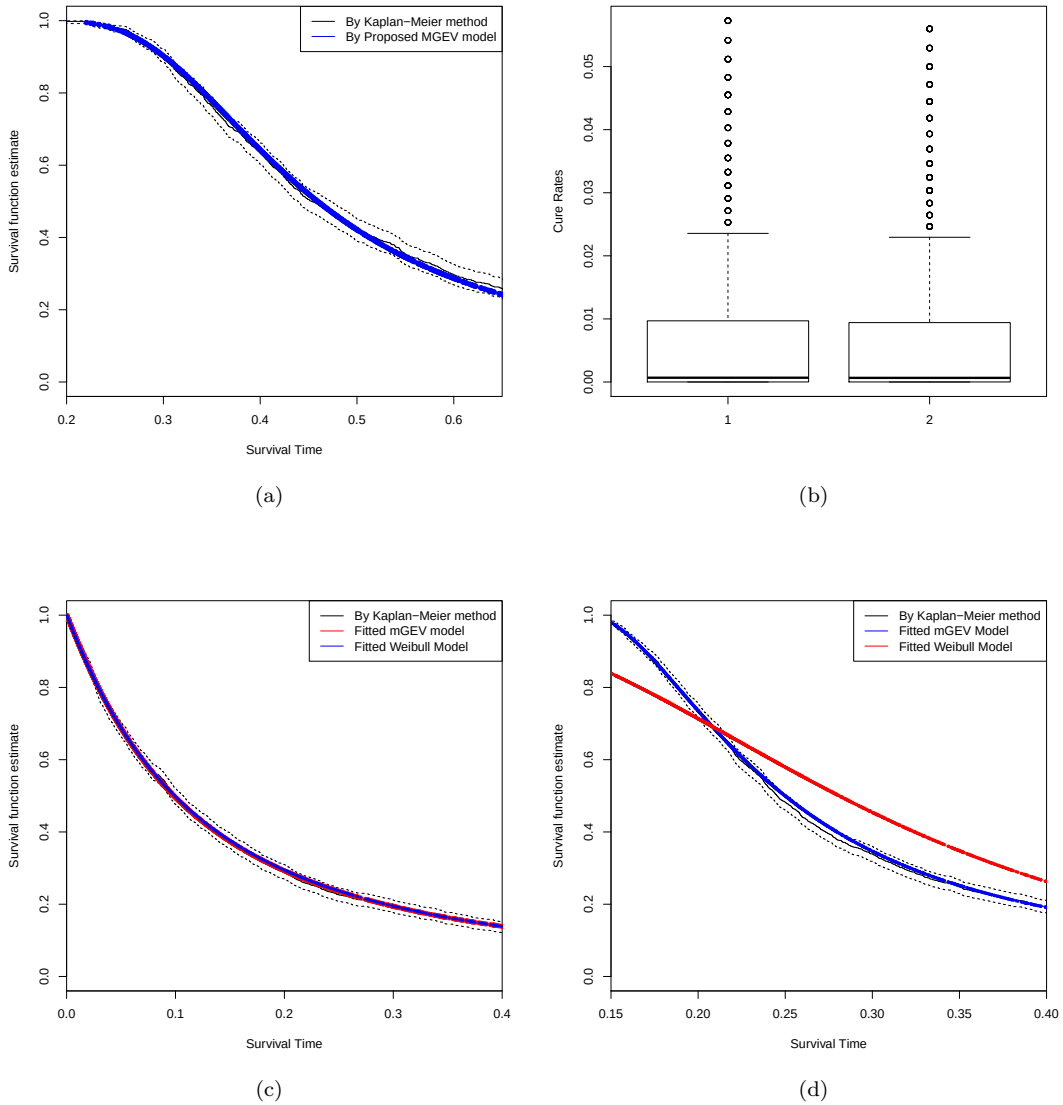


Fig. 2:
2(a) Estimated survival curves for the simulated data by Kaplan-Meier method (solid line is the estimate, dashed lines are 95% confidence band for the survival function) and the proposed model(the blue line).
2(b) Box plots of the cure rates for the simulated data. 1: MLE; 2: Bayesian estimates.
2(c) Estimated survival Curves for the simulated model Weibull(α=1.03,λ=1) by Kaplan-Meier method (solid line is the estimate, dashed lines are 95% confidence band for the survival function) and the fitting models mGEV($\mu = 0, \sigma = 1, \xi$) (the red line) and the Weibull(α,λ=1) (the blue line).
2(d) Estimated survival Curves for the simulated model mGEV($\mu = 0, \sigma = 1, \xi = 0.5$) by Kaplan-Meier method (solid line is the estimate, dashed lines are 95% confidence band for the survival function) and the proposed model Weibull(α,λ=1) (the red line) and the mGEV($\mu = 0, \sigma = 1, \xi$)(the blue line).

Table 1: MLE's & Bayesian Estimates of the model parameters: $\beta_0,\ \beta_1,\ \xi$.

Parameter	MLE	SD	Posterior Mean	95% HPD interval
Intercept	2.002	0.036	2.002	(1.939, 2.071)
Age	0.555	0.036	0.555	(0.476, 0.623)
ξ	0.293	0.015	0.293	(0.260, 0.320)

Table 2: Comparison of Posterior Inference between Weibull distribution and mGEV distribution.

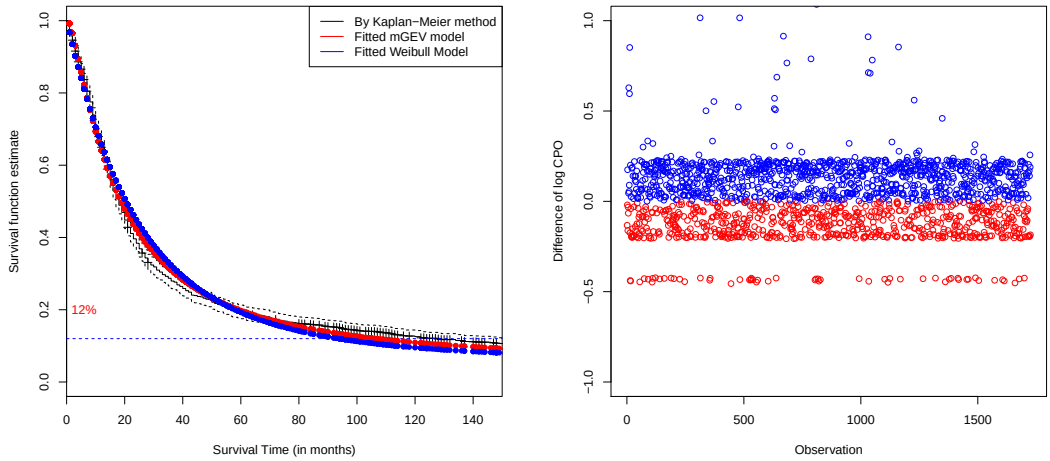
Simulated Model	Fitting Model	Parameter	Posterior Mean	HPD interval
Weibull($\alpha = 1.03, \lambda = 1$)	mGEV($\mu = 0, \sigma = 1, \xi$)	β_0	1.989	(1.936, 2.039)
		β_1	0.628	(0.581, 0.679)
		ξ	0.011	(-0.002, 0.023)
Weibull($\alpha = 1.03, \lambda = 1$)	Weibull($\alpha, \lambda = 1$)	β_0	1.999	(1.845, 2.153)
		β_1	0.645	(0.542, 0.758)
		α	1.021	(0.951, 1.101)
mGEV($\mu = 0, \sigma = 1, \xi = 0.5$)	Weibull($\alpha, \lambda = 1$)	β_0	3.086	(2.835, 3.306)
		β_1	0.665	(0.544, 0.786)
		α	2.735	(2.486, 2.934)
mGEV($\mu = 0, \sigma = 1, \xi = 0.5$)	mGEV($\mu = 0, \sigma = 1, \xi$)	β_0	1.998	(1.945, 2.044)
		β_1	0.588	(0.540,0.633)
		ξ	0.499	(0.496, 0.501)

Table 3: Model Fitting Comparison of mGEV, MGEV, Weibull and Exponentiated Weibull distributions.

		Fitted		
Generated		mGEV	MGEV	Weibull
mGEV	$LPML$	1742.649	1562.961	1301.597
	DIC	-3485.159	-3126.306	-2603.715
MGEV	$LPML$	174.949	695.359	-392.934
	DIC	-348.746	-1390.719	798.289
Weibull	$LPML$	1611.027	1590.493	1595.899
	DIC	-3222.066	-3181.773	-3192.217
Exponentiated Weibull	$LPML$	424.115	476.069	482.26
	DIC	-850.013	-952.832	-964.704

Table 4: Influence Diagnostics for the mGEV model.

Data set	Perturbed Cases	β_0	β_1	ξ	DIC	LPML
a	none	2.002	0.595	0.501	-3494.89	1747.411
b	1	1.996	0.587	0.501	-3473.727	1736.797
c	200	1.995	0.585	0.501	-3477.262	1738.623
d	600	1.998	0.592	0.501	-3479.290	1739.632
e	$1, 200$	1.994	0.592	0.501	-3461.795	1730.878
f	$1, 200, 600$	1.991	0.574	0.501	-3441.219	1720.494



(a) (b)

Fig. 3:
3(a) Estimated survival curves for the GBM data by Kaplan-Meier method(solid line is the estimate, dashed lines are 95% confidence band for the survival function), the fitted Weibull Model(the blue line) and the proposed mGEV model(the red line).
3(b) Plot of difference of the log CPO between mGEV and Weibull Model for the GBM cancer data.

Table 5: Influence Diagnostics for the Weibull model.

Data set	Perturbed Cases	β_0	β_1	α	DIC	LPML
a	none	1.989	0.599	1.019	-3233.66	1616.325
b	1	1.986	0.597	1.021	-3209.603	1604.433
c	200	1.986	0.597	1.019	-3212.707	1605.953
d	600	1.983	0.594	1.019	-3207.631	1603.393
e	$1, 200$	1.985	0.597	1.021	-3205.160	1602.221
f	$1, 200, 600$	1.979	0.598	1.017	-3197.366	1598.312

Table 6: Summary of the GBM Cancer Data.

Survival time(y) (months)	Status(freq)	Age (years)	Gender (freq)	Radiation (freq)	Marital status(freq)
Median 18	Censored 182	Mean 31.4	Male 1053	Had 1453	Married 876
IQR 32	Death 1543	10	Female 672	None 333	Other 897

Table 7: GBM Data: Posterior Estimates of the mGEV Model Parameters with covariates.

Variable	Posterior mean	Posterior SD	95% HPD interval
μ	4.319	0.088	(4.141, 4.482)
σ	1.207	0.048	(1.115, 1.295)
ξ	0.187	0.016	(0.156, 0.216)
Age	0.032	0.002	(0.028, 0.036)
Had Radiation	-0.078	0.067	(-0.226, 0.034)
Marital Status	-0.074	0.051	(-0.179, 0.013)
Gender	0.221	0.051	(0.114, 0.309)

Table 8: GBM Data: Posterior Estimates of the Weibull Model Parameters with covariates.

Variable	Posterior mean	Posterior SD	95% HPD interval
λ	65.350	0.445	(64.705, 66.250)
α	1.106	0.019	(1.070, 1.143)
Age	0.032	0.002	(0.027, 0.035)
Had Radiation	-0.083	0.058	(-0.206, 0.030)
Marital Status	-0.075	0.053	(-0.169, 0.030)
Gender	0.222	0.048	(0.131, 0.319)

Table 9: Model Comparison between Fitted mGEV distribution and Fitted Weibull distribution.

Fitted Model	DIC	LPML
mGEV(μ, σ, ξ)	14177.07	-7088.577
Weibull(α, λ)	14297.7	-7149.243