STATISTICA, anno LXXXI, n. 2, 2021

ESTIMATION OF CUMULATIVE INCIDENCE FUNCTION IN THE PRESENCE OF MIDDLE CENSORING USING IMPROPER GOMPERTZ DISTRIBUTION

Habbiburr Rehman

Department of Statistics, Ramanujan School of Mathematical Sciences, Pondicherry University, Puducherry-605 014, India

Navin Chandra[1]

Department of Statistics, Ramanujan School of Mathematical Sciences, Pondicherry University, Puducherry-605 014, India

1. Introduction

Censoring is the important feature of lifetime modelling, and widely applied when the exact lifetimes of individuals may not be observable due to various reasons. Different types of censoring like right, left and interval censoring are commonly considered in lifetime modelling. Jammalamadaka and Mangalam (2003) introduced a modern concept of censoring scheme known as middle censoring which gets considerable attention in statistical literature. In this censoring scheme the exact lifetimes of some individuals becomes unobservable when it falls into some random censoring intervals. Some amount of work on middle censoring are found in Abuzaid *et al.* (2017) and the references therein.

Middle censoring refers to the situations where the subject is temporarily absent or withdrawn from the study, such as an individual leaves town for a temporary period and returns, if still alive. Moreover, middle censoring also occur when the observations are being taken, is closed for a period, due to an external emergency such as the outbreak of disease, war or a strike. For illustration of middle censoring, let $T_1, T_2, \ldots, T_n$ and $[U_1, V_1], [U_2, V_2], \ldots, [U_n, V_n]$ are the lifetimes and random censoring intervals respectively of the n individuals who are under observation. Under the notion of middle censoring, lifetime T is become observable if $T \notin [U, V]$ with $\Pr(U < V) = 1$, otherwise unobservable.

[1] Corresponding Author. E-mail: nc.stat@gmail.com

In lifetime studies it is often interesting to observe lifetime with associated causes (i.e. competing risks) of failure for individuals or units. Competing risks applications are frequently encountered in medical sciences, demography and engineering sciences. For example, in cancer clinical trial, complete or partial response to treatment may be the primary risk of interest, and death could be considered as the competing risk. Similarly, in liver transplantation an individual can experiences one of the three possible outcomes such as death, transplantation and withdraw from the waiting list. In these illustrations outcomes are the competing risks, because the occurrence of one event either precludes or alter the chance of the occurrence of other events. Analysis of lifetime data in the presence of competing risks utilizes the two main classical approach namely, latent failure time approach and cause specific quantity approach. Latent failure time approach is inappropriate to consider due to the independence assumption of hypothetical failure times in real life problems (Tsiatis, 1975). Cause specific quantities such cause specific hazard function and cumulative incidence function (CIF) (Fine and Gray, 1999) gets considerable attention in modelling of competing risks survival data. Because, Kaplan-Meier estimates of survival function is inappropriate in estimating survival function and its complement (Kalbfleisch and Prentice, 2002).

Recently, Wang (2016) and Ahmadi *et al.* (2017) considered the statistical analysis of middle censored competing risks data with exponential distribution. These literatures mainly focused on latent failure time approach. But modelling of competing risks using cause specific quantities in the presence of middle censoring is still sparse and would be an interesting attempt. We therefore, consider the CIF for modelling of competing risks in the presence of middle censoring. CIF is a capable quantity in assessing the effect of covariates on diseases and probability of expected time to event. CIF gives the probability of failure due to a particular cause in the presence of other competing causes acting on the individuals which is defined by

$$F_j(t;X) = \Pr(T \le t, C = j|X), j = 1,2,\ldots,p, \tag{1}$$

where T is the time to failure, $C \in \{1,2,\ldots,p\}$ be the p possible causes of failure, and X is the $m \times 1$ vector of covariates. Under the assumption that causes are mutually exclusive, the overall distribution function is defined as

$$F(t;X) = \sum_{j=1}^{p} F_j(t;X), \tag{2}$$

each $F_j(t;X)$ is an improper function in the sense that $F_j(\infty;X) < 1$, because, $F(\infty;X) = 1$ is proper. In other words, we can say that the asymptote of CIFs are less than 1, i.e. the distribution function of any cause is improper, because overall cumulative distribution function due to all causes is necessarily proper. This characteristic of CIF can be arise in various applications, for instance, proportion of death in liver transplantation tends to increase for a period of time and then plateau. Those patients who do not experience the death can be considered as cured population. Therefore, in

this situation the inference about the distribution of death can be modelled via the cure model because the cumulative probability of death is less than 1.

Various methods have been proposed for estimating the CIF in statistical literature. Fine and Gray (1999) proposed the semiparametric model for estimating the CIF by extending the Cox proportional hazards model into the competing risks setting. This model is not work for simultaneous modelling of event of interest as well as competing events. An alternative of this model, Jeong and Fine (2006) proposed the direct parameterization of CIF through improper Gompertz distribution (Gompertz, 1825) without covariates. Further, Jeong and Fine (2007) extended the direct parameterization of CIF in to the regression setting. Lee (2019) provided the quantile inference on CIF through improper Gompertz distribution function and compared with Weibull cause specific proportional hazards model. For more detail on direct parameterization of CIF one could refer to Haile *et al.* (2016) and references therein. The aforementioned studies on CIF are mainly based on right censored survival times. Hence, in this article we consider the competing risks analysis through direct parameterization of CIF under middle censoring. The direct parameterization of CIF is more flexible and has a more straightforward interpretation. Therefore, we consider a two parameter improper Gompertz distribution for modelling of CIF in Section 2.

However, Bayesian estimation of CIF based on direct parameterization is not frequently discussed in literature. Therefore, we are interested in estimating the unknown parameters as well as CIF through classical and Bayesian methods of estimation. The novelty of this article assumes the both point and interval estimations of parameters of CIF. In classical approach we consider maximum likelihood estimator (MLE) and midpoint approximation (MPA) methods of point estimation. Asymptotic confidence interval (ACI) of unknown parameters and CIFs are derived based on the asymptotic property of MLE. In addition to Bayesian method we utilized informative and non-informative priors under two loss functions such as squared error loss function (symmetric) and LINEX loss function (asymmetric). As expected, the posterior densities cannot be turns out in any explicit form so we adopt the Markov Chain Mote Carlo (MCMC) simulation algorithm for generating the posterior samples. A 95% credible intervals of unknown parameters and CIFs are obtained using MCMC posterior samples.

The structure of the paper is organised as follows. In Section 2, we formulate the CIF through improper Gompertz distribution. In Section 3, point estimates of unknown parameters and CIF are obtained using MLE and MPA methods, the ACI also obtained in this Section. The Bayes point and credible interval estimates of unknown parameters and CIF are derived in Section 4. In Section 5, we compared the different estimates using four sample sizes and two censoring scheme based on simulation study. In Section 6, we illustrate the estimation procedure using a real data. Finally in section 7, the concluding remarks are given.

2. Model assumption

In this study, for mathematically convenience we consider two mutually exclusive competing risks, $j = 1, 2$, where cause 1 is for event of interest and other competing risks are combined in cause 2. However, the generalization for more than two causes is straight forward. Without loss of generality, we consider the two parameter improper Gompertz distribution (Marshall and Olkin, 2007) for parameterizing the CIF. Therefore, the CIF under the improper Gompertz distribution function is given by

$$F_j(t;\boldsymbol{\Theta}_j) = 1 - \exp\left\{\frac{\alpha_j}{\lambda_j}\left(e^{-\lambda_j t} - 1\right)\right\}, \qquad t > 0, \alpha_j, \lambda_j > 0, \tag{3}$$

where $\boldsymbol{\Theta}_j = (\alpha_j, \lambda_j), j = 1, 2$ is the vector of parameters.

The important feature of survival data with competing risks is the assumption that subject will eventually experience the event of interest or the competing event. In this setting, the probability of never occurring an event of type j equals $\lim_{t\to\infty}(1 - F_j(t;\boldsymbol{\Theta}_j)) = \exp(-\frac{\alpha_j}{\lambda_j})$. Therefore, CIFs from both causes should add up to one as time goes to ∞ such that

$$F_1(\infty;\boldsymbol{\Theta}_1) + F_2(\infty;\boldsymbol{\Theta}_2) = 1. \tag{4}$$

The additivity constraint (4) will be well explain in the situation when the death is one of the competing risks. For example, if we had followed all the patients long enough in a liver transplant clinical trial, we would have noticed that each patient either had a transplant or died without a transplant. This additivity constraint (4) intrinsically holds in nonparametric estimations of CIFs without covariates. However, this additivity constraint has not been explicitly observed in previous studies of CIF based on direct parameterization, mentioned in Section 1. Therefore, in this paper, we consider a parametric model that explicitly takes into account the constraint (4) between CIFs. In practice, constraint (4) can be plugged by solving one parameter in terms of other parameters. Therefore, in this way number of parameters are reduced by one. For example, in model (3), $j = 1, 2$, therefore, α_2 tern out in the following form $\alpha_2 = -\lambda_2 \log[1 - \exp(-\frac{\alpha_1}{\lambda_1})]$.

3. Classical methods of estimation

In this section, MLE and MPA methods of point estimation for unknown parameters and CIF are expressed. Although, interval estimates obtained based on asymptotic property of MLE.

3.1. *Maximum likelihood estimation*

Let us consider the competing risks survival data with middle censoring. Suppose that we have $n \in \mathbb{N}$ individuals with lifetime variate T. It is assumed that T is middle censored by the random censoring interval $[U, V]$ which having a bivariate cumulative

distribution function $G(\cdot,\cdot)$. Also, for the ith individual, lifetime $T_i,(i=1,2,\ldots,n)$ and censoring interval $[U_i,V_i]$ are independent. Therefore, the observed lifetime Y_i is given by

$$Y_i=\begin{cases} T_i; & \text{if} \quad \delta_i=1 \\ [U_i,V_i]; & \text{if} \quad \delta_i=0, \end{cases}$$

where $\delta_i=\mathbf{1}(T_i\notin[U_i,V_i])$ is a censoring indicator. Further, we assume that the left end point of the censored interval U_i and width of the censored interval $W_i=V_i-U_i$ are independently and identically distributed (i.i.d.) random variables and they are independent of T_i. It is also assumed that U_i and W_i are independently follow exponential distributions, i.e. $U_i\sim\exp(\omega_1)$ and $W_i\sim\exp(\omega_2)$. For the ith individual, if $T_i\notin[U_i,V_i]$ then the causes of failure are observable, otherwise it is assumed that they can be observed in later inspection.

Without loss of generality let us consider (y_i,δ_i,j_i) are i.i.d. samples of (Y,δ,C) corresponding to n individuals under study. For the observed data (Y,δ,C) the likelihood function is then given by

$$L(\boldsymbol{\Theta})\propto\prod_{i=1}^{n}\prod_{j=1}^{2}f_j(t_i;\boldsymbol{\Theta}_j)^{\delta_i\Delta_i(j)}\left[F_j(v_i;\boldsymbol{\Theta}_j)-F_j(u_i;\boldsymbol{\Theta}_j)\right]^{(1-\delta_i)\Delta_i(j)}, \tag{5}$$

where $\boldsymbol{\Theta}=(\boldsymbol{\Theta}_1,\boldsymbol{\Theta}_2)$, $f_j(t;\boldsymbol{\Theta}_j)=dF_j(t;\boldsymbol{\Theta}_j)/dt$ and $\Delta_i(j)=\mathbf{1}(C_i=j), j=1,2$ is the indicator function for jth cause. The parameters ω_1 and ω_2 do not depend on $\boldsymbol{\Theta}_j$ and we are not interested to estimate them. Therefore, they do not affect the likelihood of interest. We assume that first $n_1=\sum_{i=1}^{n}\delta_i$ are the uncensored and remaining $n_2=n-n_1$ are censored observations respectively. Let $n_{1j}=\sum_{i=1}^{n}\delta_i\Delta_i(j)$ and $n_{2j}=\sum_{i=1}^{n}(1-\delta_i)\Delta_i(j)$ are the number of the observed events of type j with respect to uncensored and censored individuals respectively with $\sum_{j=1}^{p}\sum_{i=1}^{n}\Delta_i(j)=n$ where $n=n_1+n_2$.

Therefore, the joint likelihood function under model (3) is given by

$$\begin{aligned} L(\boldsymbol{\Theta})=K&\prod_{i=1}^{n_1}\prod_{j=1}^{2}\left[\alpha_j e^{-\lambda_j t_i}\exp\left\{\frac{\alpha_j}{\lambda_j}\left(e^{-\lambda_j t_i}-1\right)\right\}\right]^{\Delta_i(j)} \\ &\times\prod_{i=n_1+1}^{n_1+n_2}\prod_{j=1}^{2}\left[\exp\left\{\frac{\alpha_j}{\lambda_j}\left(e^{-\lambda_j u_i}-1\right)\right\}-\exp\left\{\frac{\alpha_j}{\lambda_j}\left(e^{-\lambda_j v_i}-1\right)\right\}\right]^{\Delta_i(j)}, \end{aligned} \tag{6}$$

where K is the normalizing constant which depends on ω_1 and ω_2. Hence, the corre-

sponding log-likelihood function $\ell = \log L(\boldsymbol{\Theta})$ can be written as

$$\begin{aligned}\ell(\boldsymbol{\Theta}) = \log K + \sum_{j=1}^{2} n_{1j} \log \alpha_j - \sum_{j=1}^{2} \lambda_j \sum_{i=1}^{n_{1j}} t_i + \sum_{j=1}^{2} \sum_{i=1}^{n_{1j}} \frac{\alpha_j}{\lambda_j} \left(e^{-\lambda_j t_i} - 1\right) \\ + \sum_{j=1}^{2} \sum_{i=n_1+1}^{n_1+n_{2j}} \log \left[\exp\left\{ \frac{\alpha_j}{\lambda_j} \left(e^{-\lambda_j u_i} - 1\right) \right\} - \exp\left\{ \frac{\alpha_j}{\lambda_j} \left(e^{-\lambda_j v_i} - 1\right) \right\} \right].\end{aligned} \tag{7}$$

In survival analysis, the heterogeneity among the individuals are explained by covariates or explanatory variables. In order to built the regression model we introduce covariate $\boldsymbol{X}$ with loglink function and write $\alpha_1 = \exp(\beta' \boldsymbol{X}_i)$, in the likelihood $L(\boldsymbol{\Theta})$. Where $\boldsymbol{X}_i = (1, x_{i1}, x_{i2}, \ldots, x_{im})'$ and $\beta' = (\beta_0, \beta_1, \ldots, \beta_m)$ is the vector of regression coefficients. All the parameters of both the competing risks are simultaneously included in likelihood (6) and estimated, although parameter α_2 calculated in terms of other parameters as discussed in Section 2. Then in order to acquire the MLE of unknown parameters, the system of normal equations i.e. $\frac{\partial \ell(\boldsymbol{\Theta})}{\partial \beta} = 0$ and $\frac{\partial \ell(\boldsymbol{\Theta})}{\partial \lambda_j} = 0$ has been derived. Observe that estimates of unknown parameters can not be obtained in closed forms and so we have employed a numerical technique to compute these estimates. We use the `optim` function in `R software` for obtaining MLEs and variance covariance matrix of the unknown parameters. By using the invariance property of MLE we obtained the MLE of the CIF. Suppose the $\hat{\boldsymbol{\Theta}}$ is the MLE of $\boldsymbol{\Theta}$ then the estimator of $F_j(t; \boldsymbol{\Theta}_j, X)$ is given by $\hat{F}_j(t; \hat{\boldsymbol{\Theta}}_j, X)$.

3.2. *Asymptotic confidence interval*

This section deals with deriving interval estimators of unknown parameters and CIFs based on asymptotic distribution of MLE. The MLEs of the unknown parameters are not in closed form, therefore, it is not possible to obtained exact distribution of MLEs. By following the asymptotic property of MLE, the sampling distribution of $(\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta})/\sqrt{\text{var}(\hat{\boldsymbol{\Theta}})}$ can be approximated by a standard normal distribution. The variance of MLE i.e. $\text{var}(\hat{\boldsymbol{\Theta}})$ are the diagonal elements of the asymptotic variance covariance matrix $\Sigma(\hat{\boldsymbol{\Theta}})$ which is the inverse of the Fisher information matrix $\boldsymbol{I}(\boldsymbol{\Theta})$.

The (i, j)th element I_{ij} of Fisher information matrix $\boldsymbol{I}(\boldsymbol{\Theta})$ is given by

$$I_{ij}(\boldsymbol{\Theta}) = -E\left(\frac{\partial^2 \ell(\boldsymbol{\Theta})}{\partial \boldsymbol{\Theta} \partial \boldsymbol{\Theta}'} \right), \qquad i, j = 1, 2, \ldots, m+2.$$

Unfortunately, the exact mathematical expressions for the above expectations are difficult to obtain. Therefore, the observed Fisher information matrix $\boldsymbol{I}_0(\boldsymbol{\Theta})$ can be used to

approximate Fisher information matrix $\boldsymbol{I}(\boldsymbol{\Theta})$, which is obtained by dropping the expectation operator E and it can be written as

$$I_{0_{ij}}(\boldsymbol{\Theta}) = -\left(\frac{\partial^2 \ell(\boldsymbol{\Theta})}{\partial \boldsymbol{\Theta} \partial \boldsymbol{\Theta}'}\right), \quad i, j = 1, 2, \ldots, m+2.$$

Therefore, the variance covariance matrix $\boldsymbol{\Sigma}(\hat{\boldsymbol{\Theta}})$ is obtained by substituting the MLE of $\boldsymbol{\Theta}$ into the $\boldsymbol{I}_0(\boldsymbol{\Theta})$. Thus, for a given confidence level γ, a two-sided $100(1-\gamma)\%$ ACI for $\boldsymbol{\Theta}$ can be constructed as follows

$$\left[\hat{\boldsymbol{\Theta}} - z_{\gamma/2}\sqrt{\text{var}(\hat{\boldsymbol{\Theta}})}, \hat{\boldsymbol{\Theta}} + z_{\gamma/2}\sqrt{\text{var}(\hat{\boldsymbol{\Theta}})}\right],$$

where, $z_{\gamma/2}$ is the upper $\gamma/2$ quantile of the standard normal distribution. Further, we also computed two-sided $100(1-\gamma)\%$ confidence interval for CIF which is given by

$$\left[\hat{F}_j(t;\hat{\boldsymbol{\Theta}}_j, \boldsymbol{X}) - z_{\gamma/2}\sqrt{\text{var}(\hat{F}_j(t;\hat{\boldsymbol{\Theta}}_j, \boldsymbol{X}))}, \hat{F}_j(t;\hat{\boldsymbol{\Theta}}_j, \boldsymbol{X}) + z_{\gamma/2}\sqrt{\text{var}(\hat{F}_j(t;\hat{\boldsymbol{\Theta}}_j, \boldsymbol{X}))}\right],$$

variance of CIF is obtained by using the delta method.

3.3. *Midpoint approximation estimation*

In this subsection, we employ MPA method of estimation (Chen and Lio, 2010), which is useful when the actual observation cannot be made beside an interval can be observed. Teimouri and Gupta (2012) considered the MPA method for parameter estimation of Gompertz-Makeham distribution under progressive type-I interval censoring. Recently, Wang (2016) and Yan *et al.* (2019) utilized the MPA method under middle censoring with competing risks using latent failure time approach.

In MPA method, when the actual observation is masked by the random interval $[U_i, V_i]$, then the T_i is approximated by $T_i^* = (U_i + V_i)/2$ with the assumption that the pseudo actual failure occurred at the middle of the related censoring interval and then the censoring data can be approximately expressed in form of pseudo-complete data as follows.

$$Y_i = \begin{cases} T_i; & \text{if} \quad \delta_i = 1 \\ T_i^*; & \text{if} \quad \delta_i = 0 \end{cases}$$

Under the above notion the pseudo-complete likelihood function is given by

$$L_{\text{MPA}}(\boldsymbol{\Theta}) = K \prod_{i=1}^{n} \prod_{j=1}^{p} f_j(t_i;\boldsymbol{\Theta}_j)^{\delta_i \Delta_i(j)} f_j(t_i^*;\boldsymbol{\Theta}_j)^{(1-\delta_i)\Delta_i(j)}. \tag{8}$$

Further, the log-likelihood function of $L_{\mathrm{MPA}}(\boldsymbol{\Theta})$ under proposed model is given by

$$\begin{aligned}\ell_{MPA}(\boldsymbol{\Theta}) = \log K + \sum_{j=1}^{2} n_{1j} \log \alpha_j - \sum_{j=1}^{2} \lambda_j \sum_{i=1}^{n_{1j}} t_i + \sum_{j=1}^{2} \sum_{i=1}^{n_{1j}} \frac{\alpha_j}{\lambda_j} \left(e^{-\lambda_j t_i} - 1\right) \\ + \sum_{j=1}^{2} n_{2j} \log \alpha_j - \sum_{j=1}^{2} \lambda_j \sum_{i=n_1+1}^{n_1+n_{2j}} t_i^* + \sum_{j=1}^{2} \sum_{i=n_1+1}^{n_1+n_{2j}} \frac{\alpha_j}{\lambda_j} \left(e^{-\lambda_j t_i^*} - 1\right).\end{aligned} \tag{9}$$

As in MLE, the expressions of normal equations under MPA likelihood function are not in explicit form. However, the normal equations of MPA method looks simpler than the MLE. Hence, the estimates of the parameters are turn out to be more easily.

4. Bayesian estimation

In this section, the Bayesian method is used to estimate the unknown parameters and CIF. In the Bayesian method, prior distributions of the model parameters are used to derive their corresponding posterior densities. Usually the selection of prior distribution is based on past experiences, historical data, expert suggestion, simply mathematical convenience or wholly subjective.

Here, we assume that informative priors for β and λ_j which are independently follows standard normal and gamma distributions respectively of the form

$$\pi_1(\beta) \propto e^{-\frac{1}{2}\beta^2}, \quad -\infty < \beta < \infty \tag{10}$$

and

$$\pi_{2j}(\lambda_j) \propto \lambda_j^{a_j - 1} e^{-b_j \lambda_j}, \quad \lambda_j > 0, a_j > 0, b_j > 0, j = 1, 2, \tag{11}$$

where a_j and b_j are the hyper-parameters. The hyper-parameters are assumed to be known or chosen in such a way that reflects the degree of belief about the unknown parameters. On the other side, we assume information on the unknown parameters are not known, i.e. the ignorance of information, the choice of non-informative prior is also considered, a detail is given in simulation section.

From the priors $\pi_1(\beta)$ and $\pi_{2j}(\lambda_j)$ the joint prior distribution of β and λ_j is given by

$$\pi(\boldsymbol{\Theta}) \propto \pi_1(\beta) \times \prod_{j=1}^{2} \pi_{2j}(\lambda_j),$$

$$\pi(\boldsymbol{\Theta}) \propto e^{-\frac{1}{2}\beta^2} \times \prod_{j=1}^{2} \lambda_j^{a_j - 1} e^{-b_j \lambda_j}. \tag{12}$$

Thus the joint posterior density of random variables β and λ_j is obtained through Bayes theorem by using likelihood (6) and joint prior (12) as follows

$$p(\boldsymbol{\Theta} \mid \text{data}) \propto \frac{L(\boldsymbol{\Theta} \mid \text{data})\pi(\boldsymbol{\Theta})}{\int \cdots \int L(\boldsymbol{\Theta} \mid \text{data})\pi(\boldsymbol{\Theta}) d\boldsymbol{\Theta}}. \tag{13}$$

Hence, the conditional posterior of β given λ_1, λ_2 and data is

$$p_1(\beta \mid \lambda_1, \lambda_2, \text{data}) = \frac{p(\boldsymbol{\Theta} \mid \text{data})}{\int_{-\infty}^{\infty} p(\boldsymbol{\Theta} \mid \text{data}) d\beta}, \tag{14}$$

the conditional posterior of λ_1 given β, λ_2 and data is

$$p_{21}(\lambda_1 \mid \beta, \lambda_2, \text{data}) = \frac{p(\boldsymbol{\Theta} \mid \text{data})}{\int_{0}^{\infty} p(\boldsymbol{\Theta} \mid \text{data}) d\lambda_1}, \tag{15}$$

and the conditional posterior of λ_2 given β, λ_1 and data is

$$p_{21}(\lambda_2 \mid \beta, \lambda_1, \text{data}) = \frac{p(\boldsymbol{\Theta} \mid \text{data})}{\int_{0}^{\infty} p(\boldsymbol{\Theta} \mid \text{data}) d\lambda_2}. \tag{16}$$

Clearly, it is not possible to solve the integral involve in the denominator of (13) explicitly, therefore we cannot obtain the conditional posterior densities (14), (15) and (16) in closed form. Thus it is not possible to obtain the Byes estimates of β and λ_j explicitly.

Hence, in this situation we proposed MCMC method to approximate the integrals (Robert and Casella, 2010). MCMC methods are the set of different types of algorithm which are depend on repeated sampling for obtaining the numerical results. Popularly used MCMC algorithms are Gibbs sampling algorithm (Geman and Geman, 1984) and Metropolis-Hastings (M-H) algorithm (Hastings, 1970). Since, conditional posterior densities of random variables β and λ_j are not obtained in closed form so in this situation it is preferable to use M-H algorithm.

We consider two different types of loss functions, namely, squared error (symmetric) and LINEX (asymmetric) loss functions for comprehensive comparison of Bayes estimates. Squared error loss function (SELF) for a parameter $\boldsymbol{\Theta}_j$ is defined as

$$L_1(\boldsymbol{\Theta}_j, \hat{\boldsymbol{\Theta}}_j) = (\boldsymbol{\Theta}_j - \hat{\boldsymbol{\Theta}}_j)^2.$$

Then the Bayes estimate for parameter $\boldsymbol{\Theta}_j$ and CIF $F_j(t; \boldsymbol{\Theta}_j, X)$ under SELF can be obtained as the posterior means and calculated by

$$\hat{\boldsymbol{\Theta}}_j^{\text{self}} = \frac{1}{N-M} \sum_{l=M+1}^{N} \left[\boldsymbol{\Theta}_j\right]_{\boldsymbol{\Theta}_j = \boldsymbol{\Theta}_j^{(l)}},$$

$$\hat{F}_j(t;\boldsymbol{\Theta}_j,\boldsymbol{X})^{\text{self}} = \frac{1}{N-M}\sum_{l=M+1}^{N}\left[F_j(t;\boldsymbol{\Theta}_j,\boldsymbol{X})\right]_{\boldsymbol{\Theta}=\boldsymbol{\Theta}_j^{(l)}},$$

where $\boldsymbol{\Theta}_j^{(l)}, l = 1,2,\ldots,N$ are the are MCMC random samples generated from the posterior distribution of $\boldsymbol{\Theta}_j$ and M is the number of iteration used in burn-in period.

However, we also consider LINEX loss function (LLF) as an asymmetric loss function which is given by

$$L_2(\boldsymbol{\Theta}_j,\hat{\boldsymbol{\Theta}}_j) = e^{\rho(\hat{\boldsymbol{\Theta}}_j-\boldsymbol{\Theta}_j)} - \rho(\hat{\boldsymbol{\Theta}}_j-\boldsymbol{\Theta}_j) - 1,\ \ \rho \neq 0.$$

where ρ is the hyper parameter of the LLF and magnitude of ρ reflect the degree of asymmetry. For $\rho > 0$ the LLF is quite asymmetric about 0 with overestimation being more serious than underestimation. The vice versa is true with $\rho < 0$. If ρ is close to zero then estimates under LLF are approximately equal to estimates obtained under SELF. Hence, LLF is more applicable in lifetime modelling, for instance, over estimation of survival function and failure rate function is usually much more serious than under estimation.

Under LLF the Bayes estimates of parameter $\boldsymbol{\Theta}_j$ and CIF $F_j(t;\boldsymbol{\Theta}_j,\boldsymbol{X})$ can be obtained as follows

$$\hat{\boldsymbol{\Theta}}_j^{llf} = -\frac{1}{\rho}\log\left(\frac{1}{N-M}\sum_{l=M+1}^{N-M} e^{-\rho[\boldsymbol{\Theta}]_{\boldsymbol{\Theta}_j=\boldsymbol{\Theta}_j^{(l)}}}\right),$$

$$\hat{F}_j(t;\boldsymbol{\Theta}_j,\boldsymbol{X})^{llf} = -\frac{1}{\rho}\log\left(\frac{1}{N-M}\sum_{l=M+1}^{N} e^{-\rho\left[F_j(t;\boldsymbol{\Theta}_j,\boldsymbol{X})\right]_{\boldsymbol{\Theta}_j=\boldsymbol{\Theta}_j^{(l)}}}\right).$$

In Bayesian framework for a γ level of significance, the $(1-\gamma)$ interval estimate of a parameter $\boldsymbol{\Theta}_j$ is a credible interval based on given data, that covers the parameter with $(1-\gamma)$ level of confidence. The $100(1-\gamma)\%$ Bayes credible interval (BCI), $\left[\boldsymbol{\Theta}_j^L, \boldsymbol{\Theta}_j^U\right]$, for $\boldsymbol{\Theta}_j$ is obtained by setting $\boldsymbol{\Theta}_j^L$ equal to the $\gamma/2\%$ quantile and $\boldsymbol{\Theta}_j^U$ equal to $(1-\gamma/2)\%$ quantile of the posterior sample of $\boldsymbol{\Theta}_j$. Similarly, same procedure is also adopted for obtaining BCI for $F_j(t;\boldsymbol{\Theta}_j,\boldsymbol{X})$.

5. Simulation study

In this section, numerical investigation of the proposed estimators is carried out through Monte Carlo simulation. We performed numerical comparisons of MLE, MPA and Bayes estimators with the different choices of simulated sample sizes n=25, 50, 100 and 200 at prefixed censoring scheme(CS). Each sample was replicated 500 times. The performance of all point estimators are compared numerically in terms of their average estimate (AE) and mean square error (MSE) values. Also, average length (AVL) and coverage probability (CP) are computed for interval estimates. The results from the simulation study are presented in Table 1 and Table 2.

In this scenario, we generated the survival times from the proposed Gompertz model for both the causes as given in (3) using inverse transformation method. Regression model is developed through one covariate x which is generated from the standard normal distribution. We consider $\lambda_1 = 1.5$ for cause 1 and $\lambda_2 = 1$ for cause 2. For $\alpha_1 = \exp(\beta_0 + \beta_1 x)$, true parameters are chosen as $\beta_0 = 0.3$ and $\beta_1 = 0.4$, therefore, subject to additivity constraint (4) for both CIF, we have $\alpha_2 = 0.5981$ at covariate $x = -0.3$. We also generated the censoring intervals for two CSs with following combinations of $(\omega_1, \omega_2) = (1, 4)$, and$(1, 1.3)$ which are imposing the average censoring proportion approximately 10% and 20%. Based on different sample sizes and CSs, the ML, MPA and Bayes estimates of parameters and CIF are calculated. The estimates of CIF are obtained at time $t = 0.5$ and covariate $x = -0.3$ for both the causes and denoted as F_1 and F_2 with true values $F_1 = 0.344$ and $F_2 = 0.210$. MLE and MPA point estimates of four unknown parameters $(\beta_0, \beta_1, \lambda_1, \lambda_2)$ and CIF are computed using model based log-likelihood (7) and (9) respectively. The ACI were computed from the Fisher information matrix evaluated at the MLEs.

In order to compute the Bayes estimates we use informative and non-informative priors. In case of informative prior (INP), the hyper parameters of gamma priors are calculated using the likelihood estimates of λ_1 and λ_2 based on 1000 data set of sample size 25. Now, we compute the mean and variance of the MLE of λ_1 and λ_2 and compare with the mean and variance of gamma priors. Subsequently, we get the hyper parameters values as $a_1 = 8.18, b_1 = 5.14$ and $a_2 = 6.79, b_2 = 6.22$. For regression parameters β_0 and β_1 we assumed $N(0, 1)$ as informative priors. Further in case of non-informative prior (NIP) we assume that $a_1 = b_1 = a_2 = b_2 = 0.0001$ and where β_0 and β_1 are said to be follow $N(0, 1000)$. The hyper parameter of LLF is fixed at $\rho = \pm 1.5$ known as llf-1 and llf-2. We also computed the BCI under INP and NIP and denoted as BCI-1 and BCI-2.

Further, as we mentioned in Section 4 that conditional posterior densities of unknown parameters are not turn out in any distributional form, so we adopt the MCMC procedure for generating the random samples. For this purpose we used the `BUGS software` via `R2OpenBUGS` package in `R software` (Lunn *et al.*, 2012).

We generate 10000 Markov chains for each parameter and the first 4000 samples are removed for reducing the effect of initial values. Furthermore, for minimizing the effect of the autocorrelation every second equally spaced outcome is considered i.e. thin=2. By the visualization of the convergence diagnostics plots it is realized that chains are converging nicely. Therefore, the last 6000 MCMC samples are used to obtained the Bayes estimates of $\beta_0, \beta_1, \lambda_1,$ andλ_2.

From Table 1 it is observed that as the sample size n increases, MSEs decreases for MLE, MPA and Bayes estimates, for both CSs, which verifies the consistency property of all the estimators. The Bayes estimators based on INP under both the loss functions are dominating over other estimates in terms of their MSEs. We can also see that for small sample sizes n, MPA estimators also performed well in compared to MLE. For $n = 200$ the magnitudes of MSEs of all estimators are negligible. For fixed sample size as censoring proportion is increases, MSEs increases except for some values.

TABLE 1
AVE and MSE values of the estimates are reported.

			CS1						CS2					
n	Estimates		β_0	β_1	λ_1	λ_2	F_1	F_2	β_0	β_1	λ_1	λ_2	F_1	F_2
	True		0.300	0.400	1.500	1.000	0.344	0.210	0.300	0.400	1.500	1.000	0.344	0.210
25	MLE	AVE	0.309	0.443	1.564	1.104	0.349	0.214	0.299	0.412	1.589	1.150	0.349	0.220
		MSE	0.141	0.099	0.264	0.223	0.891	0.548	0.145	0.099	0.331	0.361	0.999	0.716
	MPA	AVE	0.297	0.441	1.545	1.095	0.347	0.213	0.212	0.406	1.449	1.076	0.334	0.211
		MSE	0.139	0.098	0.253	0.211	0.873	0.537	0.142	0.098	0.246	0.273	0.911	0.656
	INP	AVE	0.226	0.422	1.529	1.038	0.335	0.223	0.218	0.382	1.537	1.051	0.337	0.225
	self	MSE	0.085	0.076	0.064	0.041	0.585	0.404	0.087	0.078	0.070	0.048	0.643	0.482
	INP	AVE	0.150	0.370	1.442	0.986	0.330	0.220	0.138	0.320	1.448	0.996	0.332	0.221
	llf-1	MSE	0.107	0.073	0.054	0.032	0.585	0.385	0.112	0.082	0.057	0.037	0.638	0.456
	INP	AVE	0.297	0.476	1.634	1.100	0.340	0.227	0.295	0.445	1.645	1.117	0.343	0.229
	llf-2	MSE	0.075	0.086	0.100	0.060	0.590	0.426	0.075	0.083	0.113	0.072	0.654	0.511
	NIP	AVE	0.189	0.472	1.570	1.076	0.326	0.219	0.182	0.431	1.594	1.130	0.329	0.224
	self	MSE	0.165	0.111	0.264	0.209	0.840	0.545	0.170	0.109	0.333	0.350	0.929	0.705
	NIP	AVE	0.076	0.413	1.399	0.959	0.321	0.215	0.061	0.359	1.415	0.987	0.322	0.219
	llf-1	MSE	0.222	0.097	0.171	0.118	0.847	0.520	0.232	0.104	0.196	0.152	0.930	0.666
	NIP	AVE	0.293	0.535	1.836	1.266	0.332	0.223	0.293	0.505	1.895	1.415	0.336	0.229
	llf-2	MSE	0.140	0.134	0.652	0.610	0.839	0.574	0.143	0.127	1.050	2.295	0.935	0.751
50	MLE	AVE	0.315	0.417	1.566	1.056	0.348	0.214	0.289	0.422	1.549	1.050	0.341	0.217
		MSE	0.070	0.049	0.144	0.104	0.466	0.270	0.067	0.040	0.128	0.089	0.428	0.277
	MPA	AVE	0.302	0.416	1.546	1.050	0.345	0.213	0.201	0.417	1.416	0.993	0.326	0.210
		MSE	0.068	0.048	0.134	0.100	0.456	0.265	0.074	0.039	0.103	0.071	0.428	0.249
	INP	AVE	0.264	0.394	1.546	1.037	0.340	0.218	0.240	0.402	1.536	1.035	0.333	0.222
	self	MSE	0.050	0.043	0.059	0.038	0.355	0.222	0.050	0.035	0.058	0.039	0.337	0.233
	INP	AVE	0.223	0.360	1.489	1.003	0.337	0.216	0.198	0.376	1.477	1.001	0.331	0.220
	llf-1	MSE	0.056	0.044	0.042	0.028	0.354	0.215	0.059	0.035	0.049	0.033	0.339	0.225
	INP	AVE	0.304	0.428	1.611	1.074	0.343	0.220	0.282	0.429	1.602	1.073	0.336	0.224
	llf-2	MSE	0.047	0.045	0.089	0.054	0.357	0.229	0.046	0.037	0.078	0.049	0.336	0.241
	NIP	AVE	0.258	0.414	1.568	1.042	0.337	0.216	0.231	0.419	1.552	1.038	0.331	0.219
	self	MSE	0.074	0.050	0.144	0.101	0.448	0.271	0.074	0.040	0.128	0.086	0.422	0.278
	NIP	AVE	0.207	0.378	1.485	0.991	0.334	0.214	0.178	0.392	1.466	0.987	0.328	0.217
	llf-1	MSE	0.083	0.050	0.111	0.078	0.448	0.263	0.087	0.039	0.100	0.067	0.426	0.269
	NIP	AVE	0.307	0.450	1.668	1.103	0.341	0.218	0.283	0.447	1.657	1.099	0.334	0.221
	llf-2	MSE	0.070	0.054	0.212	0.145	0.450	0.279	0.067	0.043	0.194	0.123	0.420	0.288
100	MLE	AVE	0.294	0.406	1.514	1.023	0.343	0.213	0.313	0.404	1.520	1.031	0.348	0.210
		MSE	0.033	0.022	0.053	0.029	0.193	0.124	0.034	0.022	0.057	0.039	0.222	0.144
	MPA	AVE	0.282	0.405	1.498	1.017	0.341	0.212	0.229	0.399	1.396	0.980	0.333	0.203
		MSE	0.032	0.022	0.050	0.028	0.190	0.123	0.037	0.022	0.054	0.032	0.213	0.135
	INP	AVE	0.264	0.395	1.515	1.024	0.337	0.216	0.287	0.391	1.520	1.031	0.344	0.213
	self	MSE	0.029	0.021	0.036	0.020	0.172	0.117	0.028	0.021	0.038	0.027	0.188	0.127
	INP	AVE	0.242	0.379	1.482	1.006	0.336	0.216	0.265	0.376	1.486	1.012	0.343	0.212
	llf-1	MSE	0.032	0.022	0.033	0.018	0.173	0.115	0.029	0.021	0.035	0.024	0.187	0.125
	INP	AVE	0.286	0.412	1.550	1.043	0.339	0.217	0.309	0.405	1.556	1.053	0.345	0.214
	llf-2	MSE	0.028	0.022	0.042	0.023	0.172	0.119	0.027	0.021	0.045	0.031	0.190	0.129
	NIP	AVE	0.262	0.407	1.514	1.016	0.336	0.215	0.284	0.399	1.521	1.026	0.343	0.211
	self	MSE	0.035	0.023	0.053	0.028	0.194	0.126	0.034	0.022	0.057	0.038	0.214	0.143

Continued

			CS1						CS2					
n	Estimates		β_0	β_1	λ_1	λ_2	F_1	F_2	β_0	β_1	λ_1	λ_2	F_1	F_2
	NIP	AVE	0.237	0.390	1.474	0.995	0.335	0.214	0.259	0.385	1.480	1.002	0.341	0.210
	llf-1	MSE	0.038	0.022	0.048	0.026	0.195	0.125	0.036	0.022	0.051	0.034	0.213	0.141
	NIP	AVE	0.286	0.423	1.557	1.039	0.338	0.215	0.308	0.414	1.566	1.052	0.344	0.212
	llf-2	MSE	0.033	0.023	0.062	0.032	0.193	0.128	0.034	0.023	0.069	0.045	0.215	0.145
200	MLE	AVE	0.304	0.401	1.518	1.012	0.345	0.211	0.296	0.406	1.515	1.012	0.342	0.213
		MSE	0.016	0.009	0.029	0.014	0.097	0.061	0.017	0.009	0.030	0.015	0.107	0.068
	MPA	AVE	0.292	0.400	1.501	1.007	0.343	0.211	0.210	0.401	1.390	0.963	0.327	0.205
		MSE	0.015	0.009	0.028	0.014	0.095	0.060	0.024	0.009	0.034	0.013	0.126	0.062
	INP	AVE	0.289	0.396	1.519	1.014	0.342	0.213	0.281	0.402	1.516	1.014	0.340	0.214
	self	MSE	0.014	0.009	0.024	0.012	0.090	0.059	0.015	0.008	0.025	0.013	0.100	0.065
	INP	AVE	0.278	0.389	1.501	1.004	0.341	0.212	0.269	0.396	1.498	1.004	0.339	0.214
	llf-1	MSE	0.015	0.009	0.023	0.011	0.090	0.058	0.016	0.008	0.023	0.012	0.100	0.065
	INP	AVE	0.300	0.403	1.537	1.024	0.343	0.213	0.292	0.409	1.535	1.025	0.340	0.215
	llf-2	MSE	0.014	0.009	0.027	0.013	0.090	0.059	0.015	0.008	0.027	0.014	0.099	0.066
	NIP	AVE	0.288	0.400	1.518	1.007	0.341	0.212	0.279	0.407	1.515	1.008	0.339	0.213
	self	MSE	0.016	0.009	0.029	0.014	0.096	0.062	0.017	0.009	0.030	0.015	0.108	0.068
	NIP	AVE	0.276	0.394	1.499	0.997	0.341	0.211	0.267	0.400	1.495	0.996	0.338	0.213
	llf-1	MSE	0.016	0.009	0.028	0.013	0.096	0.061	0.018	0.008	0.028	0.014	0.108	0.068
	NIP	AVE	0.299	0.407	1.539	1.018	0.342	0.212	0.292	0.414	1.536	1.019	0.340	0.214
	llf-2	MSE	0.016	0.010	0.032	0.015	0.096	0.062	0.017	0.009	0.033	0.016	0.107	0.069

Form Table 2 it is clear that as the sample size n increases, average length of ACI, BCI-1 and BCI-2 decreases. Similarly, in most of the cases, for fixed sample size n as the censoring proportion increases, AVL of ACI, BCI-1 and BCI-2 increases except for β_1. For all the interval estimates, we observe very stable CPs around 95%. The AVL of BCI estimates based on NIP are slightly larger than the other interval estimates.

TABLE 2
AVL and CP of ACI and Bayes credible intervals.

			CS1						CS2					
n	Estimates		β_0	β_1	λ_1	λ_2	F_1	F_2	β_0	β_1	λ_1	λ_2	F_1	F_2
25	ACI	AVL	1.434	1.089	1.935	1.609	0.360	0.275	1.495	1.196	1.980	1.741	0.388	0.299
		CP	0.946	0.926	0.950	0.936	0.930	0.910	0.960	0.960	0.956	0.934	0.946	0.910
	BCI-1	AVL	1.220	1.027	1.365	1.049	0.309	0.260	1.259	1.124	1.379	1.076	0.331	0.277
		CP	0.968	0.940	0.992	0.992	0.952	0.952	0.984	0.962	0.986	0.988	0.972	0.970
	BCI-2	AVL	1.472	1.101	1.920	1.521	0.336	0.280	1.525	1.213	1.961	1.656	0.360	0.303
		CP	0.942	0.926	0.936	0.922	0.930	0.934	0.946	0.944	0.942	0.928	0.952	0.952
50	ACI	AVL	0.998	0.852	1.330	1.038	0.258	0.191	1.025	0.751	1.355	1.049	0.257	0.191
		CP	0.938	0.952	0.936	0.932	0.934	0.926	0.960	0.946	0.950	0.932	0.960	0.910
	BCI-1	AVL	0.909	0.830	1.091	0.832	0.237	0.194	0.927	0.732	1.108	0.838	0.235	0.195
		CP	0.958	0.948	0.964	0.968	0.958	0.972	0.976	0.940	0.982	0.972	0.962	0.956
	BCI-2	AVL	1.005	0.852	1.315	0.999	0.249	0.203	1.031	0.749	1.342	1.011	0.247	0.205
		CP	0.946	0.946	0.938	0.918	0.938	0.960	0.956	0.938	0.958	0.942	0.958	0.950
100	ACI	AVL	0.702	0.582	0.917	0.677	0.170	0.124	0.708	0.550	0.924	0.717	0.181	0.130
		CP	0.954	0.946	0.944	0.946	0.950	0.916	0.952	0.942	0.954	0.938	0.948	0.908
	BCI-1	AVL	0.666	0.574	0.827	0.607	0.161	0.131	0.669	0.541	0.835	0.637	0.172	0.138
		CP	0.958	0.952	0.968	0.964	0.954	0.944	0.962	0.936	0.972	0.954	0.956	0.956
	BCI-2	AVL	0.701	0.579	0.905	0.656	0.166	0.134	0.707	0.545	0.917	0.696	0.177	0.142
		CP	0.950	0.948	0.944	0.946	0.948	0.940	0.942	0.936	0.950	0.934	0.952	0.948
200	ACI	AVL	0.493	0.369	0.644	0.479	0.123	0.089	0.503	0.368	0.655	0.492	0.126	0.091
		CP	0.948	0.946	0.940	0.960	0.940	0.930	0.940	0.948	0.940	0.960	0.932	0.938
	BCI-1	AVL	0.477	0.369	0.606	0.448	0.120	0.096	0.485	0.368	0.612	0.460	0.122	0.099
		CP	0.940	0.948	0.946	0.962	0.940	0.946	0.940	0.956	0.944	0.960	0.946	0.964
	BCI-2	AVL	0.490	0.369	0.634	0.466	0.121	0.097	0.499	0.369	0.643	0.478	0.124	0.100
		CP	0.932	0.944	0.938	0.948	0.940	0.946	0.938	0.952	0.930	0.954	0.934	0.954

6. ILLUSTRATION WITH REAL DATA

In this section, we used a real data to illustrate the proposed estimation methods discussed in the previous Sections. The data set is extracted from 'survival' package in `R software` which is available by the name of 'transplant'. This data represents the survival experience of patients who were registered for liver transplant waiting list at Mayo Clinic Rochester between 1990 and 1999. In this data frame the observations of 815 patients are included, 636(78%) received transplant, 66(8%) died while waiting, 37(5%) withdraw from the list, and 76(9%) were still waiting as of February 2002. A detail study of this data set using competing risks analysis is available in Kim *et al.* (2006).

Patients who have registered for a transplant may experience the transplant, death, and withdrawal; otherwise, they would have to wait until their next appointment. Therefore, competing risks model become more reasonable for analysing the competing variables, to compare the risk of death across the various risk factor (such as age, blood groups and sex) associated to each patients. In this illustration, we consider the transplant as an event of interest and death as competing event, the patients who withdrawal from the list are combined with censored observations. Three covariates are associated

with each patient, which are age, sex and blood groups. In the present study, we considered blood group as a covariate. Blood groups included types O, A, B, and AB. We divided the covariate blood group in two groups as group 1 (A, B and AB) and group 2 (O) and coded as 0 and 1 respectively.

Lifetimes of patients are given in days that have been rescaled in years by dividing it 365. To check the goodness of fit of Gompertz distribution for complete data set, we use K-S distance and graphical fitting. The K-S distance and plots of density, cdf, q-q and p-p are obtained at MLEs which were computed through 'fitdistrplus' package in `R software`. The K-S distance between the empirical distribution function and the fitted Gompertz distribution function is 0.0294 and the corresponding p-value is 0.4725. Since the p-value is quite high, therefore, it can be seen that Gompertz distribution can be used for this data set. Figure 1 also indicates that it is appropriate to select the Gompertz distribution.

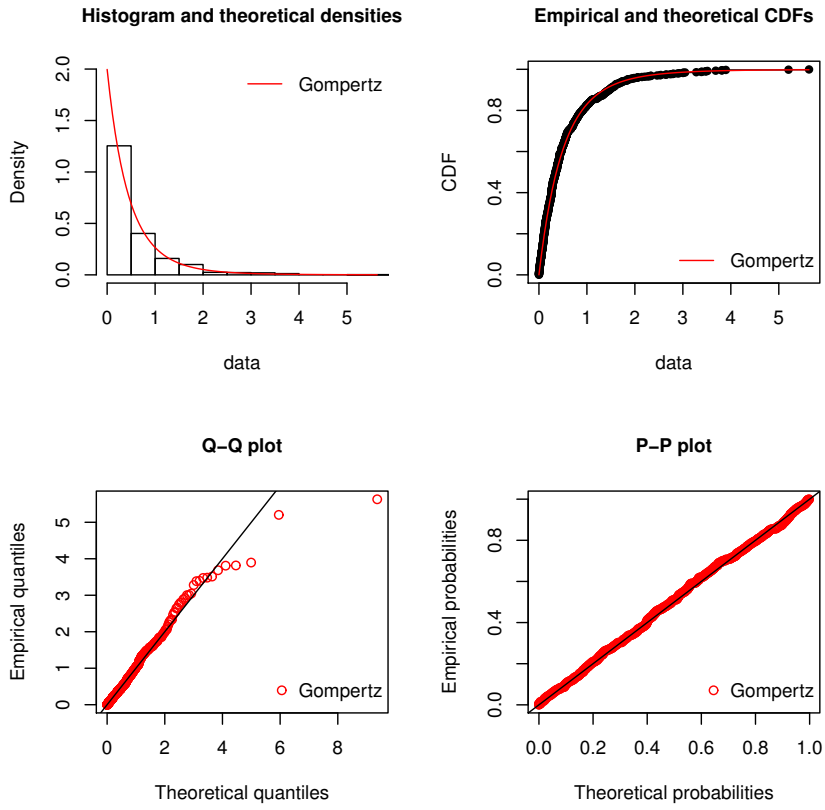

Figure 1 – Graphical fitting of the Gompertz distribution.

Next, we created an artificial data set by middle censoring, whose left end point as the observed time itself and the right end point is calculated by the sum of left end point and the width of interval which is generated from an exponential distribution with mean 10. Then, all the censored individuals are considered as the middle censored observations and the outcome variable for middle censored observation, randomly assigned from death and transplant.

We apply the proposed estimation methods on artificial data set for estimating the unknown parameters and CIFs. The Bayes estimates are obtained under SELF and LLF loss functions for non-informative prior. The point and interval estimates of $\beta_0, \beta_1, \lambda_1$, and λ_2 are presented in Table 3.

It is observed that the effect of covariate blood group is statistically significant in the sense that interval estimates (ACI and BCI) of regression coefficient of blood group does not contain zero. Although, we test the significance of blood group effect using likelihood ratio test procedure. For this we set the hypothesis of interest as $H_0 : \beta_1 = 0$ against $H_1 : \beta_1 \neq 0$. The calculated test statistic is 23.589 with $p-value$ much less than 0.001. This indicate that the covariate effect is highly significant.

However, the estimates of CIFs of transplant and death are computed for blood group 0 and 1 which are presented in Figure 2. The estimates of CIFs based on all methods behave equally. Figure 2 shows that CIF for transplant have higher value as compared to death. Also, those patients have blood group O have more risk of dying in waiting list compared to blood group A, B and AB. Note, that cumulative probability of receiving the transplant rapidly increases up to 1.5 years, thereafter CIF tends to plateau. This indicates that at the beginning of the trial, the success of transplant leads to increase in its applications, while number of available donors are remain stationary or decrease as time goes.

TABLE 3
Point and interval estimates of model parameters.

Estimator		β_0	β_1	λ_1	λ_2
MLE		0.825	-0.373	0.987	0.645
MPA		0.503	-0.399	0.61	0.301
NIP	self	0.824	-0.374	0.987	0.646
	llf-1	0.822	-0.378	0.985	0.642
	llf-2	0.827	-0.369	0.989	0.650
ACI	LL	0.712	-0.524	0.883	0.506
	UU	0.939	-0.221	1.090	0.784
BCI	LL	0.711	-0.524	0.887	0.516
	UU	0.938	-0.225	1.093	0.792

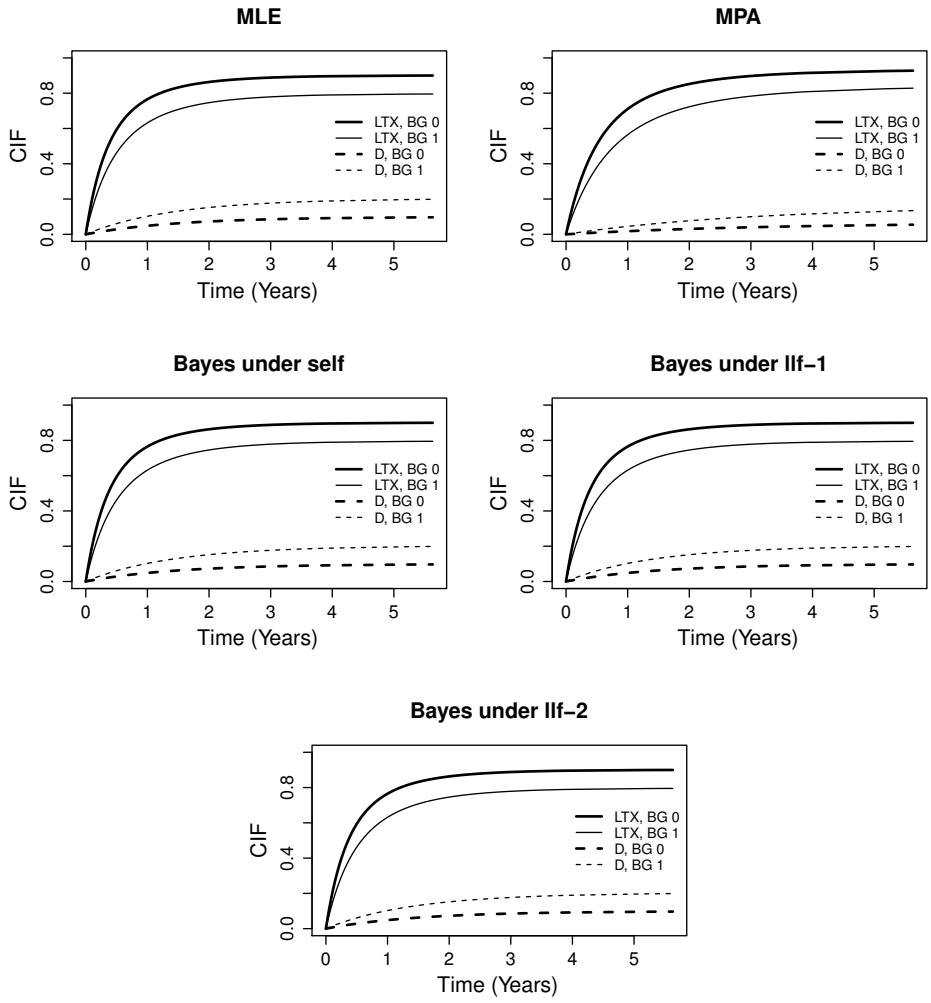

Figure 2 – Estimated CIFs of transplant (LTX) and death (D) for both the blood groups (BG).

7. Conclusion

In this article, we considered an improper Gompertz model for analysing competing risks data in the presence of middle censoring. The improper Gompertz distribution is used to model the CIF for event of interest as well as competing event subject to additivity constraint for both the events. In classical set up the MLE and MPA methods are used for obtaining the point estimates of unknown parameters and CIFs. Also, interval estimates are derived based on asymptotic property of MLE. The Bayes estimates have also been considered based on informative and non-informative priors under SELF and LLF. In simulation study it is observe that Bayes estimates based on informative priors work well among other estimates. Simulation study showed that the proposed method is efficient. In real data analysis, the impropriety of the Gompertz model well captured by CIFs for transplant and death because of the additivity constraint. Hence, ignoring the additivity constraint may lead to over estimates of cumulative probability due to all competing causes.

However, various censoring schemes such as current status censoring and double censoring have substantial statistical literature. Therefore, competing risks modelling in these censoring schemes via direct parameterization of CIF seem to be an interesting attempt. Quantile inferences are very common in survival analysis, and quantile inferences of competing risks with the middle censoring scheme may be developed elsewhere in the future.

Acknowledgements

The authors are thankful to the journal's Editor and referees for providing constructive comments to improve the content of the article.

References

A. Abuzaid, M. A. El-Qumsan, A. El-Habil (2017). *On the robustness of right and middle censoring schemes in parametric survival models*. Communications in Statistics-Simulation and Computation, 46, no. 3, pp. 1771–1780.

K. Ahmadi, M. Rezaei, F. Yousefzadeh (2017). *Statistical analysis of middle censored competing risks data with exponential distribution*. Journal of Statistical Computation and Simulation, 87, no. 16, pp. 3082–3110.

D. Chen, Y. Lio (2010). *Parameter estimations for generalized exponential distribution under progressive type-I interval censoring*. Computational Statistics & Data Analysis, 54, no. 6, pp. 1581–1591.

J. P. Fine, R. J. Gray (1999). *A proportional hazards model for the subdistribution of a competing risk*. Journal of the American statistical association, 94, no. 446, pp. 496–509.

S. GEMAN, D. GEMAN (1984). *Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 6, no. 6, pp. 721–741.

B. GOMPERTZ (1825). *On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies.* Philosophical Transactions of the Royal Society of London, , no. 115, pp. 513–583.

S. HAILE, J.-H. JEONG, X. CHEN, Y. CHENG (2016). *A 3-parameter Gompertz distribution for survival data with competing risks, with an application to breast cancer data*. Journal of Applied Statistics, 43, no. 12, pp. 2239–2253.

W. K. HASTINGS (1970). *Monte carlo sampling methods using markov chains and their applications*. Biometrika, 57, no. 1, pp. 97–109.

S. R. JAMMALAMADAKA, V. MANGALAM (2003). *Nonparametric estimation for middle-censored data*. Journal of Nonparametric Statistics, 15, no. 2, pp. 253–265.

J.-H. JEONG, J. FINE (2006). *Direct parametric inference for the cumulative incidence function*. Journal of the Royal Statistical Society: Series C (Applied Statistics), 55, no. 2, pp. 187–200.

J.-H. JEONG, J. P. FINE (2007). *Parametric regression on cumulative incidence function*. Biostatistics, 8, no. 2, pp. 184–196.

J. D. KALBFLEISCH, R. L. PRENTICE (2002). *The Statistical Analysis of Failure Time Data*, vol. 360. John Wiley & Sons, New Jersey.

W. R. KIM, T. M. THERNEAU, J. T. BENSON, W. K. KREMERS, C. B. ROSEN, G. J. GORES, E. R. DICKSON (2006). *Deaths on the liver transplant waiting list: an analysis of competing risks*. Hepatology, 43, no. 2, pp. 345–351.

M. LEE (2019). *Parametric inference for quantile event times with adjustment for covariates on competing risks data*. Journal of Applied Statistics, 46, no. 12, pp. 2128–2144.

D. LUNN, C. JACKSON, N. BEST, D. SPIEGELHALTER, A. THOMAS (2012). *The BUGS Book: A Practical Introduction to Bayesian Analysis*. Chapman and Hall/CRC, Boca Raton.

A. W. MARSHALL, I. OLKIN (2007). *Life distributions*, vol. 13. Springer-Verlag, New York.

C. P. ROBERT, G. CASELLA (2010). *Introducing Monte Carlo Methods with R*, vol. 18. Springer-Verlag, New York.

M. TEIMOURI, A. K. GUPTA (2012). *Estimation methods for the Gompertz–Makeham distribution under progressively type-I interval censoring scheme*. National Academy Science Letters, 35, no. 3, pp. 227–235.

A. TSIATIS (1975). *A nonidentifiability aspect of the problem of competing risks*. Proceedings of the National Academy of Sciences, 72, no. 1, pp. 20–22.

L. WANG (2016). *Estimation for exponential distribution based on competing risk middle censored data*. Communications in Statistics-Theory and Methods, 45, no. 8, pp. 2378–2391.

W. YAN, S. YIMIN, W. MIN (2019). *Statistical inference for dependence competing risks model under middle censoring*. Journal of Systems Engineering and Electronics, 30, no. 1, pp. 209–222.

SUMMARY

In this paper we deal with the modelling of cumulative incidence function using improper Gompertz distribution based on middle censored competing risks survival data. Together with the unknown parameters, cumulative incidence function also estimated. In classical set up, we derive the point estimates using maximum likelihood estimator and midpoint approximation methods. The asymptotic confidence interval are obtained based on asymptotic normality properties of maximum likelihood estimator. We also derive the Bayes estimates with associated credible intervals based on informative and non-informative types of priors under two loss functions such as squared error and LINEX loss functions. A simulation study is conducted for comprehensive comparison between various estimators proposed in this paper. A real life data set is also used for illustration.

Keywords: Cumulative incidence function; Improper Gompertz distribution; Middle censoring; Maximum likelihood estimator; Bayes estimators; MCMC method.