

Regression Modeling of Competing Risks Survival Data in the Presence of Covariates Based on a Generalized Weibull Distribution: A Simulation Study

Soraya Moamer

Department of Biostatistics, Faculty of Paramedical Sciences
Shahid Beheshti University of Medical Sciences, Tehran, Iran
moamer.s@sbmu.ac.ir

Ahmad Reza Baghestani

Physiotherapy Research Center, Department of Biostatistics, Faculty of Paramedical Sciences
Shahid Beheshti University of Medical Sciences, Tehran, Iran
baghestani@sbmu.ac.ir

Mohamad Amin Pourhoseingholi

Gastroenterology and Liver Diseases Research Center, Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran
amin_phg@yahoo.com

Abstract

In survival analysis or medical studies each person can be exposed to more than one type of outcomes which occurrence of one of them prevents the other outcomes' occurrence; this situation is called the competing risks. Assessing the effect of covariates on the survival time (or failure time) is one of the purposes in competing risks analysis. In this paper, we study a competing risks model in the presence of covariates when the causes of failures follow generalized Weibull distributions. Covariates are entered to the model through the scale parameter of this distribution. Also in this study the competing risks are considered to be independent. Parameter estimation has been done by the maximum likelihood approach, in a real data set and a simulation study has shown the advantages of proposed model.

Keywords: Competing risks models- Generalized Weibull Distribution- Covariates - Survival analysis

1. Introduction

In many survival studies, patients can be exposed to more than one outcome. For example, if we are interested in the analysis of data until death occurrence due to heart failure (the outcome of interest), some patients in the study may die due to causes other than heart failure (the competing event). These different causes of death are called competing risks (see Putter et al. (2007) and Pintilie (2007)). Indeed, the competing events are those which prevent the observation of main outcome or change the probability of its occurrence. The main feature of competing risks data is the presence of failure type according to event type in addition to failure time or survival time (see Porta Bleda et al. (2007)).

Modeling competing risks survival data can be carried out using a semi-parametric, non-parametric or parametric survival models (see Fine and Gray (1999)). The parametric models are studied assuming that the competing risks follow different lifetime distributions such as exponential, Lognormal and Weibull (see Sarhan (2007), Cox

(1959), Pascual (2010), Yáñez et al. (2014), Iskandar and Gondokaryono (2016)). One of the advantages of using the parametric approaches rather non-parametric and semi-parametric approaches is as follows: when the parametric model has been chosen correctly, it is possible to predict the event occurrence probability in future and have a clear picture of survival time and hazard function. Also as the survival pattern follows a special parametric model, the acquired estimates are more accurate than non- or semi-parametric approaches (see Jeong (2006)).

One of the distributions widely used in cancer research is the Weibull distribution (see Martinez et al. (2013)). However, only monotonically increasing and decreasing hazard functions can be generated from the classic two-parameter Weibull distribution. Mudholkar et al. (1996) and Mudholkar and Srivastava (1993) introduced the generalized Weibull model which includes an additional form parameter. Probability density, survival, and hazard functions of this distribution have a closed form and it is flexible so it can produce different forms of hazards such as bathtub and unimodal hazards.

When the failure pattern is analyzed considering a special cause, the researchers may care the assessment of covariate effects on failure probability of a special cause. Gray (1988) introduced a non-parametric inference for analyzing the competing risks data also (see Prentice et al. (1978)). Fine and Gray used the Cox proportional hazard model for competing risks data and presented inferences for evaluating the treatment influence and other covariates (see Kleinbaum and Klein (1996)). Also many other studies used the Fine and Gray model to evaluate the influence of covariates on the survival time (see Baghestani et al. (2014) and Lau et al. (2009)).

Wahed et al. (2009) used the generalized Weibull model introduced by Mudholkar et al. for competing risks data of breast cancer, but their model just included the main parameters but not the covariates. Alwase et al. (2009) and Sarhan et al. (2013) used generalized Weibull model in order to analyze a set of real data but their analyses were based on main parameters of generalized Weibull model and they did not enter the covariates into the model. However, as we observed in the literature, the generalized Weibull distribution has not been considered for modeling the competing risks data in presence of covariates.

The aim of this paper is to study a competing risks data in presence of covariates using generalized Weibull distribution. The paper is organized as follows: in section 2 we will explain the formulation of the likelihood functions in presence of covariates based on the generalized Weibull distribution. Then in section 3, based on a simulation study, we will assess the accuracy of estimates acquired from generalized Weibull model. In section 4 we use a competing risk real data set (medical data set) with covariates to test the generalized Weibull distribution against the Weibull distribution for fitting the data in this practical situation. Section 5 presents discussion and conclusions.

2. Model Formulation

Generalized Weibull distribution including third parameter γ , in addition to two parameters of (α, β) , possesses enough flexibility and can cover different forms of hazard function. This is an important characteristic of generalized Weibull distribution that

differs from the two-parameter Weibull distribution. Assumed that the random variable T has the generalized Weibull distribution, then the probability density function of this random variable with three parameters (α, β, γ) is as follows:

$$f(t) = \alpha\beta\gamma t^{\beta-1} e^{-\alpha t^\beta} (1 - e^{-\alpha t^\beta})^{\gamma-1} \quad \alpha > 0, \beta > 0, \gamma > 0, t > 0 \quad (1)$$

Here, β and γ are the shape parameters, while α is the scale parameter of generalized weibull distribution (GWD).

Weibull distribution can be considered as a special case of this distribution, because when $\gamma=1$, this distribution is changed to be Weibull distribution with two parameters (α, β) . Also generalized exponential distribution is a special case of this distribution when $\beta=1$.

In analyzing the competing risks data for each person there exists one type of failure (type of event:) in addition to failure time(survival time). The failure time (T) is assumed to be a continuous and positive random variable, while the failure cause (k) takes values in the finite set ($k \geq 2$) $j=1, \dots, k$. Also we assume that there is just one failure time and just one cause of failure, and competing risks are independent. Here the survival time for each person is defined as follows (see Gray (1988) and Beyersmann et al.(2011)):

$$T_i = \min(T_{1i}, T_{2i}, \dots, T_{ki}) \quad i=1, 2, \dots, n \quad (2)$$

We assume that survival time of each competing risks has generalized Weibull distribution. The probability density function for each of the competing risks (the cause of failure of j) is defined as follows:

$$f_j(t) = \alpha_j \beta_j \gamma_j t^{\beta_j-1} e^{-\alpha_j t^{\beta_j}} (1 - e^{-\alpha_j t^{\beta_j}})^{\gamma_j-1} \quad (3)$$

The hazard function is

$$h_j(t) = \frac{\alpha_j \beta_j \gamma_j t^{\beta_j-1} e^{-\alpha_j t^{\beta_j}} (1 - e^{-\alpha_j t^{\beta_j}})^{\gamma_j-1}}{1 - (1 - e^{-\alpha_j t^{\beta_j}})^{\gamma_j}} \quad (4)$$

and the survival function is

$$S_j(t) = 1 - \left(1 - e^{-\alpha_j t^{\beta_j}}\right)^{\gamma_j} \quad (5)$$

Which in the above equations $\alpha_j > 0, \beta_j > 0, \gamma_j > 0, t_j \geq 0$ and $j=1, 2, \dots, k(k \geq 2)$.

In order to assess the covariate effects on the survival time, we define the scale parameter α_j in the form of a linear combination of covariate as follows:

$$\alpha_j = \exp(\omega_{0j} + \omega_{1j}x_{1j} + \omega_{2j}x_{2j} + \dots + \omega_{lj}x_{lj}) \quad j = 1, 2, \dots, k \quad (6)$$

Thus the parameters vector is $\emptyset = (\emptyset_1, \emptyset_2, \dots, \emptyset_k)^T$, $\emptyset_j = (\omega_{0j}, \omega_{1j}, \dots, \omega_{lj}, \beta_j, \gamma_j)$ which is the vector of generalized Weibull distribution parameters for j cause of failure.

Preserving the totality of the subject matter, we can estimate the regression coefficients for the two causes of failure ($j=1, 2$) using likelihood function as follows, see Jeong and Fine (2006):

$$L(\emptyset|t) = \prod_{i=1}^n \{S_2(t_i, \emptyset_2) f_1(t_i, \emptyset_1)\}^{\delta_{i1}} \{S_1(t_i, \emptyset_1) f_2(t_i, \emptyset_2)\}^{\delta_{i2}} \\ \times \{S_1(t_i, \emptyset_1) S_2(t_i, \emptyset_2)\}^{1-\delta_{i1}-\delta_{i2}} = \\ \{f_1(t_i, \emptyset_1)\}^{\delta_{i1}} \{f_2(t_i, \emptyset_2)\}^{\delta_{i2}} \{S_1(t_i, \emptyset_1)\}^{1-\delta_{i1}} \{S_2(t_i, \emptyset_2)\}^{1-\delta_{i2}} \quad (7)$$

Where $\emptyset_1 = (\alpha_1, \beta_1, \gamma_1)$ and $\emptyset_2 = (\alpha_2, \beta_2, \gamma_2)$, and the vector of $(\delta_{i1}, \delta_{i2})$ considers the values (0,1), (1,0), and (0,0) for the observations of the cause of first failure, the cause of second failure and the observation which does not experience no event and are censored respectively.

When the survival time of each cause follows the generalized Weibull distribution, the likelihood function is defined as follows:

$$L(\emptyset|t) = \prod_{i=1}^n \left\{ \left[1 - \left(1 - e^{-\alpha_2 t_i^{\beta_2}} \right)^{\gamma_2} \right] \left[\alpha_1 \beta_1 \gamma_1 t_i^{\beta_1-1} e^{-\alpha_1 t_i^{\beta_1}} (1 - e^{-\alpha_1 t_i^{\beta_1}})^{\gamma_1-1} \right] \right\}^{\delta_{i1}} \times \\ \left\{ \left[1 - \left(1 - e^{-\alpha_1 t_i^{\beta_1}} \right)^{\gamma_1} \right] \left[\alpha_2 \beta_2 \gamma_2 t_i^{\beta_2-1} e^{-\alpha_2 t_i^{\beta_2}} (1 - e^{-\alpha_2 t_i^{\beta_2}})^{\gamma_2-1} \right] \right\}^{\delta_{i2}} \times \\ \left\{ \left[1 - \left(1 - e^{-\alpha_2 t_i^{\beta_2}} \right)^{\gamma_2} \right] \left[1 - \left(1 - e^{-\alpha_1 t_i^{\beta_1}} \right)^{\gamma_1} \right] \right\}^{1-\delta_{i1}-\delta_{i2}} \quad (8)$$

In the equation above α_1 and α_2 will be defined as the linear combination of independent variables.

$$\alpha_1 = \exp(\omega_{01} + \omega_{11}x_{11} + \dots + \omega_{l1}x_{l1}), \alpha_2 = \exp(\omega_{02} + \omega_{12}x_{12} + \dots + \omega_{l2}x_{l2})$$

Therefore, the log-likelihood function can be derived as in the following form:

$$l(\emptyset|t) = \\ 2 \sum_{i=1}^n \delta_{i1} \ln \left[1 - \left(1 - e^{-\alpha_2 t_i^{\beta_2}} \right)^{\gamma_2} \right] \\ + 2 \sum_{i=1}^n \delta_{i1} \ln \left[\alpha_1 \beta_1 \gamma_1 t_i^{\beta_1-1} e^{-\alpha_1 t_i^{\beta_1}} (1 - e^{-\alpha_1 t_i^{\beta_1}})^{\gamma_1-1} \right] \\ + 2 \sum_{i=1}^n \delta_{i2} \ln \left[1 - \left(1 - e^{-\alpha_1 t_i^{\beta_1}} \right)^{\gamma_1} \right] \\ + 2 \sum_{i=1}^n \delta_{i2} \ln \left[\alpha_2 \beta_2 \gamma_2 t_i^{\beta_2-1} e^{-\alpha_2 t_i^{\beta_2}} (1 - e^{-\alpha_2 t_i^{\beta_2}})^{\gamma_2-1} \right] \\ + 2 \sum_{i=1}^n (1 - \delta_{i1} - \delta_{i2}) \ln \left[1 - \left(1 - e^{-\alpha_2 t_i^{\beta_2}} \right)^{\gamma_2} \right] \\ + 2 \sum_{i=1}^n (1 - \delta_{i1} - \delta_{i2}) \ln \left[1 - \left(1 - e^{-\alpha_1 t_i^{\beta_1}} \right)^{\gamma_1} \right] \quad (9)$$

Then in order to acquire the parameters estimates, the first derivation has been calculated from the first likelihood function toward each of the parameters. As there are no closed form solutions for these equations, the numerical method technique such as Newton-Raphson method, is used for computing the MLE of the parameters.

Also the variance of each of the parameters estimated is equal to the inverse of the observed Fisher information matrix, that is, the inverse of the matrix of second derivatives of the log-likelihood function locally at $\widehat{\beta}_j, \widehat{\gamma}_j$ and $\widehat{\omega}_{Lj}$.

$$\frac{\partial l}{\partial \beta_j} = 0, \quad \frac{\partial l}{\partial \gamma_j} = 0, \quad \frac{\partial l}{\partial \omega_{Lj}} = 0 \quad j = 1, 2 \quad \text{and} \quad L = 0, 1, \dots, l \quad (10)$$

3. A Simulation Study

A simulation study was done in order to assess the accuracy of the estimated parameters from generalized Weibull model and to compare their performance with the Weibull model under different sample sizes. The simulation program has been written in the software R. For the amount of different samples ($n=200$, $n=500$, $n=1000$ and $n=2000$) in 500 repetition, first it is the time to each of the two competing events ($j=1$ and $j=2$) was assumed to depend on one covariate X and to define the scale parameter α_j in the form of a linear combination of this covariate.

thus an covariate X from the normal distribution with the mean 0.6 and the standard deviation 0.1 was produced and then $\alpha_1 = \exp(\theta_0 + \theta_1 X)$ and $\alpha_2 = \exp(\theta_2 + \theta_3 X)$ were calculated in which $\theta_0 = 0$, $\theta_1 = 1.2$, $\theta_2 = 0.1$, and $\theta_3 = 0.5$. Also the values of $\gamma_2, \gamma_1, \beta_2, \beta_1$ were considered as 1. So we have here two scenarios.

Scenario 1 is dedicated to the generalized Weibull model ($\theta_0 = 0$, $\theta_1 = 1.2$, $\theta_2 = 0.1$, $\theta_3 = 0.5$, $\gamma_1 = 0.8$, $\gamma_2 = 1.2$, $\beta_2 = 1$, $\beta_1 = 1$) and scenario 2 is dedicated to the Weibull model ($\theta_0 = 0$, $\theta_1 = 1.2$, $\theta_2 = 0.1$, $\theta_3 = 0.5$, $\gamma_1 = 1$, $\gamma_2 = 1$, $\beta_2 = 1$, $\beta_1 = 1$).

Then we employed these steps for simulation as the following:

- 1- Simulation of survival time T using the cumulative hazards for both causes of failure.

$$H_1(t) + H_2(t) = u \quad (11)$$

where u is a random variable that has a exponential distribution with parameter 1 [$u \sim \text{Exp}(1)$]. Then the root of the following equation which is the same survival time (t) can be calculated and it was obtained with the “uniroot” R function.

$$\varphi = -\log \left[1 - \left(1 - e^{-\alpha_1 t^{\beta_1}} \right)^{\gamma_1} \right] - \log \left[1 - \left(1 - e^{-\alpha_2 t^{\beta_2}} \right)^{\gamma_2} \right] - u = 0 \quad (12)$$

- 2- For the time value which is simulated in step1, we can determine the functions $h_1(t)$, $h_2(t)$ as functions of time which are dependent on covariate (X).
- 3- Then, we refer to the article by Beyersmann et al. (2009) on Simulating competing risks data in survival analysis, and for the simulation survival time T , we perform a binomial experiment with the probability $\frac{h_1(t)}{h_1(t) + h_2(t)}$ for cause of failure 1 (event type 1), in this way if the produced random number equals 1, the first cause ($j=1$), otherwise the second cause of failure ($j=2$) will be determined.
- 4- For the production of times of censor, if the survival time is greater than 2, the acquired time is considered as censoring times.

Thus, we can fit the two models of generalized Weibull and Weibull on the simulated data. The maximum likelihood estimates of the parameter vectors were calculated by “optim” R function and by applying the Newton–Raphson algorithm. The simulation results for scenarios 1 and scenarios 2 are shown in table1 and table2 , respectively. The average estimates, MSE (mean of square error) and (SD) standard deviation acquired from the simulation of generalized Weibull model are shown in tables 1 and 2.

Tables1 shows, for all parameters $(\theta_0, \theta_1, \gamma_1, \beta_1, \theta_2, \theta_3, \gamma_2, \beta_2)$ as it can be seen the SD and MSE of the estimates have been decreased when the sample size increases. Also the average estimates is very close to the true values which show the generalized Weibull model as a good one in estimating the parameters. In table 1, the mean of the estimates for parameter of the form γ_1, γ_2 in generalized Weibull competing risks model gets close to the true values ($\gamma_1 = 0.8, \gamma_2 = 1.2$) by increasing sample size and they have a very small difference.

The results of the second simulation are summarized in Table 2. The average estimates, the SD and the MSE of the parameters estimates were obtained with the Weibull model for the simulated data. As it be seen, MSE of the estimates is decreased when the sample size increases except in the case where β_1 and β_2 .

4. A Real Data Set Example

In this section the generalized Weibull model has been applied on a real data set and also it has been compared with Weibull model. colorectal cancer patients that have been registered that have registered in the cancer registry center of Research institute of Gastroenterology and Liver Disease, Shahid Beheshti University of Medical Sciences, Tehran, Iran from 2004 to 2013 were used in this study and their survival situation was identified.

For colorectal cancer data, the first cause of failure (death due to colorectal cancer, $j=1$) considered as the main event in survival model, and the second cause (death due to other events, $j=2$) considered as the competing risks. Also the patients who were alive till the end of the study had been considered as the right censoring. We assessed the effects of sex(x_1), age at diagnosis(x_2), family history (x_3), and BMI (x_4) on the survival time have been studied. We use the generalized Weibull distribution against the Weibull distribution for fitting the data in the Presence of these covariates.

For the two models $j=1,2$ and $\alpha_j = \exp(\omega_{0j} + \omega_{1j}x_{1j} + \omega_{2j}x_{2j} + \omega_{3j}x_{3j} + \omega_{4j}x_{4j})$. Thus the vectors for parameters include $\phi_1=(\omega_{01}, \omega_{11}, \omega_{21}, \omega_{31}, \omega_{41}, \beta_1, \gamma_1)$ for the first cause of failure (death due to the colorectal failure) and $\phi_2=(\omega_{02}, \omega_{12}, \omega_{22}, \omega_{32}, \omega_{42}, \beta_2, \gamma_2)$ for the second cause of failure (death due to other events). The Fisher information matrix will be a matrix of 14×14 . In Table 3 we present for both models, the maximum likelihood estimates for $\widehat{\omega}_{Lj}, \beta_j$ and $\gamma_j, j = 1, 2$ and $L=0, \dots, 4$.

The estimate of parameters was done through the maximum likelihood approach using the software R; and as there are no closed form solutions for the equations, the numerical

Newton-Raphson approach was applied. In this study, in order to compare (Akaike, 1983) generalized Weibull competing risks model and Weibull model the Akaike Information Criterion (AIC) has been used.

From 388 patients, it was observed censored time for 271 patients (69.8%), death due to colorectal cancer for 104 patients (26.8%), and death due to other events ($j=2$) for 13 patients (3.4%). The maximum likelihood estimate and standard deviation of each parameters of the model are presented in table 3. According to the results of table 3 in generalized Weibull competing risks model, all the variables exist in the model have an impact on colorectal cancer mortality. While in the Weibull competing risks, just the effect of sex was statistically significant for mortality of colorectal cancer.

According to the results of table 3 the AIC of generalized Weibull is lesser than the Weibull model and thus the generalized Weibull model is more suitable for this data set. Moreover, based on the likelihood ratio test statistics (LR), the efficiency of competing risk generalized Weibull model for the real data set of colorectal cancer was evaluated. Following this, the null hypothesis and the alternative hypothesis are defined as follows:

$H_0: \gamma_1 = 1, \gamma_2 = 1$, Competing risks with Weibull distributions (W)

$H_1: \gamma_1 \neq 1, \gamma_2 \neq 1$, Competing risks with generalized Weibull distributions (GW)

Under the null hypothesis, The value $X_L = 2L_{GW} - 2L_W = 26.08$ has a chi-square distribution with two degrees of freedom and the p-value is less than 0.001. Here, L_W , and L_{GW} are the log-likelihood functions under H_0 , and H_1 , respectively. Hence, based on the values of X_L , and p-value, the generalized weibull distribution fits better than the weibull distribution for these data.

5. Discussion and Conclusion

In this paper, we extend Sarhan and Wahed approach that is the parametric approach to the analysis of competing risks data based on generalized Weibull model. A set of real data and simulation study were used for illustrative purposes.

Our simulation study showed that the estimate of generalized Weibull model parameters are very close to real values and MSE and SD of model parameter drop as the sample size has been increased. The analysis of real data of colorectal cancer showed that based on the AIC s, generalized Weibull distribution is more suitable than Weibull distribution for this set of data and the (se) of estimated data of generalized Weibull model is lesser than the Weibull model.

Yet, many studies have been accomplished in the field of parametric competing risks, see Kundu and Basu (2000), Park (2005), Shayan et al.(2011), Iskandar (2016), Moamer et al. (2017). Fine and Jeong (2007) evaluated the covariate effects in competing risks data based on Gompertz. Mazucheli et al. (2011) based on Lindley competing risks model, and evaluated the covariate effects on survival time. Other studies that analyzed the competing risks using generalized Weibull, included just the main parameters without the presence of independent variables. As evaluating the influence of covariates on survival

time is one of the purposes of competing risks analysis, thus in this study we tried to evaluate the covariate effects on the survival time in generalized Weibull model, we defined the scale parameter in the form of a linear combination of covariate. Also and we estimated the parameters of each failure using the maximum likelihood approach.

Although Bayesian estimation has been used widely in survival analysis, but this approach so far has been limited to single risk models. Therefore for further research, in order to estimate the unknown parameters of generalized Weibull model data in the presence of competing risks, the Bayesian approach can be applied instead the maximum likelihood approach. Typically for parametric survival models, the scale parameter is reparameterized in terms of covariates (predictor variables), we suggest the parameters γ and β will be defined based on a linear combination of covariates instead α parameter scale.

References

1. H. Putter, M. Fiocco and R. B. Geskus, 2007. Tutorial in biostatistics: Competing risks and multi-state models. *Statistics in medicine*, 26(11), 2389-2430.
2. M. Pintilie, 2007. Analysing and interpreting competing risk data. *Statistics in medicine*, 26(6), 1360-7.
3. N. Porta Bleda, G. Gómez Melis, C. Rosingana, M. Luz and N. Malats i Riera, 2007. Competing risks methods.
4. J. P. Fine and R. J. Gray, 1999. A proportional hazards model for the subdistribution of a competing risk. *Journal of the American statistical association*, 94, 496-509.
5. A. M. Sarhan, 2007. Analysis of incomplete, censored data in competing risks models with generalized exponential distributions. *IEEE Transactions on Reliability*, 56(1), 132-138.
6. D. R. Cox, 1959. The analysis of exponentially distributed life-times with two types of failure, *Journal of the Royal Statistical Society. Series B (Methodological)*, 21, 411-421.
7. F. Pascual, 2010. Accelerated life test planning with independent lognormal competing risks. *Journal of Statistical Planning and Inference*, 140(4), 1089-1100.
8. S. Yáñez, L. A. Escobar and N. González, 2014. Characteristics of two competing risks models with weibull distributed risks. *Revista de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales*, 38(148), 298-311.
9. I. Iskandar and Y. S. Gondokaryono, 2016. Competing risk models in reliability systems, a weibull distribution model with bayesian analysis approach. *IOP Conference Series: Materials Science and Engineering*, IOP Publishing.
10. J. H. Jeong, 2006. A new parametric family for modelling cumulative incidence functions: Application to breast cancer data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 169(2), 289-303.

11. E. Z. Martinez, J. A. Achcar, A. A. Jácome and J. S. Santos, 2013. Mixture and non-mixture cure fraction models based on the generalized modified weibull distribution with an application to gastric cancer data. *Computer methods and programs in biomedicine*, 112(3), 343-355.
12. G. S. Mudholkar, D. K. Srivastava and G. D. Kollia, 1996. A generalization of the weibull distribution with application to the analysis of survival data. *Journal of the American Statistical Association*, 91(436), 1575-1583.
13. G. S. Mudholkar and D. K. Srivastava, 1993. Exponentiated weibull family for analyzing bathtub failure-rate data. *IEEE Transactions on Reliability*, 42(2), 299-302.
14. R. L. Prentice, J. D. Kalbfleisch, A. V. Peterson Jr, N. Flournoy, V. Farewell and N. Breslow, 1978. The analysis of failure times in the presence of competing risks. *Biometrics*, 541-554.
15. D. Kleinbaum and M. Klein, 1996. *Survival analysis: A self-learning text*. Springer, New York .
16. A. R. Baghestani, T. Daneshvar, M. A. Pourhoseingholi and H. Asadzade, 2014. Survival of colorectal cancer patients in the presence of competing-risk. *Asian Pac J Cancer Prev*, 15, 6253-6255.
17. B. Lau, S. R. Cole and S. J. Gange, 2009. Competing risk regression models for epidemiologic data. *American journal of epidemiology*, kwp107.
18. A. S. Wahed, T. M. Luong and J. H. Jeong, 2009. A new generalization of weibull distribution with application to a breast cancer data set. *Statistics in medicine*, 28(16), 2077-2094.
19. A. M. Sarhan, M. Alameri and I. Al-Wasel, 2013. Analysis of a competing risks model with generalized weibull distributions. *Pak. J. Statist*, 29(3), 271-281.
20. I. A. Alwasel, 2009. Statistical inference of a competing risks model with modified weibull distributions. *International Journal of Mathematical Analysis*, 3(19), 905-918.
21. R. J. Gray, 1988. A class of k-sample tests for comparing the cumulative incidence of a competing risk. *The Annals of statistics*, 1141-1154.
22. J. Beyersmann, A. Allignol and M. Schumacher, 2011. *Competing risks and multistate models with r*. Springer Science & Business Media.
23. J. H. Jeong and J. Fine, 2006. Direct parametric inference for the cumulative incidence function. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 55(2), 187-200.
24. J. Beyersmann, A. Latouche, A. Buchholz and M. Schumacher, 2009. Simulating competing risks data in survival analysis. *Statistics in medicine*, 28(6), 956-971.
25. H. Akaike, 1983. Information measures and model selection. *Bulletin of the International Statistical Institute*, 50(1), 277-291.
26. D. Kundu and S. Basu, 2000. Analysis of incomplete data in presence of competing risks. *Journal of Statistical Planning and Inference*, 87(2), 221-239.

27. C. Park, 2005. Parameter estimation of incomplete data in competing risks using the em algorithm. *IEEE Transactions on Reliability*, 54(2), 282-290.
28. Z. Shayan, S. Ayatollahi and N. Zare, 2011. A parametric method for cumulative incidence modeling with a new four-parameter log-logistic distribution. *Theoretical Biology and Medical Modelling*, 8(1), 43.
29. I. Iskandar, 2016. Competing risk models in reliability systems, a gamma distribution model with bayesian analysis approach. *IOP Conference Series: Materials Science and Engineering*, IOP Publishing.
30. S. Moamer, A. Baghestani, M. A. Pourhoseingholi, N. Hajizadeh, F. Ahmadi and M. Norouzinia, 2017. Evaluation of prognostic factors effect on survival time in patients with colorectal cancer, based on weibull competing-risks model, *Gastroenterology and Hepatology from bed to bench*, 10(1), 54.
31. J.-H. Jeong and J. P. Fine, 2007. Parametric regression on cumulative incidence function. *Biostatistics*, 8(2), 184-196.
32. J. Mazucheli and J. A. Achcar, 2011. The lindley distribution applied to competing risks lifetime data. *Computer methods and programs in biomedicine*, 104(2), 188-192.

Table1: Mean, average mean square error (MSE) and standard deviation (SD) for $\theta_0=0$, $\theta_1 = 1.2$, $\gamma_1 = 0.8$, $\beta_1 = 1$, $\theta_2 = 0.1$, $\theta_3 = 0.5$, $\gamma_2 = 1.2$ and $\beta_2 = 1$ in generalized Weibull competing risks model

	Sample Size	Parameter							
		θ_0	θ_1	γ_1	β_1	θ_2	θ_3	γ_2	β_2
Mean	200	0.272	0.912	1.273	0.938	-0.663	1.436	1.511	1.641
	500	0.091	1.053	0.977	0.994	0.193	0.368	1.573	1.047
	1000	0.071	1.072	0.911	1.007	0.202	0.395	1.508	1.132
	2000	0.068	1.064	0.873	1.023	0.166	0.478	1.454	0.934
SD	200	0.615	0.931	1.012	0.274	2.732	3.816	1.106	1.561
	500	0.445	0.63	0.374	0.195	0.568	0.752	0.997	0.383
	1000	0.314	0.449	0.222	0.144	0.38	0.494	0.639	0.435
	2000	0.202	0.323	0.147	0.103	0.255	0.342	0.342	0.133
MSE	200	0.453	0.950	1.244	0.078	8.050	15.440	1.321	2.849
	500	0.199	0.418	0.161	0.048	0.323	0.576	1.124	0.147
	1000	0.099	0.212	0.06	0.021	0.155	0.254	0.49	0.201
	2000	0.042	0.115	0.022	0.011	0.065	0.117	0.171	0.018

Table2: Mean, average mean square error (MSE) and standard deviation (SD) for $\theta_0=0$, $\theta_1 = 1.2$, $\gamma_1 = 1$, $\beta_1 = 1$, $\theta_2 = 0.1$, $\theta_3 = 0.5$, $\gamma_2 = 1$ and $\beta_2 = 1$ in Weibull competing risks model

	Sample Size	Parameter							
		θ_0	θ_1	γ_1	β_1	θ_2	θ_3	γ_2	β_2
Mean	200	-0.742	2.422	0.968	1.548	0.157	0.459	1.431	0.951
	500	-0.02	1.184	1.002	1.074	0.006	0.579	1.096	1.086
	1000	-0.041	1.21	0.966	1.058	-0.004	0.596	1.011	1.075
	2000	-0.102	1.27	0.889	1.09	0.016	0.566	0.962	1.049
SD	200	1.653	2.640	0.505	1.428	0.903	1.396	0.813	0.376
	500	0.449	0.623	0.932	0.24	0.636	0.878	0.6	0.432
	1000	0.381	0.535	0.287	0.166	0.449	0.625	0.401	0.293
	2000	0.236	0.344	0.15	0.105	0.28	0.387	0.212	0.145
MSE	200	3.284	8.476	0.256	2.340	0.818	1.951	0.848	0.144
	500	0.202	0.389	0.153	0.063	0.413	0.778	0.37	0.194
	1000	0.147	0.286	0.084	0.031	0.212	0.535	0.161	1.091
	2000	0.066	0.023	0.035	0.821	0.086	0.154	0.047	0.212

Table 3: Parameter estimation (standard error) for Generalized Weibull and Weibull competing risks model

Model	Parameter	Colorectal cancer MLE(se.)	Competing risk MLE(se.)	AIC(-2 × log-like)
Generalized Weibull	ω_0	1.031(0.014)*	0.771(0.054)*	1410.23(1396.23)
	ω_1	0.0015(0.0002)*	0.005(0.0009)*	
	ω_2	0.003(0.0005)*-	-0.031(0.002)*	
	ω_3	0.771(0.008)*	0.037(0.031)	
	ω_4	-0.008(0.02)*	0.240(0.072)*	
	β	0.175(0.008)*	0.275(0.033)*	
	γ	31.180(1.610)*	23.172(3.113)*	
Weibull	ω_0	-1.775(0.052)*	-3.640(0.277)*	1434.32(1422.32)
	ω_1	0.008(0.0009)*	0.0316(0.004)*	
	ω_2	-0.0001(0.002)	-0.117(0.012)*	
	ω_3	0.048(0.030)	0.0526(0.164)	
	ω_4	-0.056(0.078)	1.259(0.353)*	
	β	0.671(0.027)*	1.310(0.129)*	
* significance level: 5%				