Original Research Paper

Volume - 11 | Issue - 03 | March - 2021 | PRINT ISSN No. 2249 - 555X | DOI : 10.36106/ijar

Statistics

PENALIZED LOG-LIKELIHOOD ESTIMATION FOR COX-FRAILTY MODEL WITH NONINFORMATIVE BIVARIATE CURRENT STATUS DATA

Alhassan Faisal Department of Statistics, Faculty of Mathematical Sciences, C. K. Tedam University of Technology and Applied Sciences, P. O. Box 24, Navrongo, Ghana.

ABSTRACT A Penalized Maximum Likelihood Estimation (PMLE) procedure is proposed for Cox proportional hazards frailty model with noninformative bivariate current status data. An integrated splines (I-splines) was used to approximate the two unknown baseline cumulative hazard functions of the failure times. The one-parameter gamma frailty distribution was used to model the correlation between the two failure times. An easy to implement computational algorithm is proposed to estimate the regression and splines parameters. Bayesian technique as proposed by Wahba (1983) was employed for the variance estimation. The statistical properties of the estimated parameters were studied through extensive simulation and it was observed that the PMLEs were consistent, asymptotically normal and efficient. In addition, the estimators were robust to the choice of knots, censoring rates and type of frailty distribution used. The proposed methodology is further demonstrated through the analysis of the tumorigenicity experiment data by Lindsey and Ryan (1994).

KEYWORDS : Bivariate Current Status data, Splines, Proportional Hazards Model, Penalized Maximum Likelihood Estimation

1. INTRODUCTION

Bivariate current status data arise when each subject in a sample is observed only once to ascertain the presence or otherwise of two events under study. Because the two events share the same environment, they are expected to be similar in terms of the hazard of the events. Therefore, the assumption of the failure times being independent of each other may not hold in a bivariate data setting. Ignoring this correlation in the modeling process can lead to bias and misleading conclusions. One example of bivariate current status data is the tumorigenicity studies experiment (Lindsey & Ryan , 1994), which investigated time until the onset of tumors (lung and bladder) in mice. The study lasted 33 months. Female mice were exposed to two doses of a suspected carcinogen; the control or high dose groups. The mice were pathologically examined at either sacrifice or natural death times to ascertain the presence or absence of these two tumors. Thus, each mouse has only one observation time, which is at death, and the only knowledge available to the investigator about the onset times of these two tumors is either their presence or otherwise prior to death. This data results in a bivariate current status data. When the examination time is assumed to be independent of the two failure times, then a noninformative bivariate current status data are obtained. The most commonly used approach by researchers, which is herein adopted, to account for the relationship between survival times in a multivariate current status data setting is the Frailty model approach.

Many authors have proposed different estimation approaches for noninformative multivariate current status data. For example, Chen (2007) proposed an efficient Expectation Maximization (EM) estimation procedure for multivariate current status data using the normal-frailty proportional hazard model. Wang, et al. (2015) proposed an EM and Bayesian estimation approaches for bivariate current status using gamma-frailty Proportional Hazard models. Hens et al., (2009) analyzed multivariate current status using normal-frailty to model the correlation between the survival times. Lin and Wang (2011) presented a Bayesian Proportional Odds model with a shared gamma frailty for clustered and bivariate current status data. These researchers treat the frailties as random effects and assume to follow a parametric distribution. For more details on frailty modeling, refer to Hougaard (2000) and Wienke (2012).

In this paper, a PMLE method is proposed for Cox model with noninformative bivariate current status data. Section 2 presents the proposed constructed likelihood function. The spline-based penalized log-likelihood function is discussed in Section 3. The selection procedure for the smoothing parameters for the two failure times is presented in section 4. The choosing and positioning the knots for the spline-based approximation is discussed in Section 5. Section 6 contains the proposed computational algorithm. Determination of the standard errors and credible interval of the PMLEs is presented in section 7. Monte Carlo simulations to assess the performance of the PMLE approach is discussed in Section 8. For illustrative purposes, the PMLE method is applied to the tumorigenicity experiment dataset in Section 9. Concluding remarks is presented in Section 10.

2. PROPOSED MODELS AND LIKELIHOOD FUNCTION CONSTRUCTION

Suppose a study involves n independent subjects, where each subject is observed only once to ascertain the current state of two events. Noting from Wang and Ding (2000), if the two failure times T_{i1} and T_{i2} are examined from the same subject, then their examination times C_i's can be assumed to be equal, i.e. $C_{i1} = C_{i2} = C_i$. To model the correlation between T_{i1} and T_{i2}, it is assumed that both times are conditionally independent given the subject frailty b_i. Assume the p-covariates $\mathbf{Z}_i$ are time-independent, then evaluating the effects of $\mathbf{Z}_i$ on T_j, $j = 1, 2$ using the Cox frailty model is written as

$$\Lambda_{T_j}(t_i|\mathbf{Z}_i, b_i) = b_i\Lambda_{0j}(t_i)\exp(\boldsymbol{\beta}^T\mathbf{Z}_i) \quad (1)$$

where $\Lambda_{0j}(t_i)$ represents the unknown baseline cumulative hazard function (CHF) of T_j, $j = 1, 2$, $\boldsymbol{\beta}$ denotes the regression parameters. The conditional survival and cumulative distribution functions are respectively written as follows;

$$S_{T_j}(t_i|\mathbf{Z}_i, b_i) = \exp\left(-b_i\Lambda_{0j}(t_i)\exp(\boldsymbol{\beta}'\mathbf{Z}_i)\right) \quad (2)$$

$$F_{T_j}(t_i|\mathbf{Z}_i, b_i) = 1 - \exp\left(-b_i\,\Lambda_{0j}(t_i)\exp(\boldsymbol{\beta}'\mathbf{Z}_i)\right) \quad (3)$$

It can be noted that (1) assumes each T_j has its baseline CHF but have the same covariate effect. The b_i is assumed to follow a gamma distribution with density function

$$g(b_i;\theta) = \frac{b_i^{\left(\frac{1}{\theta}-1\right)} exp\left(-\frac{b_i}{\theta}\right)}{\theta^{1/\theta}\Gamma\left(\frac{1}{\theta}\right)} \quad (4)$$

where the frailty parameter $\theta > 0$ is used to quantify the degree of relationship between the T_1 and T_2. High value of θ indicate a stronger positive correlation and vice-versa. Denote C_i as the examination time and the failure time T_{ij} $(j = 1, 2)$ indicator as

$$\delta_{ij} = \begin{cases} 1 & if\ T_{ij} \le C_i \Rightarrow j^{th}\ event\ is\ left-censored\ at\ C_i \\ 0 & if\ T_{ij} > C_i \Rightarrow j^{th}\ event\ is\ right-censored\ at\ C_i \end{cases}$$

It can be seen that the likelihood function for the complete data set would consist of four cases depending on the categories defined by δ_{ij}. If T_{i1} and T_{i2} are assumed to be independent only when b_i and $\mathbf{Z}_i$ are given, then the joint conditional likelihood associated with these four categories are presented below;

(i) $A_1 = \{i : C_i,\ \delta_{i1} = 0,\ \delta_{i2} = 0,\ \mathbf{Z}_i,\ b_i\}$. The joint conditional likelihood can be written as;

$$\begin{aligned} L_1(\boldsymbol{\Phi}|\mathbf{Z}_i, b_i) &= P(\delta_{i1} = 0,\ \delta_{i2} = 0|\mathbf{Z}_i, b_i) = P(T_{i1} > C_i,\ T_{i2} > C_i|\mathbf{Z}_i, b_i) = P(T_{i1} > C_i|\mathbf{Z}_i, b_i)P(T_{i2} > C_i|\mathbf{Z}_i, b_i) \\ &= S_{T_1}(c_i|\mathbf{Z}_i, b_i)S_{T_2}(c_i|\mathbf{Z}_i, b_i) = \exp\left(-b_i\Lambda_{01}(c_i)\exp(\boldsymbol{\beta}^T\mathbf{Z}_i)\right)\exp\left(-b_i\Lambda_{02}(c_i)\exp(\boldsymbol{\beta}^T\mathbf{Z}_i)\right) \\ &= \exp\left\{-b_i\left(\Lambda_{01}(c_i)\exp(\boldsymbol{\beta}^T\mathbf{Z}_i) + \Lambda_{02}(c_i)\exp(\boldsymbol{\beta}^T\mathbf{Z}_i)\right)\right\} \quad (5) \end{aligned}$$

The unconditional likelihood function can be obtained by integrating out the b_i in (5) using the Laplace transformation concept (Klein et al., 1992) as follows;

$$\begin{aligned} L_1 = L_1(\boldsymbol{\Phi}|\mathbf{Z}_i) &= \int_0^\infty \exp\left\{-b_i\left(\Lambda_{01}(c_i)\exp(\boldsymbol{\beta}^T\mathbf{Z}_i) + \Lambda_{02}(c_i)\exp(\boldsymbol{\beta}^T\mathbf{Z}_i)\right)\right\} g(b_i)\,db_i \\ &= \left\{1 + \theta\left(\Lambda_{01}(c_i)\exp(\boldsymbol{\beta}^T\mathbf{Z}_i) + \Lambda_{02}(c_i)\exp(\boldsymbol{\beta}^T\mathbf{Z}_i)\right)\right\}^{-\frac{1}{\theta}} \quad (6) \end{aligned}$$

where $\mathbf{\Phi} = \left(\boldsymbol{\beta}^T, \Lambda_{01}(\cdot),\ \Lambda_{02}(\cdot),\ \theta\right)$ are the parameters to be estimated.

(ii) $A_2 = \{i : C_i,\ \delta_{i1} = 1,\ \delta_{i2} = 0,\ \mathbf{Z}_i,\ b_i\}$. The joint conditional likelihood is written as

$$L_2(\mathbf{\Phi}|\mathbf{Z}_i, b_i) = P(\delta_{i1} = 1,\ \delta_{i2} = 0|\mathbf{Z}_i, b_i) = P(T_{i1} \le C_i, T_{i2} > C_i|\mathbf{Z}_i, b_i) = P(T_{i1} \le C_i|\mathbf{Z}_i, b_i)P(T_{i2} > C_i|\mathbf{Z}_i, b_i)$$

$$= F_{T_1}(c_i|\mathbf{Z}_i, b_i)S_{T_2}(c_i|\mathbf{Z}_i, b_i) = \left[1 - S_{T_1}(c_i|\mathbf{Z}_i, b_i)\right]S_{T_2}(c_i|\mathbf{Z}_i, b_i) = S_{T_2}(c_i|\mathbf{Z}_i, b_i) - S_{T_1}(c_i|\mathbf{Z}_i, b_i)S_{T_2}(c_i|\mathbf{Z}_i, b_i)$$

The unconditional likelihood can be obtained as follows;

$$L_2 = L_2(\mathbf{\Phi}|\mathbf{Z}_i) = \int_0^\infty S_{T_2}(c_i|\mathbf{Z}_i, b_i)g(b_i)db_i - \int_0^\infty S_{T_1}(c_i|\mathbf{Z}_i, b_i)S_{T_2}(c_i|\mathbf{Z}_i, b_i)g(b_i)\,db_i$$

$$= \left[\int_0^\infty \exp\left\{-b_i\left(\Lambda_{02}(c_i)\exp(\boldsymbol{\beta}^T\mathbf{Z}_i)\right)\right\}g(b_i)\,db_i\right] - L_1 = \left[1 + \theta\left(\Lambda_{02}(c_i)\exp(\boldsymbol{\beta}^T\mathbf{Z}_i)\right)\right]^{-\frac{1}{\theta}} - L_1 \qquad (7)$$

(iii) $A_3 = \{i : C_i,\ \delta_{i1} = 0,\ \delta_{i2} = 1,\ \mathbf{Z}_i,\ b_i\}$. The conditional joint likelihood can be written as

$$L_3(\mathbf{\Phi}|\mathbf{Z}_i, b_i) = P(\delta_{i1} = 0,\ \delta_{i2} = 1|\mathbf{Z}_i, b_i) = P(T_{i1} > C_i, T_{i2} \le C_i|\mathbf{Z}_i, b_i) = P(T_{i1} > C_i|\mathbf{Z}_i, b_i)P(T_{i2} \le C_i|\mathbf{Z}_i, b_i)$$

$$= S_{T_1}(c_i|\mathbf{Z}_i, b_i)F_{T_2}(c_i|\mathbf{Z}_i, b_i) = S_{T_1}(c_i|\mathbf{Z}_i, b_i)\left(1 - S_{T_2}(c_i|\mathbf{Z}_i, b_i)\right) = S_{T_1}(c_i|\mathbf{Z}_i, b_i) - S_{T_1}(c_i|\mathbf{Z}_i, b_i)S_{T_2}(c_i|\mathbf{Z}_i, b_i)$$

The unconditional likelihood is therefore

$$L_3 = L_3(\mathbf{\Phi}|\mathbf{Z}_i) = \int_0^\infty S_{T_1}(c_i|\mathbf{Z}_i, b_i)g(b_i)db_i - \int_0^\infty S_{T_1}(c_i|\mathbf{Z}_i, b_i)S_{T_2}(c_i|\mathbf{Z}_i, b_i)g(b_i)\,db_i$$

$$= \left[\int_0^\infty \exp\left\{-b_i\left(\Lambda_{01}(c_i)\exp(\boldsymbol{\beta}^T\mathbf{Z}_i)\right)\right\}g(b_i)\,db_i\right] - L_1 = \left[1 + \theta\left(\Lambda_{01}(c_i)\exp(\boldsymbol{\beta}^T\mathbf{Z}_i)\right)\right]^{-\frac{1}{\theta}} - L_1 \qquad (8)$$

(4) $A_4 = \{i : C_i,\ \delta_{i1} = 1,\ \delta_{i2} = 1,\ \mathbf{Z}_i,\ b_i\}$. The conditional likelihood can be written as;

$$L_4(\mathbf{\Phi}|\mathbf{Z}_i, b_i) = P(\delta_{i1} = 1,\ \delta_{i2} = 1|\mathbf{Z}_i, b_i) = P(T_{i1} \le C_i, T_{i2} \le C_i|\mathbf{Z}_i, b_i) = P(T_{i1} \le C_i|\mathbf{Z}_i, b_i)P(T_{i2} \le C_i|\mathbf{Z}_i, b_i)$$

$$= F_{T_1}(c_i|\mathbf{Z}_i, b_i)F_{T_2}(c_i|\mathbf{Z}_i, b_i) = \left(1 - S_{T_1}(c_i|\mathbf{Z}_i, b_i)\right)\left(1 - S_{T_2}(c_i|\mathbf{Z}_i, b_i)\right)$$

$$= 1 - S_{T_2}(c_i|\mathbf{Z}_i, b_i) - S_{T_1}(c_i|\mathbf{Z}_i, b_i) + S_{T_1}(c_i|\mathbf{Z}_i, b_i)S_{T_2}(c_i|\mathbf{Z}_i, b_i)$$

The unconditional likelihood is therefore

$$L_4 = L_4(\mathbf{\Phi}|\mathbf{Z}_i) = \int_0^\infty \left(1 - S_{T_2}(c_i|\mathbf{Z}_i, b_i) - S_{T_1}(c_i|\mathbf{Z}_i, b_i) + S_{T_1}(c_i|\mathbf{Z}_i, b_i)S_{T_2}(c_i|\mathbf{Z}_i, b_i)\right)g(b_i)db_i$$

$$= \int_0^\infty g(b_i)db_i - \int_0^\infty S_{T_2}(c_i|\mathbf{Z}_i, b_i)g(b_i)db_i - \int_0^\infty S_{T_1}(c_i|\mathbf{Z}_i, b_i)g(b_i)db_i$$

$$+ \int_0^\infty S_{T_1}(c_i|\mathbf{Z}_i, b_i)S_{T_2}(c_i|\mathbf{Z}_i, b_i)g(b_i)db_i$$

$$= 1 - \left[1 + \theta\left(\Lambda_{01}(c_i)\exp(\boldsymbol{\beta}^T\mathbf{Z}_i)\right)\right]^{-\frac{1}{\theta}} - \left[1 + \theta\left(\Lambda_{02}(c_i)\exp(\boldsymbol{\beta}^T\mathbf{Z}_i)\right)\right]^{-\frac{1}{\theta}} + L_1 \qquad (9)$$

The full log-likelihood functions can thus be written as follows;

$$L_f(\mathbf{\Phi}) = \prod_{i=1}^{n}\left[(L_1)^{(1-\delta_{i1})(1-\delta_{i2})}(L_2)^{\delta_{i1}(1-\delta_{i2})}(L_3)^{(1-\delta_{i1})\delta_{i2}}(L_4)^{\delta_{i1}\delta_{i2}}\right] \qquad (10)$$

$$l(\mathbf{\Phi}) = \sum_{i=1}^{n}\left[(1-\delta_{i1})(1-\delta_{i2})\log L_1 + \delta_{i1}(1-\delta_{i2})\log L_2 + (1-\delta_{i1})\delta_{i2}\log L_3 + \delta_{i1}\delta_{i2}\log L_4\right] \qquad (11)$$

3. SPLINE-BASED PENALIZED MAXIMUM LOG-LIKELIHOOD ESTIMATION

To guarantee the smoothness of $\Lambda_{01}(\cdot)$ and $\Lambda_{02}(\cdot)$, each was penalized by its penalty term defined respectively as $\frac{\kappa_1}{2}\int \Lambda_{01}''^2(t)dt$ and $\frac{\kappa_2}{2}\int \Lambda_{02}''^2(t)dt$. The penalized likelihood, can thus be defined as

$$pl(\mathbf{\Phi}) = l(\mathbf{\Phi}) - \frac{\kappa_1}{2}\int \Lambda_{01}''^2(t)dt - \frac{\kappa_2}{2}\int \Lambda_{02}''^2(t)dt \tag{12}$$

where $\kappa_1 \geq 0$ and $\kappa_2 \geq 0$ control the smoothness level of $\Lambda_{01}(\cdot)$ and $\Lambda_{02}(\cdot)$ respectively. Maximizing (12) gives the PMLE $\hat{\mathbf{\Phi}}$. However, exact computations of the infinite-dimensional parameters of $\Lambda_{01}(t)$ and $\Lambda_{02}(t)$ may pose computational challenges, but can be approximated using cubic I-splines (Ramsay 1988). Thus, each baseline CHF is estimated using;

$$\left.\begin{aligned} \hat{\Lambda}_{01}(t) &\approx \sum_{j=1}^{m} \hat{\alpha}_j I_j(t),\ \alpha_j \geq 0\ \forall j \\ \hat{\Lambda}_{02}(t) &\approx \sum_{j=1}^{m} \hat{\eta}_j I_j(t),\ \eta_j \geq 0\ \forall j \end{aligned}\right\} \tag{13}$$

where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \mathrm{K}, \alpha_m)^T$ and $\boldsymbol{\eta} = (\eta_1, \eta_2, \mathrm{K}, \eta_m)^T$ are the spline coefficients and $\mathbf{I}_i = (I_1(t_i), I_2(t_i), \mathrm{L}, I_m(t_i))^T$ is the I-spline basis vector. To reduce the computational burden of (12), the two penalty terms can further be expressed as follows (Eilers & Marx, 1996);

$$\frac{\kappa_1}{2}\int \Lambda_{01}''^2(t)dt \approx \frac{\kappa_1}{2}\sum_{j=3}^{m}(\Delta^2\alpha_j)^2 = \frac{\kappa_1}{2}\sum_{j=3}^{m}(\alpha_j - 2\alpha_{j-1} + \alpha_{j-2})^2 = \frac{\kappa_1}{2}\boldsymbol{\alpha}^T\mathbf{A}^T\mathbf{A}\boldsymbol{\alpha} = \frac{\kappa_1}{2}\boldsymbol{\alpha}^T\mathbf{R}_1\boldsymbol{\alpha} \tag{14}$$

$$\frac{\kappa_2}{2}\int \Lambda_{02}''^2(t)dt \approx \frac{\kappa_2}{2}\sum_{j=3}^{m}(\Delta^2\eta_j)^2 = \frac{\kappa_2}{2}\sum_{j=3}^{m}(\eta_j - 2\eta_{j-1} + \eta_{j-2})^2 = \frac{\kappa_2}{2}\boldsymbol{\eta}^T\mathbf{C}^T\mathbf{C}\boldsymbol{\eta} = \frac{\kappa_2}{2}\boldsymbol{\eta}^T\mathbf{R}_2\boldsymbol{\eta} \tag{15}$$

where $\mathbf{A}$ and $\mathbf{C}$ are defined penalty matrices each of size $(m-2)\times m$. Thus, equation (12) can be re-written as

$$pl(\mathbf{\Phi}) = l(\mathbf{\Phi}) - \frac{\kappa_1}{2}\boldsymbol{\alpha}^T\mathbf{R}_1\boldsymbol{\alpha} - \frac{\kappa_2}{2}\boldsymbol{\eta}^T\mathbf{R}_2\boldsymbol{\eta} \tag{16}$$

Thus, given κ_1 and κ_2, the PMLEs, $\hat{\mathbf{\Phi}} = (\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\eta}}, \hat{\theta})^T$ can be obtained by maximizing (16).

4. SELECTION OF THE SMOOTHING PARAMETERS κ_1 AND κ_2

Estimating κ_1 and κ_2 simultaneously via cross-validation score (O' Sullivan, 1988) is associated with an extensive computational grid search. For that matter, separate maximization of their respective semiparametric penalized log-likelihood is proposed. Thus, to choose κ_1, the dataset $(C_i, \delta_{i1}, \mathbf{Z}_i)$ is used to compute it using the penalized log-likelihood function

$$l(\boldsymbol{\beta}, \boldsymbol{\alpha}) = \sum_{i=1}^{n}\left(\delta_{i1}\log\left(1 - \exp\left(-\sum_{j=1}^{m}\alpha_j I_j(t)\exp(\boldsymbol{\beta}'\mathbf{Z}_i)\right)\right) - (1-\delta_{i1})\sum_{j=1}^{m}\alpha_j I_j(t)\exp(\boldsymbol{\beta}'\mathbf{Z}_i)\right). \tag{17}$$

$$pl(\boldsymbol{\beta}, \boldsymbol{\alpha}) = l(\boldsymbol{\beta}, \boldsymbol{\alpha}) - \frac{\kappa_1}{2}\boldsymbol{\alpha}^T\mathbf{R}_1\boldsymbol{\alpha} \tag{18}$$

Let $\boldsymbol{\theta}_1 = (\boldsymbol{\beta}, \boldsymbol{\alpha})$, then the approximate cross-validation score (CVS) for (18) is

$$CVS(\kappa_1) \approx -\frac{1}{n}l(\hat{\boldsymbol{\theta}}_1) + \frac{1}{n}trace\left[\left(\hat{\mathbf{H}}_{pl}(\hat{\boldsymbol{\theta}}_1)\right)^{-1}\hat{\mathbf{H}}_l(\hat{\boldsymbol{\theta}}_1)\right]. \tag{19}$$

where $\hat{\mathbf{H}}_{pl}(\hat{\boldsymbol{\theta}}_1)$ and $\hat{\mathbf{H}}_l(\hat{\boldsymbol{\theta}}_1)$ are Hessian matrices of (18) and (17) respectively. Then maximize (19) using grid search to obtain κ_1 and a good initial guess of $\boldsymbol{\theta}_1$. Similarly, to choose κ_2 the dataset $(C_i, \delta_{i2}, \mathbf{Z}_i)$ is used with its penalized log-likelihood function written as

$$l(\boldsymbol{\beta}, \boldsymbol{\eta}) = \sum_{i=1}^{n}\left(\delta_{i2}\log\left(1 - \exp\left(-\sum_{j=1}^{m}\eta_j I_j(t)\exp(\boldsymbol{\beta}'\mathbf{Z}_i)\right)\right) - (1-\delta_{i2})\sum_{j=1}^{m}\eta_j I_j(t)\exp(\boldsymbol{\beta}'\mathbf{Z}_i)\right) \tag{20}$$

$$pl(\boldsymbol{\beta}, \boldsymbol{\eta}) = l(\boldsymbol{\beta}, \boldsymbol{\eta}) - \frac{\kappa_2}{2}\boldsymbol{\eta}^T\mathbf{R}_2\boldsymbol{\eta}\ . \tag{21}$$

Let $\boldsymbol{\theta}_2 = (\boldsymbol{\beta}, \boldsymbol{\eta})^T$, then the approximate CVS for (21) is

$$CVS(\kappa_2) \approx -\frac{1}{n} l(\hat{\boldsymbol{\theta}}_2) + \frac{1}{n} trace\left[\left(\hat{\mathbf{H}}_{pl}(\hat{\boldsymbol{\theta}}_2)\right)^{-1} \hat{\mathbf{H}}_l(\hat{\boldsymbol{\theta}}_2)\right]. \tag{22}$$

where $\hat{\mathbf{H}}_{pl}(\hat{\boldsymbol{\theta}}_2)$ and $\hat{\mathbf{H}}_l(\hat{\boldsymbol{\theta}}_2)$ are the Hessian matrices of (21) and (20) respectively. Maximize (22) using grid search to obtain κ_2 and a good initial guess of $\boldsymbol{\theta}_2$. After obtaining κ_1 and κ_2, (16) is then maximized using the complete dataset to obtain the PMLEs $\hat{\boldsymbol{\Phi}} = (\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\eta}}, \hat{\theta})^T$. The BFGS algorithm (Fletcher, 1987) was used for optimization of all the given objective functions.

5. DETERMINATION OF NUMBER OF KNOTS AND THEIR POSITIONS

The number of knots was determined using the relation $q_n = \lfloor n^{1/3} \rfloor$, where $\lfloor n^{1/3} \rfloor$ is the largest integer below $n^{1/3}$, and the knots positioned at quantiles of the observation times C_i.

6. COMPUTATIONAL ALGORITHM

Equation (16) is maximized using the following proposed computing algorithm;

Step1: Choose a set of say 50 grid points for κ_{1i}, equally spaced from 0.001 to 1.5 and let $i = 1$

Step 2: With κ_{1i}, obtain $\hat{\boldsymbol{\theta}}_1$ of (18) using the BFGS optimization technique.

Step 3: Compute $CVS(\kappa_{1i})$ using (19), let $i = i + 1$

Step 4: Iterate between steps 2 and 3 until all the 50 grid points are exhausted

Step 5: Select the κ_1 that maximize (19) from step 4

Step 6: Recall step 1 and with κ_{2i}, obtain $\hat{\boldsymbol{\theta}}_2$ of (21) using the BFGS optimization technique

Step 7: Compute $CVS(\kappa_{2i})$ using (22) and let $i = i + 1$

Step 8: Repeat steps 6 and 7 until all the 50 grid points are exhausted

Step 9: Select κ_2 that maximized (22) from step 8

Step 10: Let $\hat{\boldsymbol{\Phi}}^{(0)} = (\hat{\boldsymbol{\theta}}_1^{(0)}, \hat{\boldsymbol{\theta}}_2^{(0)}, \hat{\theta}^{(0)})^T$ be initial values, where $\hat{\boldsymbol{\theta}}_1^{(0)}$ and $\hat{\boldsymbol{\theta}}_2^{(0)}$ are obtained from steps 5 and 9 respectively and $\hat{\theta}^{(0)} = 1$. Maximize (16) using the BFGS algorithm to obtain the final PMLEs $\hat{\boldsymbol{\Phi}} = (\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\eta}}, \hat{\theta})^T$.

Step 11: Estimate $\Lambda_{01}(t)$ and $\Lambda_{02}(t)$ using (13).

The positivity of $\boldsymbol{\alpha}$, $\boldsymbol{\eta}$ and θ were ensured by using a square transformation.

7. OBTAINING STANDARD ERRORS AND CREDIBLE INTERVALS

The Bayesian approach (Wahba, 1983) was employed to obtain credible intervals for $\hat{\boldsymbol{\Phi}}$. Thus, $\boldsymbol{\Phi}$ is assumed to be a random variable where (16) is treated as posterior log-likelihood and the penalty terms as the prior log-likelihood. Since the posterior distribution is conditional upon observing the sample, its distribution is used to make statements about $\boldsymbol{\Phi}$. If r is the required number of parameters to estimate, then the conditions which usually guarantee asymptotic normality of ordinary maximum likelihood estimates can easily be adapted to develop the asymptotic properties of the PMLE. Thus, the PMLE $\hat{\boldsymbol{\Phi}}$ asymptotically follows a multivariate posterior Gaussian distribution with mean $\boldsymbol{\Phi}_0$ and variance partitioned as

$$Var(\hat{\boldsymbol{\Phi}}) = I(\hat{\boldsymbol{\Phi}})^{-1} = \begin{bmatrix} \hat{H}^{-1}_{\boldsymbol{\beta\beta}} & . & . & . \\ . & \hat{H}^{-1}_{\boldsymbol{\alpha\alpha}} & . & . \\ . & . & \hat{H}^{-1}_{\boldsymbol{\eta\eta}} & . \\ . & . & . & \hat{H}^{-1}_{\theta\theta} \end{bmatrix}$$

where $\hat{H}^{-1}_{\boldsymbol{\beta\beta}} \in \Re^{p\times p}$, $\hat{H}^{-1}_{\boldsymbol{\alpha\alpha}} \in \Re^{m\times m}$, $\hat{H}^{-1}_{\boldsymbol{\eta\eta}} \in \Re^{m\times m}$ and $\hat{H}^{-1}_{\theta\theta} \in \Re^{1\times 1}$

Following O'Sullivan (1988), and Silverman (1985) a 95% Bayesian credible interval for $\Lambda_{01}(t)$ and $\Lambda_{02}(t)$ are as follows;

$$\left.\begin{array}{l} \hat{\Lambda}_{01}(t) \pm 1.96SE\left(\hat{\Lambda}_{01}(t)\right) \Rightarrow \sum_{j=1}^{m} \hat{\alpha}_j I_j(t) \pm 1.96\sqrt{Var\left(\sum_{j=1}^{m} \hat{\alpha}_j I_j(t)\right)} \\ \mathbf{I}^T(t)\boldsymbol{\alpha} \pm 1.96\sqrt{\mathbf{I}^T(t)Var(\boldsymbol{\alpha})\mathbf{I}(t)} \Rightarrow \mathbf{I}^T(t)\boldsymbol{\alpha} \pm 1.96\sqrt{\mathbf{I}^T(t)\hat{H}^{-1}_{\boldsymbol{\alpha\alpha}}\,\mathbf{I}(t)} \end{array}\right\} \quad (23)$$

$$\left.\begin{array}{l} \hat{\Lambda}_{02}(t) \pm 1.96SE\left(\hat{\Lambda}_{02}(t)\right) \Rightarrow \sum_{j=1}^{m} \hat{\eta}_j I_j(t) \pm 1.96\sqrt{Var\left(\sum_{j=1}^{m} \hat{\eta}_j I_j(t)\right)} \\ \mathbf{I}^T(t)\boldsymbol{\eta} \pm 1.96\sqrt{\mathbf{I}^T(t)Var(\boldsymbol{\eta})\mathbf{I}(t)} \Rightarrow \mathbf{I}^T(t)\boldsymbol{\eta} \pm 1.96\sqrt{\mathbf{I}^T(t)\hat{H}^{-1}_{\boldsymbol{\eta\eta}}\,\mathbf{I}(t)} \end{array}\right\} \quad (24)$$

where $\mathbf{I}^T(t) = \left(I_1(t), I_2(t), \ldots, I_m(t)\right)$ is the I-spline vector. Similarly, the 95% credible intervals for β_j are computed as

$$\hat{\beta}_j \pm 1.96SE(\hat{\beta}_j) \Rightarrow \hat{\beta}_j \pm 1.96\sqrt{\hat{H}^{-1}_{\boldsymbol{\beta\beta},j}}. \quad (25)$$

where $\sqrt{\hat{H}^{-1}_{\boldsymbol{\beta\beta},j}}$ denotes the square root of the j^{th} diagonal value of $\hat{H}^{-1}_{\boldsymbol{\beta\beta}}$. Lastly, the 95% credible intervals for θ is obtained using $\hat{\theta} \pm 1.96SE(\hat{\theta}) \Rightarrow \hat{\theta} \pm 1.96\sqrt{\hat{H}^{-1}_{\theta\theta}}$. (26)

8. SIMULATION STUDIES

To gain a better understanding of the empirical properties of the PMLE, extensive Monte Carlo simulations were conducted using R software. Three different sample sizes of $n =$ 100, 500 and 1,000 were used, with left censoring rates of $\pi_c = 20\%$ *and* 50% for each n. The b_i was generated from the gamma distribution $\Gamma\left(\frac{1}{\theta}, \theta\right)$ with $\theta = 0.5$ and that of the $T_{ij}\ (j = 1, 2)$ from $\Lambda_{T_j}(t_i|\mathbf{Z}_i, b_i) = b_i\Lambda_{0j}(t_i)\exp(\beta_1 Z_{i1} + \beta_2 Z_{i2})$ where $\Lambda_{0j}(t_i) = t_i$, $\beta_1 = 1$, $\beta_2 = 0.6$, $Z_{i1} \sim Bernoulli(0.5)$ and $Z_{i2} \sim U(0, 1)$. The observation times C_i was generated from an exponential distribution within the interval $[0, 2]$ with an appropriate parameter to yield the desired π_c. Therefore, the observed data $\left(C_i, \delta_{i1}, \delta_{i2}, Z_{i1}, Z_{i2}\right)$ is obtained as follows:

(i) Generate T_{i1}, T_{i2} and C_i from their specified distributions stated above

(ii) If $T_{i1} > C_i$ and $T_{i2} > C_i$, then set $\delta_{i1} = 0$ and $\delta_{i2} = 0$; $\Rightarrow (C_i, 0, 0, Z_{i1}, Z_{i2})$

(iii) If $T_{i1} \le C_i$ and $T_{i2} > C_i$, then set $\delta_{i1} = 1$ and $\delta_{i2} = 0$; $\Rightarrow (C_i, 1, 0, Z_{i1}, Z_{i2})$

(iv) If $T_{i1} > C_i$ and $T_{i2} \le C_i$, then $\delta_{i1} = 0$ and $\delta_{i2} = 1$; $\Rightarrow (C_i, 0, 1, Z_{i1}, Z_{i2})$

(v) If $T_{i1} \le C_i$ and $T_{i2} \le C_i$, then set $\delta_{i1} = 1$ and $\delta_{i2} = 1$; $\Rightarrow (C_i, 1, 1, Z_{i1}, Z_{i2})$

To estimate $\Lambda_{01}(t)$ and $\Lambda_{02}(t)$, cubic I-splines was used with the appropriate knots placed at quantiles within [0, 2]. The κ_1 and κ_2 were estimated using the approach proposed. Monte Carlo simulations were conducted based on the PMLE with 500 replications. The first replicate was used to estimate κ_1 and κ_2 and kept constant to generate the remaining 499 datasets. Table 1 presents the summary statistics which includes the average estimates (AEST). It is observed that (i) for fixed π_c, the BIAS, sample standard deviations (SSD), average estimated standard errors (AESE), mean square error (MSE) are decreasing with increasing n, indicating that the PMLEs are exhibiting the consistency property. (ii) the SSD and AESE values are almost equal when compared, demonstrating that the variance estimation procedure is satisfactory (iii) the coverage probabilities (CP) agree very well with the nominal value of 95% under the different n and π_c. (iv) the mean integrated square error (MISE) of $\hat{\Lambda}_{01}(t)$ and $\hat{\Lambda}_{02}(t)$ values show relatively small values, indicating that the proposed PMLE method is generally satisfactory.

In order to provide a numerical justification for the asymptotic normality of the regression and frailty parameters, histograms of β_1, β_2 and θ were constructed for the different n and π_c, and all of them clearly exhibit the normality distribution (sample shown in Figure 1). In addition, plots of the true $\Lambda_{01}(t)$ and $\Lambda_{02}(t)$ and their estimated curves under the different π_c and n were done (sample shown in Figure 2). It is observed that the estimated and the true curves are very close to each other under the different n and π_c, indicating the existence, if any, of little bias in the PMLEs.

To ascertain the effects of knots on the PMLE performance, an analysis was performed for each simulating setting with an arbitrarily chosen number of knots 3, 5, and 8. The summary statistics from the 500 replications of $n = 500$ and 1000 are presented in Tables 2 and 3. It is observed that the estimated values obtained from these knots are almost the as those obtained using the relation $q_n = \lfloor n^{1/3} \rfloor$, revealing that the proposed PMLE method is robust to the choice of knots.

In order to assess the robustness of our procedure to misspecification of the frailty distribution, the simulation studies was conducted using the proposed method but generated the b_i from log-normal distribution with $\theta = 0.5$. Table 4 displays the results. It is observed that the values obtained from the lognormal (L-Normal) are generally similar to those obtained from the correctly specified gamma (Gamma) except a little variation in the values of the frailty parameter where there seem to be an under estimation. Overall, it can be said that the PMLE method is robust to misspecification of the distribution of the frailty variable.

Table 1: Simulation results on the Regression and frailty Parameters and baseline CHFs for the PMLE method based on 500 replications.

n	Interior knot	True Parameters	AEST	BIAS	SSD	AESE	MSE	CP (%)	MISE
			Left censoring rate of 20%						
100	4	$\beta_1 = 1$	1.133	0.133	0.535	0.443	0.304	95.3	
		$\beta_2 = 0.6$	0.734	0.134	0.824	0.675	0.697	95.4	
		$\theta = 0.5$	0.572	0.072	0.721	0.501	0.526	95.0	
		$\Lambda_{01}(t)$						96.7	0.550
		$\Lambda_{02}(t)$						96.4	0.586
500	7	$\beta_1 = 1$	1.032	0.032	0.179	0.190	0.033	97.0	
		$\beta_2 = 0.6$	0.640	0.040	0.324	0.308	0.107	95.6	
		$\theta = 0.5$	0.520	0.020	0.263	0.208	0.070	95.4	
		$\Lambda_{01}(t)$						97.4	1.038
		$\Lambda_{02}(t)$						97.4	1.019
1000	10	$\beta_1 = 1$	1.010	0.010	0.131	0.132	0.017	95.2	
		$\beta_2 = 0.6$	0.595	-0.005	0.213	0.215	0.045	95.5	
		$\theta = 0.5$	0.503	0.003	0.198	0.138	0.039	95.1	
		$\Lambda_{01}(t)$						97.8	1.142
		$\Lambda_{02}(t)$						97.7	1.180
			Left censoring rate of 50%						
100	4	$\beta_1 = 1$	1.090	0.090	0.375	0.360	0.149	96.8	
		$\beta_2 = 0.6$	0.749	0.149	0.606	0.563	0.390	95.0	
		$\theta = 0.5$	0.484	-0.016	0.383	0.360	0.147	95.1	
		$\Lambda_{01}(t)$						97.7	0.894
		$\Lambda_{02}(t)$						96.8	1.254
500	7	$\beta_1 = 1$	1.024	0.024	0.163	0.152	0.027	95.8	
		$\beta_2 = 0.6$	0.615	0.015	0.240	0.240	0.058	96.2	
		$\theta = 0.5$	0.494	-0.006	0.159	0.108	0.025	95.0	
		$\Lambda_{01}(t)$						94.6	1.045

		$\Lambda_{02}(t)$						95.5	1.423
1000	10	$\beta_1 = 1$	1.021	0.021	0.104	0.107	0.011	96.0	
		$\beta_2 = 0.6$	0.614	0.014	0.174	0.170	0.030	95.2	
		$\theta = 0.5$	0.507	0.007	0.110	0.075	0.012	95.4	
		$\Lambda_{01}(t)$						95.1	1.956
		$\Lambda_{02}(t)$						95.6	1.779

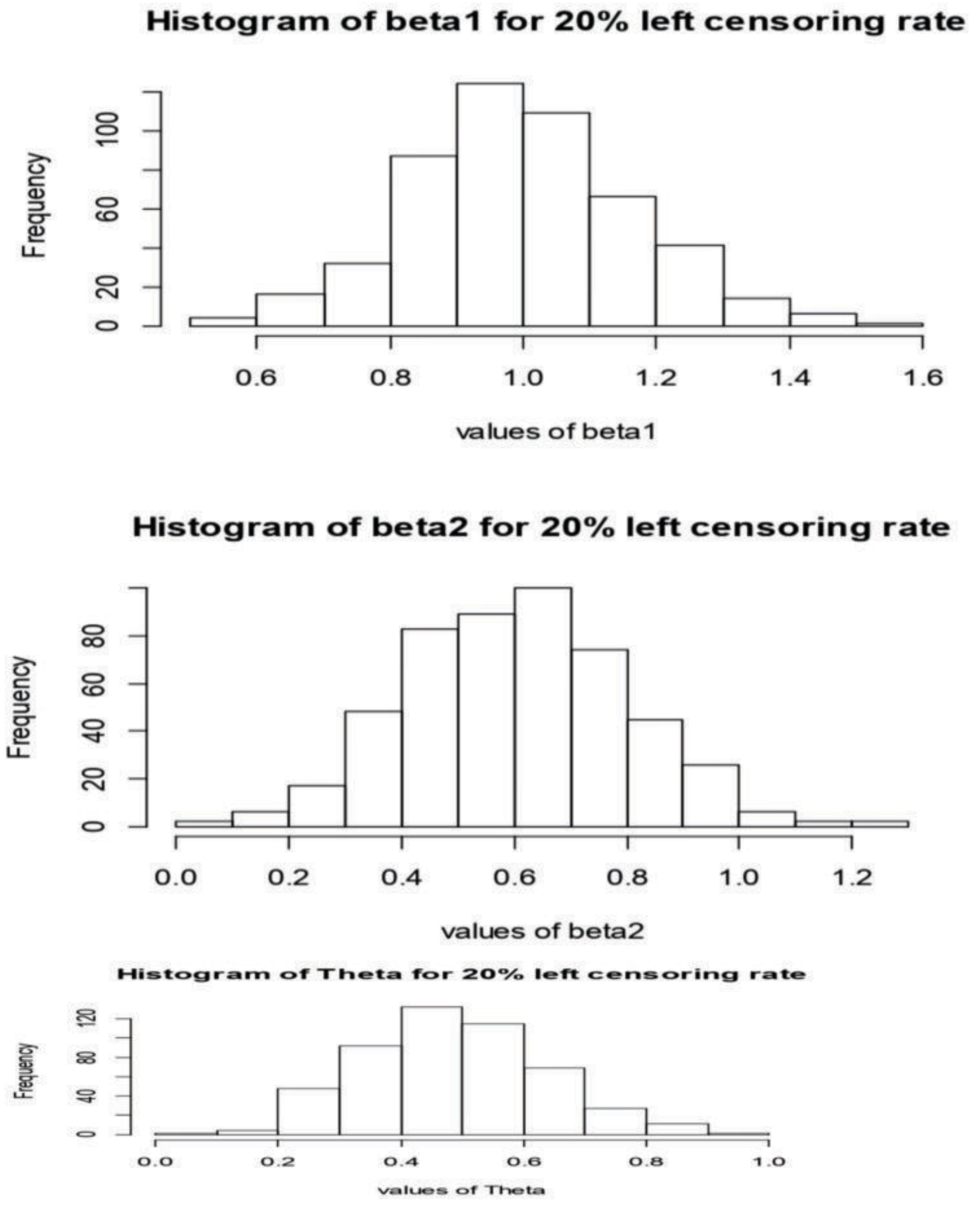
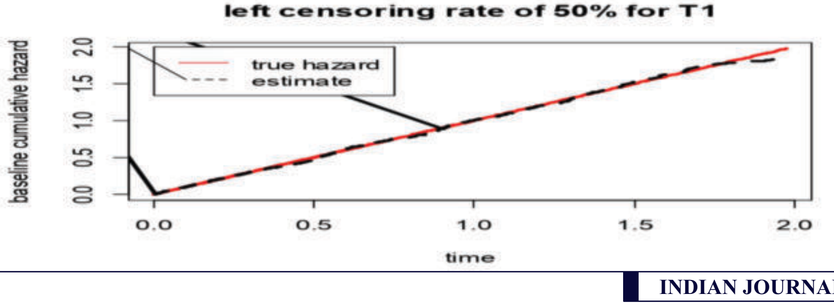


Figure 1: Histogram for the PMLEs $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\theta}$ for $n = 100$ based on 500 replications.



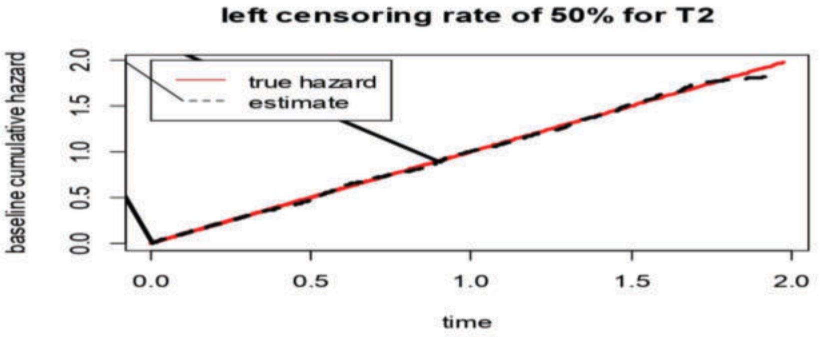


Figure 2: True $\Lambda_{01}(t)$ (top, red) and $\Lambda_{02}(t)$ (bottom, red), PMLEs (black dashes) for $n = 1000$ based on 500 replications.

Table 2: Simulation results on the Regression and frailty Parameters and MISE of $\Lambda_{01}(t)$ and $\Lambda_{02}(t)$ with different knots for $n = 500$ based on 500 replications

Interior knot	True Parameters	AEST	BIAS	SSD	AESE	MSE	CP (%)	MISE
	Left censoring rate of 20%							
3	$\beta_1 = 1$	1.015	0.015	0.190	0.188	0.036	95.0	
	$\beta_2 = 0.6$	0.611	0.011	0.311	0.306	0.097	95.0	
	$\theta = 0.5$	0.511	0.011	0.274	0.212	0.075	95.0	
	$\Lambda_{01}(t)$						95.3	0.993
	$\Lambda_{02}(t)$						95.2	1.052
5	$\beta_1 = 1$	1.012	0.012	0.186	0.188	0.035	95.0	
	$\beta_2 = 0.6$	0.605	0.005	0.330	0.306	0.109	95.4	
	$\theta = 0.5$	0.510	0.010	0.288	0.225	0.083	95.0	
	$\Lambda_{01}(t)$						96.9	1.010
	$\Lambda_{02}(t)$						97.1	1.005
8	$\beta_1 = 1$	1.017	0.017	0.187	0.188	0.035	96.0	
	$\beta_2 = 0.6$	0.624	0.024	0.309	0.304	0.096	95.0	
	$\theta = 0.5$	0.517	0.017	0.283	0.213	0.080	95.1	
	$\Lambda_{01}(t)$						97.6	1.058
	$\Lambda_{02}(t)$						97.6	1.077
	Left censoring rate of 50%							
3	$\beta_1 = 1$	1.018	0.018	0.150	0.151	0.023	95.0	
	$\beta_2 = 0.6$	0.623	0.023	0.249	0.241	0.063	95.0	
	$\theta = 0.5$	0.505	0.005	0.151	0.107	0.023	95.8	
	$\Lambda_{01}(t)$						94.7	1.742
	$\Lambda_{02}(t)$						94.9	1.204
5	$\beta_1 = 1$	1.007	0.007	0.151	0.150	0.023	95.1	
	$\beta_2 = 0.6$	0.630	0.030	0.241	0.239	0.059	95.8	
	$\theta = 0.5$	0.486	-0.014	0.147	0.108	0.022	95.4	
	$\Lambda_{01}(t)$						94.7	1.945
	$\Lambda_{02}(t)$						95.0	1.842
8	$\beta_1 = 1$	1.027	0.027	0.155	0.152	0.025	95.0	
	$\beta_2 = 0.6$	0.627	0.027	0.239	0.242	0.058	96.2	
	$\theta = 0.5$	0.508	0.008	0.157	0.108	0.025	95.1	
	$\Lambda_{01}(t)$						94.7	1.577
	$\Lambda_{02}(t)$						94.9	1.916

Table 3: Simulation results on the Regression and frailty Parameters and MISE of $\Lambda_{01}(t)$ and $\Lambda_{02}(t)$ with different knots for $n = 1000$ based on 500 replications

Interior knot	True Parameters	AEST	BIAS	SSD	AESE	MSE	CP (%)	MISE
	Left censoring rate of 20%							
3	$\beta_1 = 1$	1.014	0.014	0.131	0.133	0.017	95.4	
	$\beta_2 = 0.6$	0.612	0.012	0.221	0.217	0.049	95.6	
	$\theta = 0.5$	0.520	0.020	0.197	0.138	0.039	95.1	
	$\Lambda_{01}(t)$						94.9	1.191
	$\Lambda_{02}(t)$						95.0	1.169
5	$\beta_1 = 1$	1.010	0.010	0.136	0.132	0.019	95.0	
	$\beta_2 = 0.6$	0.611	0.011	0.211	0.215	0.045	95.8	
	$\theta = 0.5$	0.510	0.010	0.184	0.137	0.034	94.6	
	$\Lambda_{01}(t)$						96.2	1.177

	$\Lambda_{02}(t)$						96.7	1.100
8	$\beta_1 = 1$	1.013	0.013	0.136	0.132	0.019	95.6	
	$\beta_2 = 0.6$	0.614	0.014	0.229	0.216	0.052	95.0	
	$\theta = 0.5$	0.502	0.002	0.198	0.139	0.039	94.8	
	$\Lambda_{01}(t)$						97.2	1.194
	$\Lambda_{02}(t)$						97.2	1.308
Left censoring rate of 50%								
3	$\beta_1 = 1$	1.015	0.015	0.107	0.106	0.012	95.2	
	$\beta_2 = 0.6$	0.615	0.016	0.162	0.168	0.026	95.0	
	$\theta = 0.5$	0.498	-0.002	0.101	0.075	0.010	94.8	
	$\Lambda_{01}(t)$						94.6	1.499
	$\Lambda_{02}(t)$						95.0	1.591
5	$\beta_1 = 1$	1.006	0.006	0.102	0.106	0.010	95.4	
	$\beta_2 = 0.6$	0.610	0.010	0.176	0.169	0.0031	95.2	
	$\theta = 0.5$	0.505	0.005	0.105	0.074	0.011	95.2	
	$\Lambda_{01}(t)$						94.9	1.515
	$\Lambda_{02}(t)$						95.1	1.703
8	$\beta_1 = 1$	1.008	0.008	0.103	0.106	0.011	96.2	
	$\beta_2 = 0.6$	0.616	0.016	0.163	0.169	0.027	96.0	
	$\theta = 0.5$	0.502	0.002	0.105	0.075	0.011	94.7	
	$\Lambda_{01}(t)$						94.5	1.895
	$\Lambda_{02}(t)$						94.8	1.179

Table 4: Simulation results of regression and frailty parameters and MISE of $\Lambda_{01}(t)$ and $\Lambda_{02}(t)$ with misspecified frailty distribution based on 500 replications

True parameter	statistical Indicators	$n = 100$		$n = 500$		$n = 1000$	
		Left censoring rate of 20%					
		Interior knot = 4		Interior knot = 7		Interior knot = 10	
Frailty distribution		Gamma	L-Normal	Gamma	L-Normal	Gamma	L-Normal
$\beta_1 = 1$	AEST	1.133	1.175	1.032	0.997	1.010	0.999
	BIAS	0.133	0.175	0.032	-0.003	0.010	-0.001
	SSD	0.535	0.515	0.179	0.172	0.131	0.119
	AESE	0.443	0.389	0.190	0.167	0.132	0.118
	MSE	0.304	0.295	0.033	0.029	0.017	0.014
	CP (%)	95.3	95.7	97.0	94.7	95.2	95.0
$\beta_2 = 0.6$	AEST	0.734	0.742	0.640	0.623	0.595	0.608
	BIAS	0.134	0.142	0.040	0.023	-0.005	0.008
	SSD	0.824	0.749	0.324	0.286	0.213	0.197
	AESE	0.675	0.579	0.308	0.263	0.215	0.190
	MSE	0.697	0.581	0.107	0.082	0.045	0.039
	CP (%)	95.2	94.8	95.6	94.5	95.0	94.5
$\theta = 0.5$	AEST	0.572	0.344	0.520	0.250	0.503	0.234
	BIAS	0.072	-0.156	0.020	-0.250	0.003	-0.266
	SSD	0.721	0.503	0.263	0.199	0.198	0.139
	AESE	0.501	0.459	0.208	0.267	0.138	0.179
	MSE	0.526	0.277	0.070	0.102	0.039	0.090
	CP (%)	94.5	94.7	95.4	94.9	95.1	94.6
$\Lambda_{01}(t)$	MISE	0.550	0.581	1.038	1.187	1.142	1.703
	CP (%)	96.7	96.3	97.4	97.4	97.8	97.5
$\Lambda_{02}(t)$	MISE	0.586	0.554	1.019	1.132	1.180	1.590
	CP (%)	96.4	96.3	97.4	97.6	97.7	97.4
Left censoring rate of 50%							
$\beta_1 = 1$	AEST	1.090	1.099	1.024	1.001	1.021	1.009
	BIAS	0.090	0.099	0.024	0.001	0.021	0.009
	SSD	0.375	0.353	0.163	0.138	0.104	0.095
	AESE	0.360	0.313	0.152	0.133	0.107	0.095
	MSE	0.149	0.135	0.027	0.019	0.011	0.009
	CP (%)	96.8	95.3	95.8	94.5	96.0	95.8
$\beta_2 = 0.6$	AEST	0.749	0.673	0.615	0.615	0.614	0.616

	BIAS	0.149	0.073	0.015	0.015	0.014	0.016
	SSD	0.606	0.540	0.240	0.212	0.174	0.149
	AESE	0.563	0.487	0.240	0.208	0.170	0.147
	MSE	0.390	0.296	0.058	0.045	0.030	0.022
	CP (%)	94.5	94.8	96.2	95.0	95.2	94.8
$\theta = 0.5$	AEST	0.484	0.224	0.494	0.197	0.507	0.210
	BIAS	-0.016	-0.276	-0.006	-0.303	0.007	-0.290
	SSD	0.383	0.272	0.159	0.111	0.110	0.085
	AESE	0.360	0.420	0.108	0.158	0.075	0.099
	MSE	0.147	0.150	0.025	0.104	0.012	0.091
	CP (%)	94.5	94.7	94.8	95.0	95.1	94.7
$\Lambda_{01}(t)$	MISE	0.894	2.217	1.045	2.854	1.956	2.121
	CP (%)	95.2	94.5	95.0	94.9	95.1	95.0
$\Lambda_{02}(t)$	MISE	1.254	2.374	1.423	2.463	1.779	2.054
	CP (%)	95.0	94.5	95.0	95.0	95.6	95.0

9. PRACTICAL ILLUSTRATION

The PMLE method was applied to tumorigenicity data that contains lung and bladder tumors of 671 female mice (Lindsey & Ryan, 1994). The goal of this study is to identify whether a suspected environment accelerates the onset of both lung and bladder tumors in the mice. To achieve this, the mice were randomly assigned either to a control group (387 mice) or High-dose group (284 mice). Each animal was examined at its death time to ascertain the current status of these two sites. Animals found to have developed the tumors, the exact onset times will not be known to the researcher, and in that case, the onset times is said to be left-censored; otherwise, they are said to be right-censored. At the end of 33 months of follow-up, 121 and 124 mice had left censored lung and bladder tumor onset times respectively. The death time C_i is assumed to be noninformative of the onset times of lung and bladder tumors. Mice assigned to the control treatment are coded $Z_i = 0$ and 1 for those assigned to the high-dose. The analysis of the lung and bladder tumors dataset was jointly carried out to satisfy the bivariate current status data. To approximate $\Lambda_{01}(t)$ and $\Lambda_{02}(t)$, the cubic I-splines was used with 8 knots placed at quantiles of C_i. The smoothing parameters κ_1 and κ_2 were determined using the approach proposed. It can be observed from the results presented in Table 5 that the PMLE of the effect of environment on growth of both lung and bladder tumors gave a p-value $p = 8.882 \times 10^{-16}$, indicating that the high-dose treatment has a significant effect on the development of both lung and bladder tumors. The estimated frailty parameter is $\hat{\theta} = 2.081$, which suggests a strong correlation between the onset times of lung and bladder tumors. Thus, exposure to high-dose environment increases the risk of developing both lung and bladder tumors and their occurrence rates are positively significantly correlated.

Table 5: Parameter Estimates of effect of the High-dose based on a Gamma frailty

Statistics	Regression and frailty parameters	
	β	θ
Estimate	1.806	2.081
SE	0.218	0.162
95% CI	(1.378, 2.234)	(1.764, 2.398)
t - value	8.267	12.880
p – value	8.881784e-16	0.000

10. CONCLUDING REMARKS

In this study, a PMLE approach was proposed to fit a bivariate current status data with noninformative observation time. The proportional hazards gamma frailty model was employed, where the gamma distribution was used to model the correlation between the two failure times of interest. A proposed log-likelihood function was developed and the quadratic penalty function for each baseline function was subtracted from this log-likelihood function to obtain the penalized log-likelihood function. The baseline CHFs for the two failure times were approximated using the cubic I-spline (Ramsay, 1988). The smoothing parameters were selected by maximizing their respective penalized log-likelihood using the cross-validation score (O'Sullivan, 1988). The PMLEs were estimated using a combination of Broyden-Fletcher-Goldfarb-Shanno (BFGS) quasi-Newton algorithm (Fletcher, 1987).

The statistical properties of the PMLEs were studied through extensive simulations and the results obtained suggest that the estimators performed quite and reasonably well. The estimators were consistent, asymptotically normal and efficient. The proposed method of optimal selection of knots and smoothing parameters proved to be satisfactory. In addition, the estimators were robust to the choice of knots, rate of left censoring and type of frailty distribution used. For illustration purposes, the proposed method was applied to real data on tumorigenicity experiment by Lindsey and Ryan (1994). The findings of the real data coincide very well with previous findings by other authors including Lindsey and Ryan (1994).

REFERENCES

1. Chen, M. H. (2007). *Statistical Analysis of Multivariate Interval-censored Failure time Data.* PhD Thesis. University of Missouri, Columbia.
2. Eilers, P. H., & Marx, B. D. (1996). Flexible smoothing with B-splines and penalties. *Statistical Science*, *11*, 89 – 121.
3. Fletcher, R. (1987). *Practical Methods of Optimization (2nd edn.)*. John Wiley & Sons, New York.
4. Hens, N., Wienke, A., Aerts, M., & Molenberghs, G. (2009). The correlated and shared gamma frailty model for bivariate current status data: An illustration for cross-sectional serological data. *Statistics in Medicine, 28*, 2785 – 2800.
5. Hougaard, P. (2000). *Analysis of multivariate survival data*. Springer: New York.
6. Klein, J. P., Moeschberger, M. L., Li, Y. H., & Wang, S. T. (1992). Estimating rand-om effects in the Framingham heart study; Survival Analysis: State of the Art, Kluwer Academic: Boston, Massachusetts, 99 – 120.
7. Lindsey, J.., & Ryan, L. (1994). A comparison of continuous- and discrete-time three-state models for rodent tumorigenicity experiments. *Environmental Health Pe-spectives Supplements, 102*, 9 – 17.
8. Lin, X., & Wang, L. (2011). Bayesian proportional odds models for analyzing current status data: univariate, clustered and multivariate. *Communications in Statistics – Simulation and Computation*, *40*: 1171 – 1181.
9. O'Sullivan (1988). Fast computation of fully automated log-density and log-hazard estimators. *SIAM Journal of Science and Statistical Computation, 9*: 363 – 379.
10. Ramsay (1988). Monotone regression splines in action. *Statistical Science, 3*: 425-441
11. Silverman, B. W. (1985). Some aspects of the spline smoothing approach to nonpara-metric regression curve fitting. *Journal of Royal Statistical Society B*, *47,* 1 – 52.
12. Wahba, G. (1983). Bayesian confidence intervals for the cross-validated smoothing spline. *Journal of the Royal Statistical Society, Series B, 45*: 133 – 150.
13. Wang, N., Wang, L., & McMahan, C. S. (2015). Regression analysis of bivariate current status data under the Gamma-frailty PH using EM algorithm. *Computational Statistics & Data Analysis*, *83*: 140 – 150.
14. Wienke, A. (2012). *Frailty Models in Survival Analysis*. Chapman & Hall
15. Wang, W., & Ding, A. A. (2000). On assessing the association for bivariate current status data. Biometrika, *87*, 879 – 893.