Available at
http://pvamu.edu/aam
Appl. Appl. Math.
ISSN: 1932-9466
Vol. 12, Issue 2 (December 2017), pp. 753 – 765

Applications and Applied Mathematics:
An International Journal
(AAM)

Proportional Reversed Hazard Rate Models with Exponential Baseline

G. Barmalzan & [2]H. Saboori

Department of Statistics
University of Zabol
Sistan and Baluchestan, Iran
[1]Ghobad.barmalzan@gmail.com, [2]h.saboori@uoz.ac.ir.com

Received: August 6, 2017; Accepted: November 16, 2017

Abstract

The proportional hazard regression models have been used extensively in survival analysis to understand and exploit the relationship between survival time and covariates. For left censored survival times, reversed hazard rate functions are more appropriate. In this paper, we discuss a parametric proportional reversed hazard rates model using exponential baseline. The estimation for the parameters are discussed. We also assess the performance of the proposed procedure based on a large number of Monte Carlo simulations. Finally, we illustrate the proposed method using a real case example and then we show that it provides a good and better fit than the usual proportional hazards model.

Keywords: Covariate; Generalized Exponential Distribution; Monte Carlo Simulations; Weibull Distribution; Proportional Reversed Hazard Regression Models

MSC 2010 No.: 62N05, 62N86

1. Introduction

Even through hazard rate and reversed hazard rates share many similar aspects, reversed hazard rate functions are far less frequently used in applications. The reversed hazard rate function has attracted the attention of researchers only relatively recently (Finkelstein (2002)). Although initially introduced in actuarial research, the reversed hazard rate has mainly been applied in reliability engineering (Desai et al. (2011)) until now. Despite being similar to the hazard rate function, its typical behavior makes it suitable for assessing waiting times, hidden failures, inactivity times and the study of systems, including optimizing reliability and the probability of successful functioning (Block et al. (1998), Chandra and Roy (2001), Xie et al. (2002), and Poursaeed (2010)).

Reliability engineering, however, is not the only field where this tool has proved useful. Reversed hazard rate has also been employed for analyzing right-truncated and left-censored data. Kalbfleisch and Lawless (1991) and Gross and Huber-Carol (1992) used it in the field of medicine, and Townsend and Wenger (2004) found it helpful in modelling information processing capacity. Likewise, stochastic comparison of order statistics is another subject where this function has found a niche (Shaked and Shanthikumar (2006)). In this regard, Cheng and Zhu (1993) characterized the best strategy for allocating servers in a tandem system, and Finkelstein (2002) discussed possible applications to the ordering of random variables using the proportional reversed hazard rate. Finally, Kijima (1998) used the reversed hazard rate in the study of continuous time Markov Chains, and Gupta et al. (2004) and Razmkhah et al. (2012) showed how the Fisher information and Shannon entropy measures could be computed using the reversed hazard rate function. With all these applications, however, reversed hazard rate remains outside the mainstream of statistical literature.

Recently, Veres-Ferrer and Jose (2014) showed a close relationship that exists between the reversed hazard rate and elasticity, a concept broadly used in economics and physics, and used an example to illustrate how elasticity can provide statisticians with a more natural way of incorporating their knowledge, hypotheses and intuitions into certain inference processes. This will make it possible to obtain distributions that would be hardly imaginable or justifiable when expressed directly in terms of either the density or distribution functions.

In survival studies, covariates or explanatory variables are usually employed to represent heterogeneity in a population. The main objective in such situations is to understand and exploit the relationship between lifetime and covariates. Regression models are useful in such contexts to present the effect of covariates on lifetime. These models can be formulated in many ways and several types are in common use. Parametric regression models for lifetime involve specification for the distribution of a lifetime T given a vector of covariates Z. The most commonly used parametric model is the Weibull regression model, which satisfies the proportional relationship between reversed hazard rate functions of the lifetimes of two subjects. The maximum likelihood technique is usually employed to find estimates of the parameters of the model. For more properties and applications of parametric regression models, one should refer to Lawless (2003).

In survival studies, there are many occasions where lifetime data are left censored. For example, suppose T represents the lifetime of a cancer patient after diagnosis of the disease. If the doctor determine that the patient would not live more than k months, then the lifetime of this patient is left censored. On such occasions, a reversed hazard rate is more appropriate than a hazard rate to analyze lifetime data due to the fact that estimators of hazard rates are unstable when data are left censored. Barlow et al. (1963) introduced the reversed hazard rate of T as follows:

$$\tilde{r}(t) = \lim_{\Delta t \to 0} \frac{P(t - \Delta t \leq T | T \leq t)}{\Delta t}.$$

The proportional reversed hazard rates have been used extensively in reliability and survival analysis. For example, it has been used in various contexts such as the estimation of distribution function under left censoring (Lawless (2003)), stochastic ordering (Keilson and Sumitha (1982)), characterization of lifetime distributions (Block et al. (1998), Finkelstein (2002), and Nair et al. (2005)), studying aging behavior (Gupta et al. (1998), and Lai and Xie (2006)), repair and maintenance strategies (Marshall and Olkin (2007)), the mixed proportional hazards model (Horny (2009)), and stress hybrid hazards model (Tojeiro and Louzada (2012)). Sengupta and Nanda (2011) introduced the proportional reversed hazards rate model in a semiparametric setup. In the present work, we introduce a fully proportional reversed hazard rate models when baseline distribution of reversed hazard rates is exponential. A simulation study indicates that the proposed approach is performing well.

The rest of the paper is organized as follows. In Section 2, we introduce a parametric regression model using the generalized exponential distribution and then present the proposed model has the property that the reversed hazard rate for the lifetime of pair of subjects is proportional. In this section, the estimation of the parameters of the model is also discussed. Checking the proposal reversed hazard rate assumption is derived in Section 3. We present graphical test for specified model in Section 4. Simulation studies are conducted in Section 5 to assess the finite sample behavior of the estimators. The proposed model is applied to real life data in Section 6 to illustrate its utility. Finally, Section 7 provides the major conclusions of the study.

2. Statistical Model

Introducing covariates into the statistical model should lead to better prediction of lifetimes and improved understanding of the process that is being studied. In medical applications where the variable of interest might be the patient's survival time, obvious covariates to consider would include age, sex, severity of disease, smoking status, results of blood tests and other laboratory data, and so on.

There are several ways of introducing covariates into parametric models. We start from the well-known statistical model that incorporates covariates. This is the linear regression model, in which the value of a continuous dependent variable Y is related to a set of covariates z_i ($i = 1, \cdots, p$) as

$$y = \beta_0 + \beta_1 z_1 + \cdots + \beta_p z_p + \epsilon = \beta' \mathbf{z} + \epsilon,$$

where the constant β_0 has been absorbed into the vector of regression coefficients by taking $z_0 \equiv 1$. The residuals ϵ corresponding to different observations are assumed to be independent and identically distributed as $N(0, \sigma^2)$. Assuming that the covariates z_i are non-stochastic, we have $Y \sim N(\mu, \sigma^2)$, where $\mu = \mu(z) = \beta' \mathbf{z}$.

Thus, the covariates have entered the model by affecting the value of the parameter μ of the distribution of Y. This idea can be extended to other situations, involving distributions other than the normal. The main interest is in the extension to generalized exponential distribution and so we will develop regression models here for lifetime data. We will show the regression model that coincide for this distribution, in general lead to one types of models, namely, proportional reversed hazard rate models.

Generalized exponential distribution is one of the commonly used distributions in reliability, life-testing and survival analysis. A random variable T is said to have the generalized exponential distribution with shape parameter $\alpha > 0$ and scale parameter $\lambda > 0$ (denoted by $T \sim GE(\alpha, \lambda)$) if its cumulative distribution and probability density functions are given by respectively

$$F(t) = \left(1 - e^{-\lambda\, t}\right)^{\alpha} \quad , \quad f(t) = \alpha\, \lambda\, e^{-\lambda\, t} \left(1 - e^{-\lambda\, t}\right)^{\alpha - 1}, \qquad x > 0.$$

Gupta and Kundu (1999) showed the GE distribution admits a log-convex density for $0 < \alpha \leq 1$ and a log-concave density for $\alpha > 1$. Like the gamma and Weibull distributions, the reversed hazard rate function of the GE distribution is decreasing for $0 < \alpha < 1$, constant for $\alpha = 1$, and increasing for $\alpha > 1$. Indeed, the reversed hazard rate function of the Weibull distribution increases from zero to ∞ for $\alpha > 1$, and decreases from 1 to zero when $\alpha < 1$; while the reversed hazard rate function of $GE(\alpha, \lambda)$ increases from zero to $1/\lambda$ when $\alpha > 1$, and decreases from $1/\lambda$ to zero when $\alpha < 1$. On the other hand, the GE distribution has a bounded reversed hazard rate function which becomes it as a suitable model in situations that a regular maintenance program are held on the items in population; see Gupta and Kundu (1999). Due to nice survival functions, the GE distribution has also found the significant applications in the analysis of censored life data sets. For comprehensive discussions on the GE distribution, one may refer to Gupta and Kundu (2003, 2004), and Kundu et al. (2005). Interested readers may also refer to Gupta and Kundu (2007), and Nadarajah (2011) for reviews of the recent developments on the GE distribution.

Now, we introduce the effect of covariates $\mathbf{z}$ on the parameters of the cumulative distribution function, resulting in

$$F(t) = \left(1 - e^{-\lambda\, t}\right)^{\alpha(\mathbf{z})},$$

where the shape parameter α now depends on the covariates $\mathbf{z}$. This is the common form of the model, although it is possible to allow instead the scale parameter λ to depend on $\mathbf{z}$, or to let both parameters to depend on $\mathbf{z}$ simultaneously. Let $\alpha = e^{\beta' \mathbf{z}}$. Observe, however, that the function $\alpha = e^{\beta' \mathbf{z}}$ is positive, a restriction that is necessary here. The reversed hazard rate function is readily obtained

$$\tilde{r}(t|\, \mathbf{z}) = \frac{\lambda\, e^{-\lambda\, t}}{1 - e^{-\lambda\, t}}\, e^{\beta' \mathbf{z}} = \tilde{r}_0(t)\, g(\beta, \mathbf{z}),$$

where $\tilde{r}_0(t) = \frac{\lambda\, e^{-\lambda\, t}}{1 - e^{-\lambda\, t}}$ and $g(\beta, \mathbf{z}) = e^{\beta' \mathbf{z}}$. Expressing the model in this manner allows for the interpretation of the effect of the covariates in another way. It is as if the hazard function has been multiplied by the factor $e^{\beta' \mathbf{z}}$. For two units with different vectors of covariates $\mathbf{z_1}$ and $\mathbf{z_2}$, the ratio of the reversed hazard rate functions is

$$\frac{\tilde{r}(t|\, \mathbf{z_1})}{\tilde{r}(t|\, \mathbf{z_2})} = \frac{\frac{\lambda\, e^{-\lambda\, t}}{1 - e^{-\lambda\, t}}\, e^{\beta' \mathbf{z_1}}}{\frac{\lambda\, e^{-\lambda\, t}}{1 - e^{-\lambda\, t}}\, e^{\beta' \mathbf{z_2}}} = e^{\beta' (\mathbf{Z_1} - \mathbf{Z_2})},$$

which is a constant and independent of t. If, for example, one hazard is 45% of the other at time zero, then it is always 45% of the other at every time point. When the hazards of any two units have this property of being in constant ratio to each other, we have the so-called proportional hazards model. This model works with the basic premise that, at different levels of the covariate, the reversed hazard rates over time always remain proportional and are not function of time. This is a critical assumption of this case. This results in models that are closed under maxima.

Remark.

The generalized exponential distribution belongs to the proportional reversed hazard rate model provided scale parameter λ does not change with covariates but the shape parameter α changes with the covariates.

Because of the major importance of the generalized exponential distribution in applications, we describe here in detail the maximum likelihood fit of a generalized exponential regression model. The situation is to have failure times and left-censored values, with the censoring mechanism being non-informative. Then, the likelihood function based on a sample of n independent values is obtain

$$L(\beta) = \prod_{i=1}^{n} (f(t_i))^{\delta_i} \, (F(t_i))^{1-\delta_i}, \tag{1}$$

where $f(t)$ is the probability density function of the lifetime variable T, $F(t)$ is the survival function, and δ_i is an indicator variable taking on the value 0 when the value t_i is censored and 1 when it is an observed lifetime. Under the generalized exponential distribution assumption, the likelihood and log-likelihood function given in (1), respectively, is obtain as

$$L(\beta, \lambda) = \prod_{i=1}^{n} \left(e^{\beta' \mathbf{z_i}} \lambda e^{-\lambda\, t_i} \left(1 - e^{-\lambda\, t_i}\right)^{e^{\beta' \mathbf{z_i}} - 1} \right)^{\delta_i} \left(\left(1 - e^{-\lambda\, t_i}\right)^{e^{\beta' \mathbf{z_i}}} \right)^{1-\delta_i},$$

and

$$\begin{aligned} l(\beta, \lambda) &= \sum_{i=1}^{n} \delta_i \log \left(e^{\beta' \mathbf{z_i}} \lambda e^{-\lambda\, t_i} \left(1 - e^{-\lambda\, t_i}\right)^{e^{\beta' \mathbf{z_i}} - 1} \right) + \sum_{i=1}^{n} (1-\delta_i) \log \left(\left(1 - e^{-\lambda\, t_i}\right)^{e^{\beta' \mathbf{z_i}}} \right) \\ &= \sum_{i=1}^{n} \delta_i \left(\log \lambda - \lambda t_i + \beta' \mathbf{z_i} - \log\left(1 - e^{-\lambda\, t_i}\right) \right) + \sum_{i=1}^{n} e^{\beta' \mathbf{z_i}} \log \left(1 - e^{-\lambda\, t_i}\right). \end{aligned} \tag{2}$$

Now, we maximize (2) to estimate the parameters β, and λ by equating the partial derivatives with respect to each parameter to zero as

$$\frac{\partial l(\beta, \lambda)}{\partial \beta} = \sum_{i=1}^{n} \delta_i \mathbf{z_i} + \sum_{i=1}^{n} \log\left(1 - e^{-\lambda\, t_i}\right) \, \mathbf{z_i} e^{\beta' \mathbf{z_i}} = 0,$$

and

$$\frac{\partial l(\beta, \lambda)}{\partial \lambda} = \sum_{i=1}^{n} \delta_i \left(\frac{1}{\lambda} - t_i - \frac{t_i\, e^{-\lambda\, t_i}}{1 - e^{-\lambda\, t_i}} \right) + \sum_{i=1}^{n} e^{\beta' \mathbf{z_i}} \, \frac{t_i\, e^{-\lambda\, t_i}}{1 - e^{-\lambda\, t_i}} = 0.$$

Since there is no closed form solution available for this system of equations, we use numerical methods to estimate the parameters. The observed information matrix is given by

$$J(\beta, \lambda) = \begin{pmatrix} \frac{-\partial^2 l}{\partial \beta\, \partial \beta'} & \frac{-\partial^2 l}{\partial \beta\, \partial \lambda} \\ & \\ \frac{-\partial^2 l}{\partial \lambda\, \partial \beta} & \frac{-\partial^2 l}{\partial \lambda^2} \end{pmatrix},$$

where we have

$$\frac{\partial^2 l}{\partial \beta\, \partial \beta'} = \sum_{i=1}^{n} \log\left(1 - e^{-\lambda\, t_i}\right) \, \mathbf{z_i} \mathbf{z_i'} e^{\beta' \mathbf{z_i}};$$

$$\frac{\partial^2 l}{\partial\beta\,\partial\lambda} = \sum_{i=1}^{n} \mathbf{z_i} e^{\beta'\mathbf{z_i}} \frac{t_i\, e^{-\lambda\, t_i}}{1-e^{-\lambda\, t_i}};$$

$$\frac{\partial^2 l}{\partial\lambda^2} = \sum_{i=1}^{n} \delta_i \left(-\frac{1}{\lambda^2} + \frac{t_i\, e^{-\lambda\, t_i}}{(1-e^{-\lambda\, t_i})^2} \right) - \sum_{i=1}^{n} \mathbf{z_i} e^{\beta'\mathbf{z_i}} \frac{t_i\, e^{-\lambda\, t_i}}{(1-e^{-\lambda\, t_i})^2}.$$

Note that the above information matrix is of order $(p+2)\times(p+2)$. Under the standard regularity conditions, the vector of estimates $(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\lambda}})$ is asymptotically $(p+2)$-variate normal with mean vector $(\boldsymbol{\beta}, \boldsymbol{\lambda})$ and dispersion matrix I^{-1}, where I is the Fisher information matrix obtained from J by taking the expected values of each entry.

3. Checking the proportional reversed hazards assumption

If the proportional reversed hazards assumption is not correct for the data, then it will not be meaningful to fit the regression model based on this assumption. The main question which may be arise here is how we can check whether this assumption is appropriate for the given data. The answer of this question is stated in this section.

Let $\tilde{H}(t) = \int_t^b \tilde{r}(x)\,dx = -\log F(t)$ where $b = \sup\{x : F(x) < 1)\}$ and $t < b$. The proportional reversed hazard rate function $\tilde{r}(t|\,\mathbf{z}) = \tilde{r}_0(t)\, e^{\beta'\mathbf{z}}$ gives the distribution function to be $F(t|\mathbf{z}) = \exp\left\{\tilde{H}_0(t)e^{\beta'\mathbf{z}}\right\}$. Then,

$$\log(-\log F(t|\mathbf{z})) - \log \tilde{H}_0(t) = \beta'\mathbf{z}$$

which means that the curve $\log(-\log F(t|\mathbf{z}))$, for any $\mathbf{z}$ is parallel to $\log \tilde{H}_0(t)$ in time. This observation suggests a simple way of checking the proportional reversed hazards assumption as follows:

(i) compute Kaplan-Meier estimates of the cumulative function $\hat{F}(t|\mathbf{z})$ for select $\mathbf{z}$;
(ii) plot $\log(-\log \hat{F}(t|\mathbf{z}))$ against $\log \tilde{H}_0(t)$ for those $\mathbf{z}$.

If all the lines for those $\mathbf{z}$ are indeed parallel to each other, then we may conclude that the assumption of proportional reversed hazards is reasonable for the data. This idea applies to any proportional reversed hazards model, but does not show us which distribution is the appropriate one if we are going to carry out a parametric regression.

4. Graphical test for specified model

Since we don't assume any particular functional form for the cumulative distribution or reversed hazard rate functions therefore we use non-parametric estimators. This freedom from assumptions is often very desirable. The main question which arises here is how we can check that a specific function F is a reasonable model assumption for the data? One simple way is to obtain the Kaplan-Meier estimate $\hat{F}(t)$ and plot it against a suitable function of t, in order to see if it takes nearly the same shape as the assumed function $F(t)$.

For the generalized exponential distribution, we know that

$$F(t|\mathbf{z}) = \left(1 - e^{-\lambda\, t}\right)^{\beta'\mathbf{z}} = \exp\left\{\beta'\mathbf{z}\log\left(1 - e^{-\lambda\, t}\right)\right\},$$

and then,

$$\log(-\log F(t|\mathbf{z})) = -\log\left(1 - e^{-\lambda\, t}\right) + \beta'\mathbf{z}.$$

This means that a scatter plot of the values of $\log(-\log \hat{F}(t|\mathbf{z}))$ against $-\log\left(1 - e^{-\hat{\lambda}\, t}\right)$ should give a straight line if the assumption of a generalized exponential distribution is correct where $\hat{\lambda}$ is maximum likelihood estimator of λ. See Figure 2.

5. Simulation Study

In this section, we carried out an extensive Monte Carlo simulation study for evaluating the performance of the proposed estimation method. We considered four sample sizes: $n = 50$, $n = 100$ and $n = 250$. To introduce the covariate effect, we divided the total sample size into four groups, where group j were assigned a covariate value of j for $j = 1, 2, 3, 4$. The comparison is based on mean square error (MSE) for each parameter. The following steps has to be followed:

1. Specify the sample size n and the values of the parameters β and λ.
2. Generate random samples with size $n_1, \cdots, n_4$, respectively, from $GE(\exp(\beta z_i), \lambda), i = 1, 2, 3, 4$, such that $n_1 + \cdots + n_4 = n$.
3. Randomly, twenty percent of the total sample are left censored.
4. Calculate the MLE of the two parameters.
5. Repeat steps $2 - 3$, for N iterations.
6. Calculate the mean square error (MSE) for each parameter.

In this simulation, the MSEs were computed by generating $N = 250$ replications of samples of size $n = 50, 100, 250$ from the GE distribution with various choices of the parameters values (β, λ) (see Table 1). The assessment based on simulation study is that the MSEs for each parameter decreases with increasing the sample size.

In many applications there is qualitative information about the reversed hazard rate shape, which can help with selecting a particular model. In this context, a device called the total time on test (TTT) plot (Aarset, 1987) is useful. The TTT plot is obtained by plotting

$$T_n(\frac{r}{n}) = \frac{\sum_{i=1}^{r} t_{i:n} + (n-r)t_{r:n}}{\sum_{i=1}^{n} t_{i:n}}, \tag{3}$$

against $r = 1, ..., n$ where $t_{i:n},\ (i = 1, ..., n)$ are the order statistics of the sample. It is a straight diagonal for constant reversed hazard rates, it is convex for decreasing reversed hazard rates and concave for increasing reversed hazard rates. It is first convex and then concave if the reversed

Table 1. Maximum likelihood estimate and MSE (in parentheses) of parameter.

		$n=50$		$n=100$		$n=250$	
β	λ	$\hat{\beta}$	$\hat{\lambda}$	$\hat{\beta}$	$\hat{\lambda}$	$\hat{\beta}$	$\hat{\lambda}$
0.5	0.5	0.5070 (0.0257)	0.5276 (0.0080)	0.4969 (0.0134)	0.5133 (0.0036)	0.5156 (0.0046)	0.5181 (0.0013)
0.5	1.5	0.5180 (0.0239)	1.5869 (0.0576)	0.4981 (0.0112)	1.5313 (0.0230)	0.5034 (0.0040)	1.5239 (0.0116)
1	0.5	1.0406 (0.2831)	0.5616 (0.2760)	1.0253 (0.0172)	0.5162 (0.0026)	1.0008 (0.0047)	0.5008 (0.0007)
1	1.5	1.0021 (0.0308)	0.1544 (0.0374)	1.0114 (0.01562)	1.5214 (0.01922)	1.0022 (0.0071)	1.5113 (0.0088)
1.5	0.5	6.0599 (347.4700)	1.1377 (4.8254)	1.7114 (4.4645)	0.5595 (0.3127)	1.5049 (0.0090)	0.5028 (0.0007)
1.5	1.5	1.5078 (0.0698)	1.5736 (0.0614)	1.5325 (0.0300)	1.5239 (0.0230)	1.5092 (0.0119)	1.5195 (0.0089)

hazard rate is bathtub-shaped. It is first concave and then convex if the reversed hazard rate is upside-down bathtub.

In Figure 1, the right plot is the TTT-scaled plot of the generated sample from $GE(\exp(\beta z_i), \lambda), i = 1, 2, 3, 4$, which shows that the reversed hazard rates of data is increasing. The left plot is the Kaplan-Meier estimate of empirical CDF of the sample data.

In Figure 2, since all the lines for those z_i are indeed parallel to each other, then we may conclude that the assumption of proportional reversed hazards is reasonable for the data, because the curve $\log(-\log F(t|z_i))$, for any $z_i, i = 1, 2, 3, 4$ is parallel to $\log \tilde{H}_0(t)$ in time.

6. Real Data

In this section, we consider an extract of left censored data from an Australian twin study (Duffy et al. (1990)). The data consist of information on the age of appendectomy of monozygotic and dizygotic twins. There are observations with missing age at onset and therefore the data are left censored. The covariate, namely, Zygocity, has values from 1 to 6. This data set consists of 7616 observations of which 1718 are left censored. We use this data for generalized exponential distribution to illustrate the utility of the parametric reverse hazard rate model. We use the simplex method to estimate the parameters. Since the parameter values are unknown and to avoid the effect of inappropriate initial values, we consider different initial values and choose the estimates which have maximum likelihood. Estimates of the parameters are $\hat{\beta} = 0.2052$ and $\hat{\lambda} = 0.0523$. From empirical Fisher information matrix the standard error of parameters β and λ are, respectively, $6.0231e - 05$ and $7.6977e - 06$. The 95 percent confidence interval for β is $(0.2050, 0.2053)$ indicates that the regression coefficient corresponding to Zygocity is significantly different than zero; that is, the effect of Zygocity is not negligible. This conclusion is also verified through the likelihood ratio test statistic value 2430.548 having a P-Value of 0.

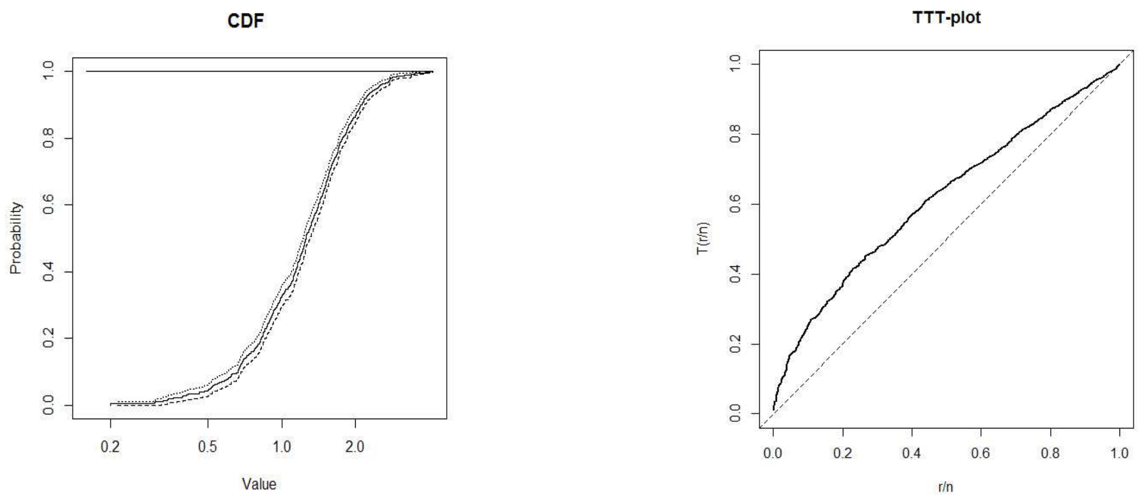

Figure 1. Left plot is the Kaplan-Meier estimation of CDF and right plot is the TTT-Scaled plot for simulation data.

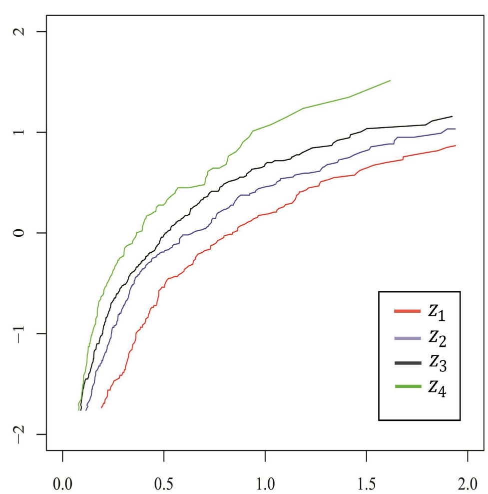

Figure 2. Checking the proportional reversed hazards assumption for simulation data.

The empirical Fisher information matrix of real data is:

$$J(\beta, \lambda) = \begin{pmatrix} 79013.09 & -422516.2 \\ -422516.2 & 4904695.2 \end{pmatrix}.$$

In the Figure 3 the left plot is the Box Plot of age on onset in six group of covariate variable (zygocity) and the right plot is the Kaplan-Meier estimation of CDF for real data. Figure 4 presents the TTT-scaled plot of real data that is shown the reversed hazard rates of real data is increasing. Also, Figure 5 is shown in the levels 2 and 6 of covariates, data can be applied in the proportional reversed hazards assumption, and the other level of data are not applied to this assumption.

Next, we use a Cox-Snell residual plot to assess the goodness of fit. The Cox-Snell residual is

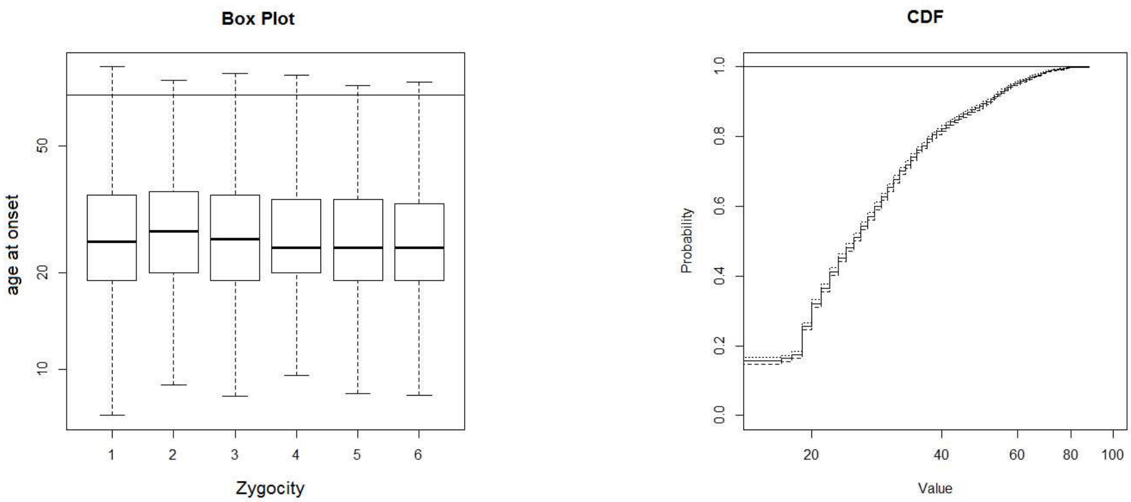

Figure 3. Left plot is the Box Plot of six group of covariate and right plot is the Kaplan-Meier estimation of CDF for real data.

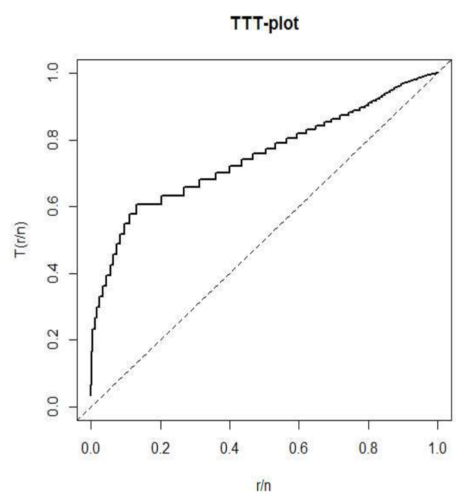

Figure 4. The TTT-Sacled plot for real data.

defined by

$$r_i = -\log(F(t_i)) = \exp(\beta \mathbf{z}_i) * \log(1 - e^{-\lambda t_i}). \tag{4}$$

If the model fits the data, then the residuals should have a standard exponential distribution, so that a hazard plot of residuals versus the Nelson-Aalen estimator of the cumulative hazard of the residuals will be a straight line with slope one. A plot of Cox-Snell residuals against the Nelson-Aalen estimates of the cumulative hazard rate of residuals is given in Figure 6, which shows that the fit is reasonably good.

In the following table, for the above real data, we have compared generalized exponential distribution (as a proportional reversed hazard rate model) with Weibull distribution (as a proportional reversed hazard rate model) and by using Akaike information criterion (AIC) and Bayesian in-

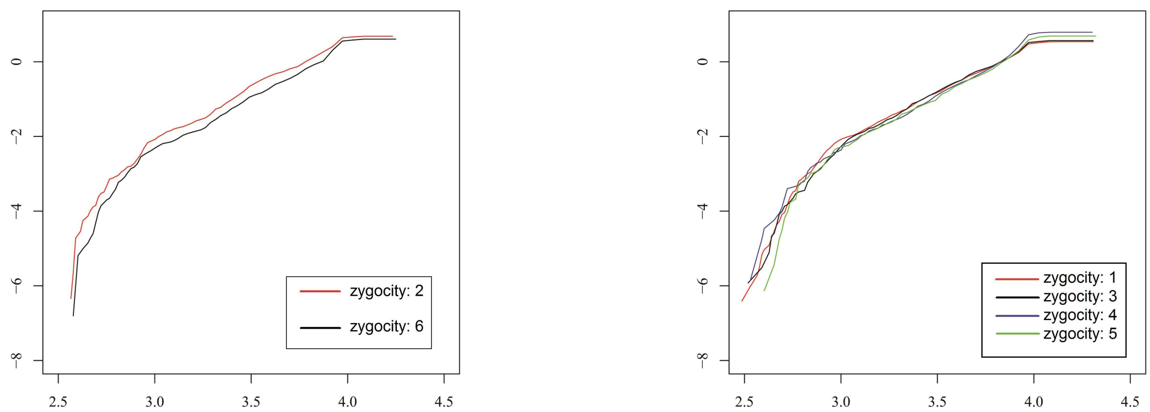

Figure 5. Checking the proportional reversed hazards assumption for real data.

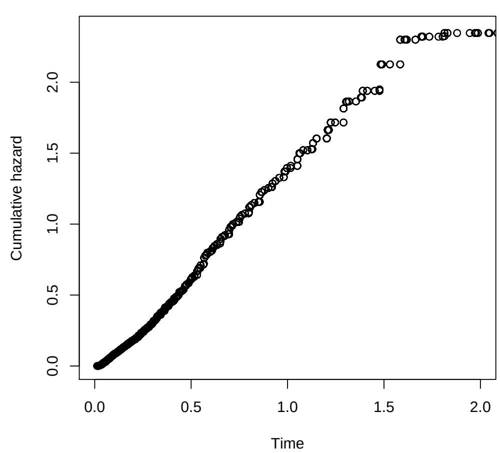

Figure 6. Plot of cumulative reversed hazard rates of Cox-Snell residuals versus residuals.

formation criterion (BIC), we show that generalized exponential distribution provides a good and better fit than the Weibull distribution.

7. Conclusion and Remarks

In this paper, we discuss a parametric proportional reversed hazard rates model using exponential baseline. The estimation and construction of confidence intervals for the parameters are discussed. We also assess the performance of the proposed procedure based on a large number of Monte Carlo simulations. Finally, we illustrate the proposed method using a real case example and then we show that it provides a good and better fit than the usual proportional hazards model.

Table 2. Maximum likelihood estimate, MSE (in parentheses) of parameters, -2ℓ, AIC and BIC for generalized exponential and Weibull distribution.

Model	$\hat{\lambda}$	$\hat{\beta}$	$\hat{\eta}$	-2ℓ	AIC	BIC
GE	0.0523	0.2052	-	54464.48	54468.48	54482.36
	(7.70e-06)	(6.02e-05)	-	-	-	-
Weibull	-	1.2145	0.4269	71086.96	71090.96	71104.84
		(0.0001)	(1.42e-05)	-	-	-

Acknowledgement:

We thank both referees and the Editor in Chief Professor Aliakbar Montazer Haghighi, for their valuable suggestions and constructive comments which helped us in revising the paper in the present form.

REFERENCES

Aarset, M.V. (1987). How to identify bathtub hazard rate, IEEE Transactions on Reliability, Vol. 36, pp. 106–108.

Barlow, R.E., Marshall, A.W. and Proschan, P. (1963). Properties of probability distributions with monotone hazard rate, Annals of Mathematical Statistics, Vol. 34, pp. 375–389.

Block, H.W., Savits, T.H. and Singh, H. (1998). The reversed hazard rate function, Probability in the Engineering and Informational Sciences, Vol. 12, pp. 69–90.

Chandra, N.K. and Roy, D. (2001). Some results on reverse hazard rate, Probability in the Engineering and Informational Sciences, Vol. 15, pp. 95–102.

Cheng, D.W. and Zhu, Y. (1993). Optimal order of servers in a tandem queue with general blocking, Queueing System, Vol. 14, pp. 427–437.

Desai, D., Mariappan, V. and Sakhardande, M. (2011). Nature of reversed hazard rate: an investigation, International Journal of Performability Engineering, Vol. 7, pp. 165–171.

Duffy, D. and Martin, N.G. (1990). Mathews, Appendectomy in Australian twins, American Journal of Human Genetics, Vol. 47, pp. 590–592.

Finkelstein, M.S. (2002). On the reversed hazard rate, Reliability Engineering and System Safety, Vol. 78, pp. 71–75.

Gross, S.T. and Huber-Carol, C. (1992). Regression models for truncated survival data, Scandinavian Journal of Statistics, Vol. 19, pp. 193–213.

Gupta, R.C., Gupta, P.L., and Gupta, R.D. (1998). Modeling failure time data by Lehman alternatives, Communications in Statistics, Vol. 27, pp. 887–904.

Gupta, R.D. and Kundu, D. (1999). Generalized exponential distributions, Australian & New Zealand Journal of Statistics, Vol. 41, pp. 173–188.

Gupta, R.D. and Kundu, D. (2003). Discriminating between the Weibull and the GE distributions, Computational statistics & data analysis, Vol. 43, pp. 179–196.

Gupta, R.D. and Kundu, D. (2004). Discriminating between gamma and generalized exponential distributions, Statistical Computation & Simulation, Vol. 74, pp. 107–121.

Gupta, R.D. and Kundu, D. (2007). Generalized exponential distribution: existing methods and some recent developments, Journal of Statistical Planning and Inference, Vol. 37, pp. 3537–3547.

Horny, G. (2009). Inference in mixed proportional hazard models with K random effects, Statistical Papers, Vol. 50, pp. 481–499.

Kalbfleisch, J.D. and Lawless, J.F. (1991). Regression models for right truncated data with applications to AIDS incubation times and reporting lags, Statistica Sinica, Vol. 1, pp. 19–32.

Kijima, M. (1998). Hazard rate and reversed hazard rate monotonicities in continuous time Markov chains, Journal of Applied Probability, Vol. 35, pp. 545–556.

Keilson, J. and Sumitha, U. (1982). Uniform stochastic ordering and related inequalities, Canadian Journal of Statistics, Vol. 15, pp. 63–69.

Kundu, D., Gupta, R.D and A. Manglick, A. (2005). Discriminating between the log-normal and generalized exponential distributions, Journal of Statistical Planning and Inference, Vol. 127, pp. 213–227.

Lawless, J.F. (2003). *Statistical models and methods for lifetime Data*, John Wiley & Sons, New York, NY, USA.

Li, X. and Xu, M. (2008). Reversed hazard rate order of equilibrium distributions and a related aging notion, Statistical Papers, Vol. 49, pp. 749–767.

Marshall, A.W. and Olkin, I. (2007). *Life Distributions*, Springer, New York.

Nadarajah, S. (2011). The exponentiated exponential distribution: a survey, Advances in Statistical Analysis, Vol. 95, pp. 219–251.

Nair, N.U., Sankaran, P.G. and Asha, G. (2005). Characterizations of distributions using reliability concepts, Journal of Applied Statistical Science, Vol. 14, pp. 237–241.

Poursaeed, M.H. (2010). A note on the mean past and the mean residual life of a $(n-k+1)$-out-of-n system under multi monitoring, Statistical Papers, Vol. 51, pp. 409–419.

Razmkhah, M., Morabbi, H. and Ahmadi, J. (2012). Comparing two sampling schemes based on entropy of record statistics, Statistical Papers, Vol. 53, pp. 95–106.

Shaked, M. and Shanthikumar, J.G. (2007). *Stochastic Orders.* Springer, New York.

Sengupta, D. and Nanda, A.K. (2011). The proportional reversed hazards regression model, Statistical Theory and Applications, Vol. 18, pp. 461–476.

Tojeiro, C.A.V. and Louzada, F. (2012). A general threshold stress hybrid hazard model for lifetime data, Statistical Papers, Vol. 53, pp. 833–848.

Townsend J.T. and Wenger M.J. (2004). A theory of interactive parallel processing: new capacity measures and predictions for a response time inequality series, Psychological Review, Vol. 111, pp. 1003–1035.

Veres-Ferrer, E.J. and Jose, M.P. (2014). On the relationship between the reversed hazard rate and elasticity, Statistical Papers, Vol. 55, pp. 275–284.

Xie, M., Gaudoin, O. and Bracquemond, C. (2002). Redefining failure rate function for discrete distributions, International Journal of Reliability, Quality and Safety Engineering, Vol. 9, pp. 275–285.