Journal of Statistical Research *ISSN 0256 - 422 X*
2019, Vol. 53, No. 2, pp. 129-146

ON TRUNCATED POISSON EXPONENTIAL PROPORTIONAL HAZARD MODEL

AHSAN RAHMAN JAMEE*
Department of Statistics, University of Dhaka, Dhaka 1000, Bangladesh
Email: jaamee08@gmail.com

WASIMUL BARI
Department of Statistics, University of Dhaka, Dhaka 1000, Bangladesh
Email: w_bari@du.ac.bd

SUMMARY

Classical survival regression models may provide misleading results when event of interest occurs due to more than one causes. In this paper, taking all possible causes for the occurrence of event into account, a Truncated Poisson Exponential survival proportional hazard model has been proposed. An extensive simulation study has been conducted to examine the performance of the proposed survival model in the absence and presence of covariates under different percentages of censoring. The simulation results reveal that estimators of the regression parameters are consistent and efficient. To illustrate the model, under-five child survival data extracted from Bangladesh Demographic and Health Survey 2014 have been used.

Keywords and phrases: Survival time; exponential distribution; truncated Poisson distribution; proportional hazard; censoring

1 Introduction

In classical survival analysis, it is assumed that the event of interest occurs from one and only cause. But in practice, it may happen that for the occurrence of an event, more than one causes may be responsible. Such situation may arise in several areas, such as public health, actuarial science, biomedical studies, demography and industrial reliability. For example, under-5 child mortality may occur due to disease, non-disease, or other causes (Mohammad et al., 2017); in a parallel circuit, the system will fail if all the components of this circuit failed (Basu and Ghosh, 1980). Competing-risk and complementary-risk models (Basu, 1980; Basu and Ghosh, 1980; Basu and Klein, 1982) are available in literature to handle such complex situations. In competing risk model with r causes, minimum of r latent survival times is observed and hence occurrence of event due to one of these causes precludes the occurrence of event due to rest of the causes. Under this model, for each cause, cause-specific hazards model or cumulative incidence functions (CIF) approach is used for analyzing

* Corresponding author
© Institute of Statistical Research and Training (ISRT), University of Dhaka, Dhaka 1000, Bangladesh.

survival data (Pintile, 2007). On the other hand, complementary risk model is similar to competing risk models, where maximum lifetime is observed for the occurrence of an event of interest.

Analyzing survival data using competing or complementary risk model, one has to know which cause is responsible for the occurrence of the event of interest and has to fit a survival model for that cause. Therefore, a number of survival models are considered to analyze the survival data and hence complexities arise in estimating parameters and in the interpretation of results. Moreover, in practice, it is almost impossible to identify which cause is responsible for the occurrence of an event.

To overcome this problem, Kuş (2007), Louzada-Neto et al. (2011) and Ramos et al. (2019) proposed a survival model based on zero truncated Poisson distribution for number of causes and exponential distribution for survival times. Bereta et al. (2011) introduced Poisson-Weibull three parameters lifetime distribution in the presence of more than one causes responsible for the occurrence of event. To examine how a set of covariates influence the survival time Louzada et al. (2012) suggested Poisson-exponential survival regression model. As the interpretation of regression parameters of this model is not straight forward, in this paper, a Truncated Poisson Exponential (TPE) proportional hazard model has been proposed so that one can interpret the regression parameter easily.

Louzada-Neto et al. (2011) have developed Poisson-Exponential distribution with increasing failure rate based on latent complementary risks problem where there is no information about which cause was responsible for the component failure and only the maximum lifetime value among all causes is observed. Kuş (2007) proposed a two-parameter distribution known as exponential-Poisson distribution, which has decreasing failure rate. Louzada et al. (2012) proposed a class of hierarchical models with latent competing risks and different activation schemes, where the lifetime associated with a particular cause is not observable, rather only the minimum, maximum or a randomly lifetime value among all cause is observed, with the lifetimes following an exponential distribution. To examine the influence of covariates on the survival time, Louzada et al. (2012) incorporated covariates into the Poisson-Exponential distribution by replacing hazard parameter of exponential distribution and parameter of zero-truncated Poisson distribution with linear predictor using log link function. The main limitation of this suggested model is that interpretation of regression coefficients is not tractable. Regression coefficients only provide the pattern of change in survival quantities, but using the model suggested by Louzada et al. (2012) it is not possible to determine to what extent survival survival quantities change with the change in the value of covariates. As a remedy of this difficulty, following proportional hazard (PH) model (Cox, 1972), a parametric survival regression model has been proposed based on a Truncated Poisson Exponential (TPE) distribution. One can interpret regression coefficients using hazard ratio. Note that the proposed model provides decreasing hazard rate, whereas constant hazard rate is obtained from the classical exponential survival model. This is because a constant baseline hazard function is used in classical exponential model and a decreasing baseline hazard function in TPE proportional hazard model.

An extensive simulation study has been conducted to examine the performance of the TPE model as well as the proposed Truncated Poisson Exponential proportional hazard (TPEPH) model in the absence and presence of covariates, respectively. Simulation results obtained from TPEPH model have also been compared with the results obtained from the classical exponential regression model.

To illustrate the proposed model, a real data set extracted from Bangladesh Demographic and Health Survey (BDHS), 2014 data has been used, where the event of interest is survival of a child before reaching his/her fifth birthday.

The paper is organized as follows. In Section 2, TPE survival model is introduced along with inference procedure in the absence and presence of covariates, respectively. Simulation setup and results are given in Section 3. Section 4 illustrates the proposed model with a real data set. This paper concludes in Section 5 with some comments.

2 Truncated Poisson-Exponential Survival Regression Model

2.1 Absence of Covariates

Let R be a random variable denoting the number of causes related to the occurrence of an event of interest. One may assume that R has a zero-truncated Poisson distribution with probability mass function,

$$P[R = r] = \frac{e^{-\theta}\theta^r}{r!(1-e^{-\theta})} \text{ and } E[R = r] = \frac{\theta}{1-e^{-\theta}} \quad r = 1, 2, \ldots; \quad \theta > 0.$$

Let T_d $(d = 1, \ldots, r)$ denote the latent time-to-event due to the d^{th} cause for given $R = r$. Under this setup, lifetimes $T_1, \ldots, T_r$ are assumed to be independent and identically distributed with exponential probability density function, $f(t, \lambda) = \lambda e^{-\lambda t}; \quad t > 0, \ \lambda > 0$ (Kuş, 2007). The lifetime associated with every causes are not observable, but only minimum lifetime, Y^0 among all causes can be observed, i.e., $Y^0 = min(T_1, \ldots, T_r)$. The probability density function of Y^0 as well as survival function, hazard function and p^{th} quantile times under TPE distribution are (Kuş 2007; Louzada et al. , 2012).

$$f_{Y^0}(y) = \frac{\lambda\theta e^{-\lambda y + \theta e^{-\lambda y}}}{e^{\theta} - 1} \tag{2.1}$$

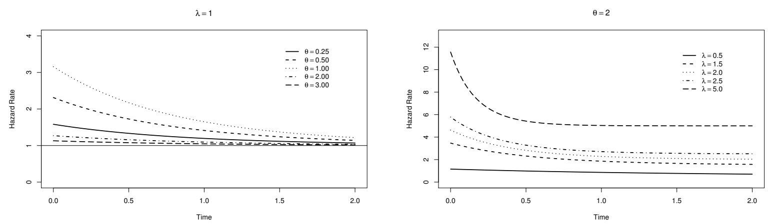
$$S_{Y^0}(y) = \frac{e^{\theta e^{-\lambda y}} - 1}{e^{\theta} - 1} \tag{2.2}$$

$$h_{Y^0}(y) = \frac{\lambda\theta \exp\left(-\lambda y + \theta e^{-\lambda y}\right)}{\exp\left(\theta e^{-\lambda y}\right) - 1} \tag{2.3}$$

$$y_p^0 = -\frac{1}{\lambda}\left[ln\left(ln[p + e^{\theta}(1-p)]\right) - ln\theta\right] \tag{2.4}$$

$y, \theta, \lambda > 0$; respectively.

Note that hazard function obtained under this model decreases as lifetime progresses, whereas the classical exponential survival model provides constant hazard rate. The graphs of hazard function given in (2.3) are presented in Figure 1.



(a) TPE hazard function for fixed λ (b) TPE hazard function for fixed θ

Figure 1: TPE hazard function

2.2 Presence of Covariates

Let $Y = \min(T_1, \ldots, T_r)$ be the survival time in the presence of a set of covariates $x = (x_1, \ldots, x_p)'$ and $\beta = (\beta_1, \ldots, \beta_p)'$ be the $(p \times 1)$ vector of regression parameters. Hence one can then define linear predictor η as $\eta = x'\beta = \sum_{j=1}^{p} x_j \beta_j$. Following multiplicative hazard model (Johnson and Kotz, 2005) in the presence of covariates x, the hazard function can be written as

$$h_Y(y|x) = h_{Y^0}(y) \times \exp(\eta), \tag{2.5}$$

where $h_{Y^0}(y)$ is given in (2.3). This model is known as proportional hazard (PH) model when covariates are time independent (Cox, 1972). Under this PH model hazard and survival functions are given as

$$h_Y(y) = \frac{\lambda\theta e^{-\lambda y + \theta e^{-\lambda y} + x'\beta}}{e^{\theta e^{-\lambda y}} - 1} \tag{2.6}$$

$$S_Y(y) = \left[\frac{e^{\theta e^{-\lambda y}} - 1}{e^{\theta} - 1}\right]^{\exp(x'\beta)}, \tag{2.7}$$

respectively. Hence the probability density function is $f_y(y) = h_Y(y) \times S_Y(y)$. Note that the hazard ratio (HR) for covariate x_j is given as $\exp(\beta_j)$, when is no interaction term with x_j is present in the model.

2.3 Inference Procedure

Under the TPEPH model the main parameter of interest is the regression vector β. The other parameters in the model λ and θ are considered as nuisance parameters. Let $\Theta = (\beta', \lambda, \theta)'$. Under this survival regression model, it is assumed that event for an individual can be occurred due to any of the available causes. Suppose that survival times are obtained from n independent individuals and this times are $Y_1, \ldots, Y_i, \ldots, Y_n$. Also let $x_i = (x_{i1}, \ldots, x_{ij}, \ldots, x_{ip})'$ be the covariate vector associated with the lifetime $Y_i, i = 1, 2, \ldots, n$. One may obtain estimates of parameter of interest Θ

by using maximum likelihood estimation technique. Under random censoring scheme the likelihood function can be written as

$$\begin{aligned} L(\Theta; y_1, \ldots, y_i, \ldots, y_n) &= \prod_{i=1}^{n} \left[h_Y(y_i)\right]^{\delta_i} S_Y(y_i) \\ &= \prod_{i=1}^{n} \frac{\lambda^{\delta_i} \theta^{\delta_i} e^{-\lambda \delta_i y_i + \theta \delta_i e^{-\lambda y_i} + \delta_i x_i' \beta}}{(e^{\theta} - 1)^{e^{x_i'\beta}}} \left[e^{\theta e^{-\lambda y_i}} - 1\right]^{e^{x_i'\beta} - \delta_i}, \qquad (2.8) \end{aligned}$$

where δ_i is the censoring indicator for the i^{th} individual taking value 1 for uncensored observation and 0 for censored observation.

Then the score function of Θ, denoted by $U(\Theta)$, can be defined as

$$U(\Theta; y) = \left[U_1(\Theta), U_2(\Theta), U_3(\Theta), \ldots, U_{j+2}(\Theta), \ldots, U_{p+2}(\Theta)\right]',$$

where

$$\begin{aligned} U_1(\Theta) &= \frac{\delta \log L(\Theta; y)}{\delta \lambda} \\ &= \frac{\sum_{i=1}^{n} \delta_i}{\lambda} - \sum_{i=1}^{n} \left[y_i \delta_i + \theta y_i \delta_i e^{-\lambda y_i}\right] - \sum_{i=1}^{n} \frac{(e^{x_i'\beta} - \delta_i)(\theta y_i e^{-\lambda y_i + \theta e^{-\lambda y_i}})}{\left(e^{\theta e^{-\lambda y_i}} - 1\right)} \qquad (2.9) \\ U_2(\Theta) &= \frac{\delta \log L(\Theta; y)}{\delta \theta} \\ &= \frac{\sum_{i=1}^{n} \delta_i}{\theta} + \sum_{i=1}^{n} \delta_i e^{-\lambda y_i} + \sum_{i=1}^{n} \frac{(e^{x_i'\beta} - \delta_i)(e^{-\lambda y_i + \theta e^{-\lambda y_i}})}{\left(e^{\theta e^{-\lambda y_i}} - 1\right)} - \sum_{i=1}^{n} \frac{e^{\theta + x_i'\beta}}{e^{\theta} - 1} \qquad (2.10) \\ U_{j+2}(\Theta) &= \frac{\delta \log L(\Theta; y)}{\delta \beta_j}; j = 1, \ldots, p \\ &= \sum_{i=1}^{n} \delta_i x_{ij} + \sum_{i=1}^{n} x_{ij} e^{x_i'\beta} \log(e^{\theta e^{-\lambda y_i}} - 1) - \sum_{i=1}^{n} x_{ij} e^{x_i'\beta} \log(e^{\theta} - 1). \qquad (2.11) \end{aligned}$$

The related information matrix, denoted by $I(\Theta)$, can be obtained as $I(\Theta) = -\frac{\delta}{\delta \Theta'} U(\Theta)$ and the maximum likelihood estimating equation is given as $U(\Theta) = 0$. For the purpose of obtaining maximum likelihood estimates of β, λ and θ, R codes using 'optimization' and 'countreg' packages have been written in RStudio. Note that estimators $\hat{\Theta}$ obtained from the likelihood function follows asymptotically normal distribution with mean Θ and variance $[I(\Theta)]^{-1}$, i.e. $\hat{\Theta} \sim N\left[\Theta, [I(\Theta)]^{-1}\right]$.

3 Simulation Study

To examine the performance of Truncated Poisson Exponential survival model in the presence and absence of covariates, an extensive simulation study has been conducted. In simulation study small as well as large sample size is considered such as $n = 100,\ 250,\ 500$ and 1000. For each sample, 1000 simulation trials have been considered so that simulated mean (SM), simulated standard error (SSE) and simulated mean squared error (MSE) of the estimate of parameters can be computed. Censoring in the data set has been introduced through random censoring scheme. To consider $100(1-\epsilon)\%$ censoring, censoring indicator, δ has been generated by using Bernoulli distribution so that $P[\delta = 1] = \epsilon$. It implies that $100(1-\epsilon)\%$ of generated survival times will be randomly considered as censored observations.

3.1 Simulation: Absence of Covariates

To generate Truncated Poisson-Exponential (TPE) survival time for different sample sizes, first, number of causes, R has been generated by using zero truncated Poisson distribution with parameter θ. Then for $R = r$, survival times $T_1, \ldots, T_r$ are generated from exponential distribution with scale parameter λ for each individual. Finally, for the i^{th} individual ($i = 1, 2, \ldots, n$) final survival time has been selected as $Y_i^0 = min(T_{i1}, \ldots, T_{ir})$. This process is repeated to generate random sample of size n. In this setup, 0%, 5%, 10%, 20%, 30% and 50% censoring have been considered and the true values of λ and θ are $\lambda = 0.5\ and\ 1.5$; $\theta = 2\ and\ 3$.

3.1.1 Results

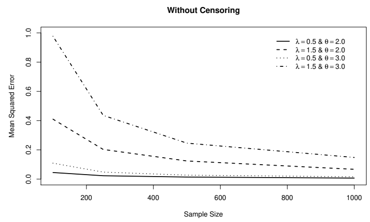
Results obtained from simulation studies have been presented graphically. In Figure 2, MSEs for $\hat{\lambda}$ and $\hat{\theta}$ for different sample sizes were plotted for different combinations of λ and θ under 0% and 10% censoring. For all other percentages of censoring, the mean squared errors for $\hat{\lambda}$ and $\hat{\theta}$ were given in Figure A1, which are reported in Appendix. It is clear from this figures that MSEs of $\hat{\lambda}$ are small, but relatively higher for small sample sizes and it decreases as sample size increases when the percentage of censoring is less than 30%. It is observed from Figure A2 that for 30% and 50% censoring the MSEs of $\hat{\lambda}$ are small, but remain almost unchanged irrespective of sample sizes.

From Figure 2 and A1, it is found that MSEs of $\hat{\theta}$ are small and these decrease as sample size increases. Note that MSEs of $\hat{\theta}$ are larger compared to that of $\hat{\lambda}$.

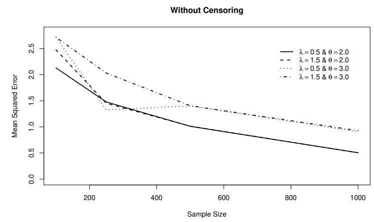
3.2 Simulation: Presence of Covariates

In this paper, covariates into survival model are incorporated following Cox PH model (Cox, 1972) using the form of baseline hazard model given in (2.3) with parameter λ and θ. Let x be a vector of covariates with regression parameters β. To generate survival times under TPEPH model, first, scale parameter λ^* in the presence of covariates x is defined as $\lambda^* = \lambda \exp(x'\beta)$. Following Section (3.1), number of causes, R is generated from zero truncated Poisson distribution with parameter θ and $R = r$ lifetimes are generated from exponential distribution with scale parameter λ^* for each individual. Final lifetimes Y is the minimum of these r lifetimes.

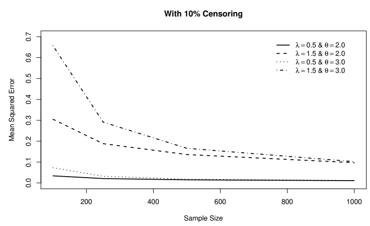
For the simulation setup, two covariates X_1 and X_2 were considered , where values of X_1 are obtained from Bernoulli distribution with success probability 0.8, i.e. $Pr[X_1 = 1] = 0.8$ and the values of X_2 are obtained from uniform distribution with parameters -1, and 1, i.e. $X_2 \sim uniform(-1, 1)$. True values of the regression parameters are considered as $\beta_1 = 0.40\ and\ \beta_2 = -0.30$. Different values for λ and θ have been chosen such as $\lambda = 0.5\ and\ 1.5$; $\theta = 2\ and\ 3$. The



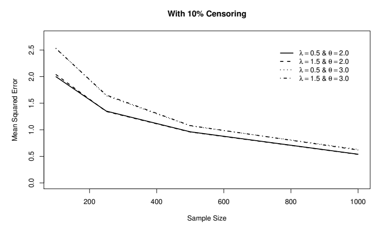
(a) Mean Squared Error of $\hat{\lambda}$: 0% censoring



(b) Mean Squared Error of $\hat{\theta}$: 0% censoring



(c) Mean Squared Error of $\hat{\lambda}$: 10% censoring



(d) Mean Squared Error of $\hat{\theta}$: 10% censoring

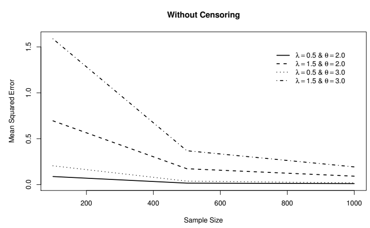
Figure 2: MSE($\hat{\lambda}$) and MSE($\hat{\theta}$) of Truncated Poisson Exponential survival model

simulation has been carried out with complete as well as 5%, 10%, 20%, 30% and 50% censored data.

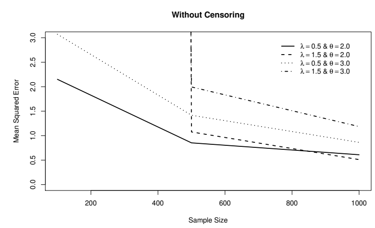
3.2.1 Results

Under the TPEPH model, the main parameter of interest is regression parameter $\beta = (\beta_1, \beta_2)'$. The parameter θ in zero truncated Poisson distribution and λ in exponential distribution may be considered as nuisance parameters. Simulated mean (SM), simulated standard error (SSE), mean squared error (MSE) and 95% coverage probability (CP) of regression parameters obtained from simulation study for different sample sizes without and with 5% censoring are reported in Table 1. Similar results for 10%, 20%, 30% and 50% censoring were reported in Appendix in Table A1 and A2. Simulation results for λ and θ are presented graphically using MSEs for different sample sizes. Figure 3 shows the MSEs for $\hat{\lambda}$ and $\hat{\theta}$ under without and 5% censoring schemes, whereas Figure A2 in Appendix represent MSEs for $\hat{\lambda}$ and $\hat{\theta}$ under 10%, 20%, 30% and 50% censoring schemes.

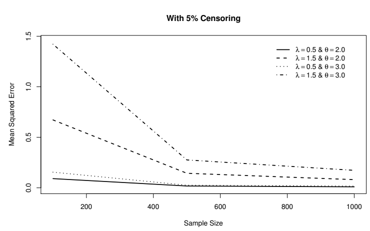
It is clear from Table 1, A1 and A2 that for all combinations of θ and λ, the MSEs of estimators $\hat{\beta}_1$ and $\hat{\beta}_2$ are small under without and all percentages of censoring considered in the survival data. It is also observed that these MSEs reduce to zero as sample size increases. It implies that these estimators are consistent. It is also found that the coverage probabilities for β_1 and β_2 are close to nominal confidence level 95%.



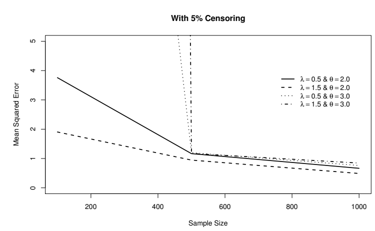
(a) MSE of $\hat{\lambda}$: Without censoring



(b) SE of $\hat{\theta}$: Without censoring



(c) SE of $\hat{\lambda}$: 5% censoring



(d) MSE of $\hat{\theta}$: 5% censoring

Figure 3: MSE($\hat{\lambda}$) and MSE($\hat{\theta}$) of Truncated Poisson Exponential Proportional Hazard model

It is clear from Table 1 that for sample size 100, $\theta = 3$ and $\lambda = 0.5$ the SM for β_1 and β_2 are 0.399 and -0.308, respectively with SSE 0.295 and 0.248; under no censoring scheme, whereas these values are 0.403 and -0.302 with SSE 0.318 and 0.235 under 5% censoring scheme. When the sample size is 1000, SM of β_1 is 0.397 with SSE 0.101 and for β_2, the SM is -0.297 with SSE 0.066 when there is no censoring considered in the survival data. Under the same setup with 5% censoring, the SM of β_1 is 0.397 with SSE 0.104 and for β_2, the SM is -0.302 with SSE 0.068. Because of convergence problem, simulation results were not found for sample size 100 with (θ, λ) combination (2,0.5) under 30% and 50% censoring schemes; and (3,0.5) under 50% censoring scheme.

Figures 3 and A2 reveal the fact that MSEs of estimators of λ and θ decrease when sample size increases under no censoring as well as all censoring schemes. It is also observed that for small sample size, the estimator of θ provides much higher amount of MSE compared to that of λ. In general, the MSEs of $\hat{\lambda}$ and $\hat{\theta}$ under TPEPH model follow the pattern that was observed for $\hat{\lambda}$ and $\hat{\theta}$ under TPE model.

Table 1: Simulated mean (SM), simulated standard Error (SSE), mean squared error (MSE) and 95% coverage probability (CP) of regression parameters in TPEPH model obtained from 1000 simulations

				Without Censoring				**5% Censoring**			
n	θ	λ	β	**SM**	**SSE**	**MSE**	**95% CP**	**SM**	**SSE**	**MSE**	**95% CP**
100	2	0.5	β_1 = 0.40	0.387	0.293	0.086	0.934	0.394	0.333	0.111	0.930
			β_2 = -0.30	-0.288	0.218	0.048	0.923	-0.294	0.221	0.049	0.930
		1.5	β_1 = 0.40	0.394	0.295	0.087	0.925	0.370	0.306	0.095	0.929
			β_2 = -0.30	-0.287	0.208	0.043	0.937	-0.307	0.229	0.052	0.940
	3	0.5	β_1 = 0.40	0.399	0.295	0.087	0.918	0.403	0.318	0.101	0.932
			β_2 = -0.30	-0.308	0.248	0.062	0.933	-0.302	0.235	0.055	0.921
		1.5	β_1 = 0.40	0.378	0.304	0.093	0.937	0.382	0.352	0.124	0.937
			β_2 = -0.30	-0.299	0.233	0.054	0.934	-0.294	0.241	0.058	0.933
500	2	0.5	β_1 = 0.40	0.398	0.129	0.017	0.953	0.397	0.141	0.020	0.951
			β_2 = -0.30	-0.298	0.093	0.009	0.936	-0.303	0.099	0.010	0.947
		1.5	β_1 = 0.40	0.393	0.134	0.018	0.938	0.394	0.142	0.020	0.946
			β_2 = -0.30	-0.300	0.094	0.009	0.943	-0.302	0.098	0.010	0.936
	3	0.5	β_1 = 0.40	0.392	0.146	0.021	0.951	0.393	0.144	0.021	0.952
			β_2 = -0.30	-0.299	0.093	0.009	0.959	-0.304	0.102	0.010	0.942
		1.5	β_1 = 0.40	0.406	0.161	0.026	0.930	0.402	0.144	0.021	0.929
			β_2 = -0.30	-0.307	0.105	0.011	0.929	-0.301	0.099	0.010	0.950
1000	2	0.5	β_1 = 0.40	0.399	0.096	0.009	0.944	0.405	0.098	0.010	0.943
			β_2 = -0.30	-0.304	0.066	0.004	0.956	-0.301	0.066	0.004	0.946
		1.5	β_1 = 0.40	0.401	0.096	0.009	0.940	0.396	0.105	0.011	0.935
			β_2 = -0.30	-0.299	0.069	0.005	0.933	-0.302	0.069	0.005	0.938
	3	0.5	β_1 = 0.40	0.397	0.101	0.010	0.943	0.397	0.104	0.011	0.930
			β_2 = -0.30	-0.297	0.066	0.004	0.945	-0.302	0.068	0.005	0.943
		1.5	β_1 = 0.40	0.403	0.102	0.010	0.945	0.394	0.107	0.012	0.929
			β_2 = -0.30	-0.299	0.071	0.005	0.936	-0.303	0.073	0.005	0.937

3.3 Model Misspecification

In this section, an attempt has been made to examine the performance of the classical exponential proportional hazard (CEPH) model when survival times are generated following TPEPH model. To generate survival data, same covariates given in Section 3.2 were selected with $\beta_1 = 0.40$, $\beta_2 = 0.50$, $\lambda = 1.25$ and $\theta = 2.25$. A simulation study with 1000 trials has been performed for sample size 500 under no censoring and 10% censoring schemes. The simulated mean (SM), simulated standard error (SSE) and mean squared error (MSE) of estimators of parameters are given in Table 2 under both TPEPH and CEPH models.

Table 2: Comparison between TPEPH and CEPH models with sample size 500 and 1000 simulations

		TPEPH			**CEPH**		
Censoring	True values	SM	SSE	MSE	SM	SSE	MSE
0%	$\beta_1 = 0.40$	0.408	0.159	0.025	0.403	0.141	0.020
	$\beta_2 = 0.50$	0.497	0.093	0.018	0.505	0.093	0.020
	$\lambda = 1.25$	1.329	0.425	0.187	2.355	0.302	1.312
	$\theta = 2.25$	2.275	1.184	1.402	–	–	–
10%	$\beta_1 = 0.40$	0.407	0.159	0.025	0.394	0.146	0.021
	$\beta_2 = 0.50$	0.489	0.100	0.018	0.498	0.099	0.019
	$\lambda = 1.25$	1.157	0.380	0.153	2.131	0.282	0.856
	$\theta = 2.25$	2.311	1.134	1.289	–	–	–

It is observed from Table 2 that regression estimators obtained from both survival models are approximately unbiased. But the estimator of λ was found approximately unbiased under TPEPH model, whereas this estimator is highly biased in CEPH model. Note that the survival quantities such as survival function, hazard function, quantile times depend on the scale parameter λ. Therefore, these quantities computed under CEPH model may provide misleading survival information, if survival data are generated through TPEPH survival model.

4 Illustration

To illustrate the proposed Truncated Poisson Exponential Proportional Hazard model, a real data set was extracted from Bangladesh Demographic and Health Survey (BDHS), 2014. The BDHS is a part of the worldwide Demographic and Health Surveys (DHS) program, which is designed to collect data on basic national indicators of social progress including fertility, family planning, and maternal and child health.

To conduct the survival analysis, under-five child survival has been considered as an event. Only last child of a mother born preceding five years of the survey was considered into the analysis. The survival time of a child was considered to be censored, if he/she experienced death before celebrating the fifth birthday and the survival time was the age at death in years. On the other hand, an event was said to occur for a child if he/she survived upto the time of interview and the age in years at the time of interview is considered as his/her survival time. The survival information along with background characteristics is collected from 6855 children and 2.55% of survival times were found to be censored.

Based on previous studies on under five child survival (Chowdhury, 2013; Baqui et al., 1998; Uddin, 2009; Paul, 1990; Mondal et al., 2009; Dancer et al., 2008), following covariates have been selected: age of mother at first birth [Age below 20, Age 20-30 and Age above 30.], mother's education level [No, Primary, Secondary and Higher] and place of residence [Rural and Urban]. Both Truncated Poisson-Exponential proportional hazard (TPEPH) model and classical exponential proportional hazard (CEPH) model have been used to find the influence of selected covariates on survival times. For the purpose of estimation under TPEPH model, the likelihood function given in (2.8) was considered with survival and hazard functions given in (2.6) and (2.7), respectively. For CEPH model 'eha' package in RStudio has been used for the purpose of estimation. The results obtained from this analysis were reported in Table 3. This table represents estimated hazard ratios (HR) for selected covariates along with 95% confidence intervals (CI) and estimates of λ and θ with standard error (SE).

Table 3: Estimates of Hazard Ratio (HR) with 95% confidence interval (CI) and p-value; estimates of λ and θ along with standard error (SE) and p-value under TPEPH and CEPH models for child survival in Bangladesh

		TPEPH		**CEPH**	
Variables	Category	HR	95% CI	HR	95% CI
Age of mother at first birth	< 20	0.934**	(0.877, 0.991)	1.001	(0.942, 1.060)
	20-30	-	-	-	-
	> 30	1.436**	(1.031, 1.841)	0.936	(0.646, 1.226)
Mother's education level	No education	-	-	-	-
	Primary	1.670***	(1.523, 1.817)	1.096**	(1.012, 1.180)
	Secondary	1.470***	(1.349, 1.591)	1.116***	(1.037, 1.195)
	Higher	1.204***	(1.065 , 1.343)	1.187***	(1.068, 1.306)
Place of residence	Rural	-	-	-	-
	Urban	0.993	(0.939, 1.047)	0.976	(0.924, 1.028)
	Parameter	**Estimate**	**SE**	**Estimate**	**SE**
	λ	0.282***	0.013	0.398***	0.016
	θ	0.420***	0.036	-	-

* p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01

The parameter θ in zero truncated Poisson distribution was estimated as 0.420 with $p-value < 0.001$. It implies that on an average there were 1.225 [$= \hat{\theta}(1 - e^{-\hat{\theta}})^{-1}$] (Johnson and Kotz, 2005) causes responsible for the occurrence of the event of interest for a child. Since the number of causes is more than one, TPEPH model fitted the survival data better compared to CEPH model. It is clear from the Table 3 that all covariates except place of residence had significant influence on survival time. One can interpret the influence of these covariates using hazard ratio. The estimates of the scale parameter λ under TPEPH and CEPH models were 0.282 ($p - value < 0.001$) and 0.398 ($p - value < 0.001$), respectively.

It is found from Table 3 that age of mother at first birth was found to have significant association with child survival in TPEPH model but not in CEPH model. Both models reveal the fact that education level of mother had a positive impact on child survival. Place of residence of child was not found to be a potential factor under both models. Interpretation of significant hazard ratios under TPEPH model is provided as this model is appropriate for analyzing the given child survival data. The hazard rate of child survival was 6.6% lower if mother's age at first birth was less than 20 years compared to that of 20-30 years. On the other hand, if mother's age at first birth was higher than 30 years, hazard rate of child survival was 43.6% greater than a child of a mother of first birth age group 20-30 years. The hazard rate of child survival increased by 67.0%, 47.0% and 20.4% if education level of mother was primary, secondary and higher, respectively compared to a mother with no education.

5 Conclusion

In classical survival analysis, only one cause is considered responsible for the occurrence of an event of interest. But in practice, an event may occur from more than one causes, where lifetimes associated with each of the causes are not observable. This is because the occurrence of event due to specific cause may preclude the occurrence of event due to other causes. Hence, the minimum of lifetimes is observed.

In this paper, a survival model has been developed assuming that the number of causes associated with an individual follows a zero truncated Poisson distribution. It is also assumed that for a given number of causes, lifetimes corresponding to causes are independent and exponentially distributed. Therefore, the marginal distribution of minimum of lifetimes has a Truncated Poisson Exponential (TPE) distribution. For simplicity in interpretation of regression parameters, covariates have been introduced into survival model by following Cox proportional hazard model (Cox, 1972). To examine the performance TPEPH model, an extensive simulation has been conducted. It is found that the estimators of regression parameters are consistent and estimators of nuisance parameters in TPE distribution are approximately consistent. The classical exponential proportional hazard (CEPH) model has also been compared with TPEPH model through a simulation study. This simulation study reveals that though the performance of estimators of regression parameters in both models are similar, the scale parameter in CEPH model was found highly biased. It implies that the survival quantities computed from CEPH model will provide misleading survival information when an event of interest may occur from more than one causes. The TPEPH and CEPH models were illustrated with under-five child survival data extracted from BDHS 2014.

Acknowledgment

We would like to thank anonymous reviewer for valuable comments and suggestions that help a lot to improve the final version of the paper. We would also like to thank National Institute of Population Research and Training (NIPORT), Bangladesh for allowing us to use the BDHS 2014 data for illustration.

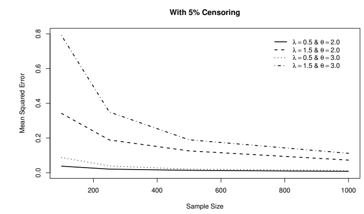
Appendix

Table A1: Simulated mean (SM), simulated standard Error (SSE), mean squared error (MSE) and 95% coverage probability (CP) of regression parameters in TPEPH model obtained from 1000 simulations with $\beta_1 = .40$ and $\beta_2 = -.30$

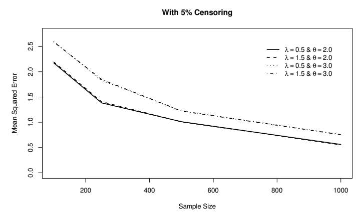
				10% Censoring				20% Censoring			
n	θ	λ	β	SM	SSE	MSE	CP	SM	SSE	MSE	CP
100	2	0.5	β_1	0.367	0.331	0.111	0.928	0.395	0.317	0.101	0.943
			β_2	-0.305	0.218	0.047	0.944	-0.301	0.222	0.049	0.943
		1.5	β_1	0.375	0.338	0.115	0.934	0.388	0.346	0.120	0.936
			β_2	-0.290	0.218	0.048	0.956	-0.313	0.236	0.056	0.932
	3	0.5	β_1	0.400	0.367	0.135	0.945	0.396	0.340	0.116	0.933
			β_2	-0.294	0.223	0.020	0.943	-0.302	0.246	0.061	0.930
		1.5	β_1	0.381	0.348	0.122	0.939	0.376	0.370	0.138	0.940
			β_2	-0.292	0.231	0.054	0.922	-0.292	0.243	0.059	0.935
500	2	0.5	β_1	0.398	0.144	0.021	0.945	0.391	0.154	0.024	0.933
			β_2	-0.302	0.099	0.010	0.959	-0.296	0.110	0.012	0.935
		1.5	β_1	0.389	0.146	0.022	0.931	0.397	0.152	0.023	0.937
			β_2	-0.294	0.096	0.009	0.951	-0.301	0.110	0.012	0.917
	3	0.5	β_1	0.394	0.142	0.020	0.944	0.393	0.158	0.025	0.946
			β_2	-0.306	0.102	0.010	0.936	-0.297	0.109	0.012	0.938
		1.5	β_1	0.395	0.155	0.024	0.910	0.389	0.162	0.026	0.932
			β_2	-0.307	0.106	0.011	0.933	-0.307	0.108	0.012	0.932
1000	2	0.5	β_1	0.396	0.105	0.011	0.935	0.394	0.114	0.013	0.924
			β_2	-0.301	0.068	0.005	0.951	-0.300	0.074	0.005	0.947
		1.5	β_1	0.397	0.107	0.011	0.938	0.391	0.106	0.011	0.940
			β_2	-0.300	0.069	0.005	0.954	-0.302	0.078	0.006	0.937
	3	0.5	β_1	0.399	0.107	0.011	0.929	0.399	0.107	0.011	0.936
			β_2	-0.305	0.071	0.005	0.938	-0.306	0.080	0.006	0.929
		1.5	β_1	0.392	0.124	0.016	0.899	0.402	0.120	0.014	0.927
			β_2	-0.330	0.074	0.006	0.925	-0.304	0.083	0.007	0.927

Table A2: Simulated mean (SM), simulated standard Error (SSE) , mean squared error (MSE) and 95% coverage probability (CP) of regression parameters in TPEPH model obtained from 1000 simulations with $\beta_1 = .40$ and $\beta_2 = -.30$

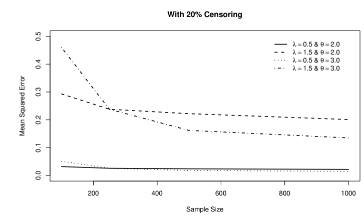
				30% Censoring				50% Censoring			
n	θ	λ	β	SM	SSE	MSE	CP	SM	SSE	MSE	CP
100	2	0.5	β_1	-	-	-	-	-	-	-	-
			β_2	-	-	-	-	-	-	-	-
		1.5	β_1	0.388	0.422	0.178	0.936	0.417	0.399	0.160	0.928
			β_2	-0.305	0.255	0.065	0.937	-0.304	0.307	0.094	0.918
	3	0.5	β_1	0.421	0.430	0.186	0.939	-	-	-	-
			β_2	-0.300	0.256	0.066	0.932	-	-	-	-
		1.5	β_1	0.363	0.437	0.192	0.925	0.426	0.505	0.256	0.934
			β_2	-0.295	0.259	0.067	0.938	-0.311	0.309	0.095	0.936
500	2	0.5	β_1	0.390	0.166	0.028	0.947	0.397	0.183	0.033	0.938
			β_2	-0.302	0.114	0.013	0.948	-0.299	0.126	0.016	0.935
		1.5	β_1	0.393	0.166	0.028	0.940	0.389	0.183	0.033	0.937
			β_2	-0.300	0.118	0.014	0.938	-0.305	0.129	0.016	0.944
	3	0.5	β_1	0.398	0.181	0.033	0.931	0.397	0.188	0.035	0.927
			β_2	-0.2971	0.116	0.013	0.937	-0.307	0.133	0.018	0.936
		1.5	β_1	0.391	0.176	0.031	0.926	0.410	0.189	0.036	0.941
			β_2	-0.304	0.114	0.013	0.931	-0.301	0.135	0.018	0.924
1000	2	0.5	β_1	0.396	0.119	0.014	0.932	0.388	0.140	0.020	0.927
			β_2	-0.296	0.081	0.007	0.927	-0.299	0.091	0.008	0.946
		1.5	β_1	0.404	0.115	0.013	0.933	0.390	0.132	0.017	0.930
			β_2	-0.300	0.078	0.006	0.940	-0.302	0.094	0.009	0.932
	3	0.5	β_1	0.407	0.113	0.013	0.941	0.401	0.140	0.019	0.928
			β_2	-0.296	0.077	0.006	0.956	-0.304	0.096	0.009	0.922
		1.5	β_1	0.397	0.121	0.015	0.931	0.397	0.146	0.021	0.928
			β_2	-0.301	0.083	0.007	0.928	-0.300	0.100	0.010	0.919



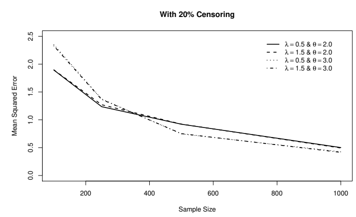
(a) MSE of $\hat{\lambda}$: 5% censoring



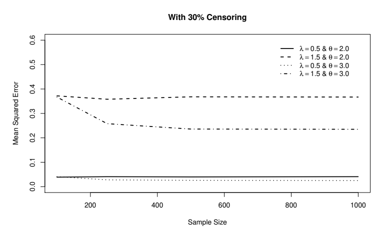
(b) MSE of $\hat{\theta}$: 5% censoring



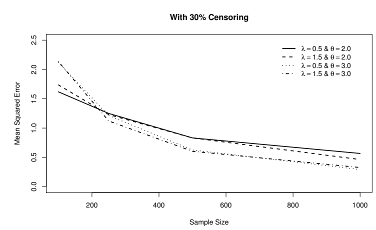
(c) MSE of $\hat{\lambda}$: 20% censoring



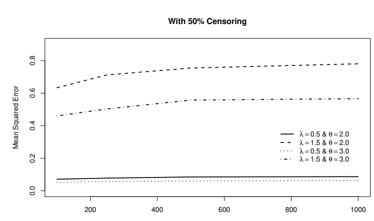
(d) MSE of $\hat{\theta}$: : 20% censoring



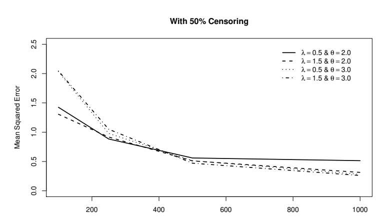
(e) MSE of $\hat{\lambda}$: 30% censoring



(f) MSE of $\hat{\theta}$: 30% censoring

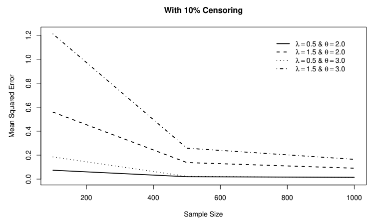


(g) MSE of $\hat{\lambda}$: 50% censoring

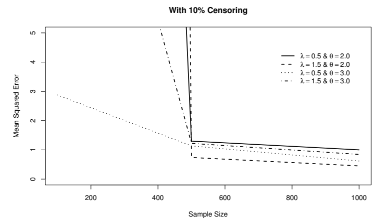


(h) MSE of $\hat{\theta}$: 50% censoring

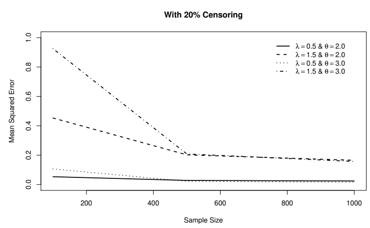
Figure A1: MSE($\hat{\lambda}$) and MSE($\hat{\theta}$) of Truncated Poisson Exponential survival model



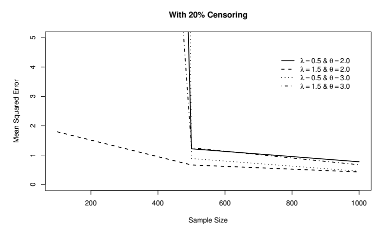
(a) MSE of $\hat{\lambda}$: 10% censoring



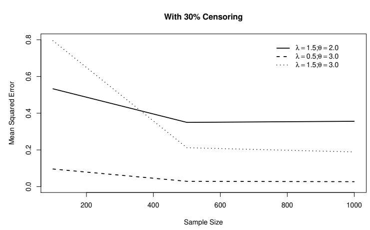
(b) MSE of $\hat{\theta}$:10% censoring



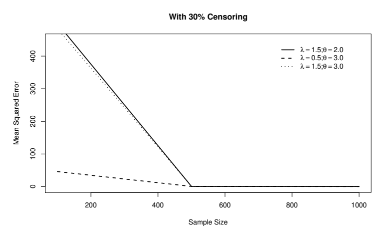
(c) MSE of $\hat{\lambda}$: 20% censoring



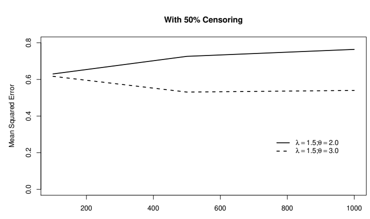
(d) MSE of $\hat{\theta}$: 20% censoring



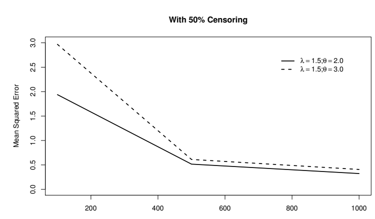
(e) MSE of $\hat{\lambda}$: 30% censoring



(f) MSE of $\hat{\theta}$:30% censoring



(g) MSE of $\hat{\lambda}$: 50% censoring



(h) MSE of $\hat{\theta}$:50% censoring

Figure A2: MSE($\hat{\lambda}$) and MSE($\hat{\theta}$) of Truncated Poisson Exponential Proportional Hazard model

References

Baqui, A. H., Black, R. E., Arifeen, S. E., Hill, K., Mitra, S. N., and Al Sabir, A. (1998). "Causes of childhood deaths in Bangladesh: Results of a nationwide verbal autopsy study", *Bulletin of the World Health Organization*, 76(2), 161-171.

Basu A.P. (1980), "Identifiability problems in the Theory of Competing and Complementary Risks", *Technical Report No. 97, Department of Statistics, University of Missouri-Columbia*.

Basu, A. P., and Ghosh, J. K. (1980). "Identifiability of distributions under competing risks and complementary risks model", *Communications in Statistics-Theory and Methods*, 9(14), 1515-1525.

Basu, A. P., and Klein, J. P. (1982). "Some recent results in competing risks theory", *Lecture Notes-Monograph Series*, 2, 216-229.

Bereta, E. M., Louzanda, F., and Franco, M. A. (2011). "The Poisson-Weibull distribution", *Adv. Applic. Statist*, 22(2), 107-118.

Chowdhury, A. H. (2013). "Determinants of under-five mortality in Bangladesh", *Open Journal of Statistics*, 3(03), 213-219.

Cox, D. R. (1972). "Regression models and life-tables", *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2), 187-202.

Dancer, D., Rammohan, A., and Smith, M. D. (2008). "Infant mortality and child nutrition in Bangladesh", *Health Economics*, 17(9), 1015-1035.

Johnson, N. L., Kemp, A. W., and Kotz, S. (2005). "Univariate discrete distributions, third edition". John Wiley & Sons. New York.

Klein, J. P. and Moeschberger, M. L. (2003). "Survival analysis: techniques for censored and truncated data, second edition". Springer-Verlag, New York.

Kuş, C. (2007). "A new lifetime distribution", *Computational Statistics & Data Analysis*, 51(9), 4497-4509.

Louzada-Neto, F., Cancho, V. G. and Barriga, G. D. (2011). "The Poisson–exponential distribution: a Bayesian approach", *Journal of Applied Statistics*, 38(6), 1239-1248.

Louzada, F., Cancho, V. G. and Barriga, G. D. (2012). "The Poisson-exponential regression model under different latent activation schemes", *Computational & Applied Mathematics*, 31(3), 617-632.

Mohammad, K.A., Fatima-Tuz-Zahura, M. and Bari, W. (2017). "Fine and Gray competing risk regression model to study the cause-specific under-five child mortality in Bangladesh", *BMC International Health and Human Rights*, 17:3.

Mondal, M. N. I., Hossain, M. K., and Ali, M. K. (2009). "Factors influencing infant and child mortality: A case study of Rajshahi District, Bangladesh", *Journal of Human Ecology*, 26(1), 31-39.

NIPORT, Mitra and Associates, and ICF (2016). "Bangladesh Demographic and Health Survey 2014", Technical report, National Institute of Population Research and Training (NIPORT), Mitra and Associates, and ICF International, Dhaka, Bangladesh, and Rockville, Maryland, USA: NIPORT, Mitra and Associates, and ICF International.

Paul, B. K. (1990). "Factors Affecting Infant Mortality in Rural Bangladesh: Results from a Retrospective Sample Survey 1", *Rural Sociology*, 55(4), 522-540.

Pintilie, M. (2007). "Analysing and interpreting competing risk data", *Statistics in Medicine*, 26(6), 1360-1367.

Ramos, P. L., Dey, D. K., Louzada, F., & Lachos, V. H. (2019). "An extended poisson family of life distribution: A unified approach in competitive and complementary risks", *Journal of Applied Statistics*, 1-17.

Uddin, M. (2009). "Child Mortality in a Developing Country: A Statistical Analysis", *Journal of Applied Quantitative Methods*, 4(3), 270-283.

Received: September 25, 2019

Accepted: November 4, 2019