Poisson Cure Rate Model with Generalized Exponential Lifetimes

by

Joynob Ara Siddiqua

A Thesis submitted to the Faculty of Graduate Studies of
The University of Manitoba
in partial fulfillment of the requirements of the degree of

Master of Science

Department of Statistics
University of Manitoba
Winnipeg

Abstract

In this thesis, we consider a competing risks scenario wherein lifetimes are potentially right censored. Instead of considering all the patients to be at risk to the event of interest, we assume that a proportion of these patients are cured and have no recurrence of the disease, known as the cure fraction. We further assume that the number of competing risks is random and follows a Poisson distribution. We consider the lifetimes of individuals to follow a two parameter generalized exponential distribution. The objective is to estimate the model parameters. Using a direct approach and also by using the expectation maximization estimation approach, we obtain maximum likelihood estimates. Standard errors of the estimates are obtained by inverting the observed Fisher information matrix. Monte Carlo simulations are used to demonstrate the performance of the two methods of estimation. Finally, we fit our model to two real data sets to illustrate the model competence.

Keywords: cure rate, generalized exponential distribution, lifetimes, competing risks, censoring, Poisson distribution.

Acknowledgment

It is my privilege to express my sincere appreciation to the Department of Statistics, University of Manitoba, for giving me the opportunity to work on this thesis. I would like to express the deepest gratitude to all of my teachers and professors, both current and previous, for enabling me with knowledge and understanding of research. Special mentions are to my most recent course teachers, Alexandre Leblanc, Elif Acar, Saumen Mandal, Liqun Wang, A. Thavaneswaran, Brad Johnson and Katherine Davies.

My utmost gratitude to my supervisor and professor, Dr. Katherine Davies, for being my mentor. She continuously and convincingly conveyed a spirit of adventure in regard to research. Without her guidance and persistent help, this dissertation would not be possible. I am grateful to her for the continuous support, extreme patience and encouragement during the time of thesis preparation.

I am very thankful to the examining committee for agreeing to evaluate my thesis work. My heartiest appreciation to Margaret, Susan and Rosa for being so helpful in this journey. For technical help and moral support, I am also very thankful to Dave Gabrielson.

Lastly and most importantly, I must express my profound gratitude to my creator, parents, sibling, family and friends, whose presence gave me strength to continue. I offer my regards and blessings to all of those who supported me in any respect during the completion of this thesis.

Dedication

To my parents and loved ones

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

The longevity of patients is of great importance in biomedical studies. For patients with long term diseases, prolonged lifetimes can be indicators of improved treatment. When studying such patients, one may find a part of them to be permanently cured who don't show any recurrence of the disease of interest. This proportion of the total target population is of medical significance and that is the motivation for developing models referred to as cure rate models.

As mentioned in Pal and Balakrishnan (2016), such cure rate models can be applied in varieties of disciplines other than biomedical studies, such as, criminology, finance, demography, manufacturing, and industrial reliability. This means that cure rate models can also be used to model the distribution of survival time or failure time of a product/item /subject for a specified population. Moreover, cure rate models are useful under the assumption of multiple mode of failures or competing causes. If there is more than one reason to experience the event of interest, then understanding and developing such cure rate model using the cured fraction is crucial. Furthermore, censoring of survival time or failure time needs to be properly addressed for developing models as such.

Our proposed model is inspired by the work of Balakrishnan and Pal (2016)where they have

proposed a cure rate model with Weibull lifetime distribution. Also, another related paper is by Pal and Balakrishnan (2016), where they considered a cure rate model with generalized gamma lifetime distribution. We will be proposing a model with generalized exponential lifetime distribution in a competing cause scenario. After introducing the proposed model, we will subsequently develop two methods of estimation for finding the estimates of the parameters of the model. The performance of these methods is then investigated numerically.

1.2 Thesis organization

In Chapter 2, some relevant concepts are discussed. First, we briefly discuss some important tools used in the analysis of survival data. We then discuss the concepts of censoring, competing risks and cure rate models. This is followed by some properties on the generalized exponential distribution. Finally, Chapter 2 closes with the description of two approaches to maximum likelihood estimation and some model discrimination tools.

In Chapter 3, we introduce our proposed model. The model assumes generalized exponential lifetimes and incorporates censoring and competing risks. In addition, in this chapter, we construct the likelihood function. In Chapter 4, we develop likelihood inference for our proposed model in two different ways. This is followed by a numerical assessment (via a simulation study and illustrative examples) in Chapter 5. Finally, in Chapter 6, some conclusions and ideas for future work are provided. Some of the R codes used to produce the numerical results and the tables in this thesis are provided in Appendix B.

Chapter 2

Preliminaries

2.1 Concepts in survival analysis

In survival analysis, the concept of lifetime is the root of all survival analysis problems. Lifetimes are positive random variables mostly dealt in reliability engineering and survival analysis and some other disciplines. Finkelstein (2008) states that a lifetime can be understood by its distribution function. Lifetime can be defined by the time before the occurrence of death, end of marriage or some other "end event". For example, manufactured items have mechanical or electronic components and in order to get information on their durability, life tests are performed. That means in a laboratory setting, such items may put into operation and observed until they fail. Therefore it is common to refer lifetimes as "failure times" as it is said to be "failed" when it stops operating satisfactorily.

Lawless (2003) discusses the case of a single continuous lifetime variable, T. Specifically, let T be a nonnegative random variable denoting the lifetimes of individuals in some population. Other than some exceptions, the lifetimes are defined over $[0, \infty)$ interval. If the probability density function (p.d.f.) for T is denoted by $f(t)$, then the cumulative distribution function (c.d.f.) is defined as

$$F(t) = Pr(T \leq t) = \int_0^t f(x)dx.$$

The probability of an individual surviving to time t is given by

$$S(t) = Pr(T \geq t) = \int_t^{\infty} f(x)dx;$$

this is called the survival function. In some contexts involving systems or lifetimes of manufactured items, $S(t)$ is also referred to as the reliability function. Here, $S(t)$ is a monotone decreasing continuous function with $S(0) = 1$ and $S(\infty) = \lim_{t\to\infty} S(t) = 0$. Sometimes, we may allow $S(\infty) > 0$, to consider settings where some individuals never fail; these are special cases. Moreover, the pth quantile of the distribution of T is the value t_p such that $Pr(T \leq t_p) = p$, $0 \leq p \leq 1$. Note t_p can be found as $t_p = F^{-1}(p)$. The pth quantile is also called the $100p$th percentile of the distribution. The 0.5 quantile is referred to as the median of the distribution.

The instantaneous rate of death or failure at time t, given that up to time t the subject has already survived, is known as the hazard function. It is also known as hazard rate or force of mortality. The hazard function $h(t)$ is an essential concept with lifetime distributions in survival analysis and is defined as:

$$h(t) = \frac{\lim_{\Delta t\to 0} Pr(t \leq T < t + \Delta t | T \geq t)}{\Delta t} = \frac{f(t)}{S(t)}.$$

Therefore, in the time interval $t < x < t + \Delta t$, the probability of a subject dying for small value of Δt, where the subject has already survived up to time t, is defined as $h(t)\Delta t$.

Another important topic in survival analysis is the Cox proportional hazards model, which assumes that all the individuals in the study population are susceptible to the event of interest. It considers a binary response which determines whether the individual is susceptible or not. According to Lawless (2003), the proportional hazards model is the best known semi-parametric lifetime regression model introduced by Cox, which has the hazard function for T given X to be of the form:

$$h(t|x) = h_0(t)exp(\beta' x),$$

where $h_0(t)$ is an arbitrary "baseline" hazard function meaning that in the absence of the covariate, hazard function $h(t|x)$ will take the form of $h_0(t)$. A convenient way of handling time varying covariates is through the hazard function. Let $X(t) = x(s), 0 \leq s \leq t$ refer to the history up to time t, with $X(\infty) = X$. It is assumed that the hazard function for T given X depends only on $X(t)$; we denote this as $h(t|x(t))$. Another flexible approach is defining a vector $w(t)$ that refers to the features of $X(t)$, and accordingly define $h(t|X(t))$ as a function of t and $w(t)$. The multiplicative formulation

$$h(t|x(t)) = h_0(t)exp(\beta' w(t))$$

is useful and is an extension of proportional hazard function. The source of survival data varies according to the application. For instance, survival data may arise in clinical trials or reliability studies, which are commonly referred to as life tests.

Other sources of survival data are medical studies. When the subjects/units aren't human, another tool used to obtain survival data are accelerated life tests, which are a special type of life tests. These tests are applied especially in the area of reliability problems. The main principle behind these tests is the imposition of stresses in order to observe the lifetimes in shorter time; for further details see Pascual et al. (2006).

2.2 Censoring

In a life test, the chronological time required to observe the full lifetime of the units on test may not be possible due to practical constraints. In such cases, "censoring" is often introduced. For instance, in clinical trials, censoring is very common since the trials are usually terminated before all individuals fail(die). Also, at various time points the individuals may enter the study (Arnold et al. (1992)).

There are four general types of censoring: Type I censoring, Type II censoring, random censoring and progressive censoring. Suppose there is a life test and n units are put on test.

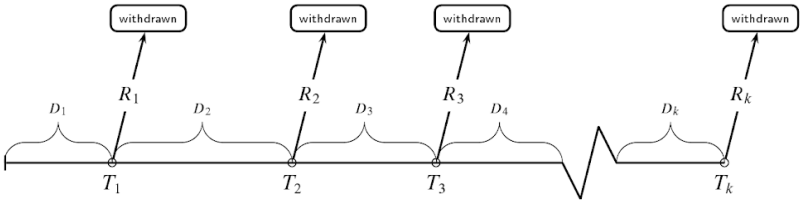
If the experiment is terminated at a predetermined time T, then items which fail prior to this specified time are observed and the recorded data is referred to as Type I censored. We say these observations are right censored in this case, since we can only observe the minimum of the lifetime and censoring time. In this case, the number of observations is random but the duration of the test is fixed.

On the other hand, suppose it is decided to terminate experiment at the time of the rth failure. We refer this type of data as a Type II right censored sample. In this case, the number of observations is fixed but the duration of the test is random. It may correspond to left censoring wherein smaller values are censored or one can have censoring of multiple regions and can only observe the maximum of the lifetime and censored time. If there is right and left censoring, then it is referred to as double censoring.

Suppose there are n lifetimes $Y_1, ..., Y_n$ with a common c.d.f. $F(y;\theta)$, and p.d.f. $f(y;\theta)$ where θ is the unknown parameter. Now, with the ith item let us associate a random variable C_i called the censoring time whose c.d.f. is F_c and it is free of θ. We define $T_i = min(Y_i, C_i)$. Then $D_i = 1$ if $T_i = Y_i$ and else $D_i = 0$. Assuming independence between Y_i and C_i, we observe the pairs (T_i,D_i), where $i = 1, 2, ..., n$. Thus each lifetime is censored by an independent time and also whether we have observed the lifetime or the censoring time is known. This is called a random censoring scheme.

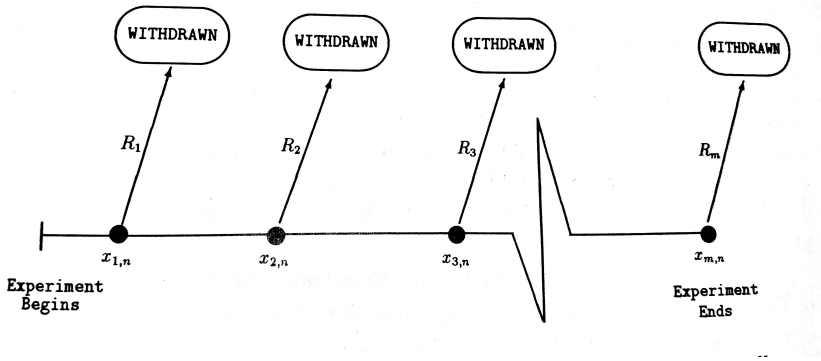
Another kind of censoring is progressive censoring, which is of two kinds: Type I progressive censoring and Type II progressive censoring. Let R be the number of items to be tested. At time T_1, we may randomly remove R_1 unfailed items among the R; at time T_2, we randomly remove R_2 unfailed items, etc., where $T_1 < T_2 < ...$ are predetermined times. If R_i items exceeds the remaining number of unfailed items at time T_i, then the experiment will be terminated. This is Type I progressive censoring scheme. All T_i and R_i values are prefixed ahead of experiment. A graphical representation of Type I progressive censoring is given in Figure 2.1 as shown by Balakrishnan and Cramer (2014):

Figure 2.1: Type I progressive scheme with interval censoring



In a Type II progressive censoring experiment, there is a prefixed progressive censoring scheme $R = (R_1, \ldots, R_m)$, where m is the number of failures to be observed. At the time of the first failure, R_1 items are randomly removed from the test. At the time of the second failure, R_2 surviving items are randomly removed from the test. This continues until the occurence of the mth failure, at which time all remaining units are removed. We refer to the resulting observed failures as a progressively Type II right censored data. A graphical representation of Type II progressive censoring is given in Figure 2.2 as shown by Balakrishnan and Aggarwala (2000):

Figure 2.2: Type II progressive scheme with right censoring



2.3 Competing risks

Individuals can, in some settings, fail in different ways, and are then assigned a mode of failure. The modes may refer to the cause of failure, in which case they are often termed competing risks. For example, an individual in a demographic study might be recorded as dying at age t from one form of cancer, cardiovascular disease, or other causes. Failure modes can also be defined in other ways, for example, to reflect costs or severity of consequences associated with failure Lawless (2003). The probability of dying because of a specific cause prior to death is attributed as risk. But when the death has actually occurred, we also refer to the cause responsible for the death as the risk.

As described by Gross and Clark (1975), rather than the disease for which a patient is under study, the patient may die due to a cause instead of the specific disease that is being studied in a clinical trial. As an example, we may consider a person who is under study due to prostate cancer and may die due to a fatal accident or heart attack. All such risks and the risk of being dead due to disease of interest are called competing risks. The authors also point out that lifetime data involving competing risks can be analyzed in different ways. First, the survival times can be analyzed separately for each cause. And in a simpler model, the existence of several competing risks can be ignored. Lastly, and the most appropriate, one can develop a model that can incorporate all the information in a competing risk data.

One common assumption for competing risk data is that the number of competing risks and the lifetimes associated with these risks are unobservable. These unobservable lifetimes can hence be considered latent variables. What is actually observed is the minimum lifetime due to these risks. In a seminal paper by Prentice et al. (1978), new methods were proposed for the analysis of failure times in the presence of competing risks. In particular, the author considered cause specific hazard function and developed estimation methods. In the initial investigations of competing risk data, such as Prentice et al. (1978), it was common to

assume the risks were independent. In fact, recent literature often still makes this assumption. For instance, in Miyakawa (1982), assuming exponential lifetimes and independent risks, they developed maximum likelihood estimators (MLEs) and uniformly minimum variance unbiased estimators. This work was followed up by Kundu and Basu (1991) wherein the results of Miyakawa (1982) were further investigated and extended to the case of Weibull lifetimes. Also, in Ishioka and Nonaka (1991), the authors estimated the parameters of the two parameter Weibull distribution in the scenario of two independent competing risks. They demonstrated the performance of estimation procedure through simulation and an illustrative example. In Hong and Meeker (2014), the authors developed inference procedures for system reliability data and applied their procedure to competing risks data. These data involved lifetimes of electronic components which were subject to two competing risks (surge and wear out). In a completely different field, Austin et al. (2016) discussed the role of competing risks and cardiovascular disease. When studying time to death due to a cardiovascular cause, death due to other causes was considered as competing risks. The authors developed various non-parametric techniques and constructed cause-specific hazard models. The authors then applied their methods in an illustrative example involving cardiac data.

It is important to note at this junction that competing risks can be dependent, and perhaps this is a more realistic assumption in certain situations. In Moeschberger (1974), the author discussed the estimation for two bivariate lifetime distributions when the causes of failure are dependent. The paper also included an illustrative example which involved the failure of small electrical appliances. As pointed out in the paper, it is suggested that dependent causes of failure often arise in physical situations. For instance, dependent competing risks may exist in the engineering and reliability fields. For example, there is what is referred to as a common shock model. As described in Eryilmaz (2018), a shock can be viewed as a load which might refer to mechanical stress, a voltage or internally generated stress such as temperature. In

common shock models, in addition to independent shocks, there is a common shock which can impact all components simultaneously. For some work in the literature on common shock models, we refer the reader to Gåsemyr and Natvig (1995). In addition to this, Moeschberger and Klein (1995) provide a survey of the literature on the various statistical methods that have been developed for the situation where the competing risks are not independent. Since then, many works have focused on the statistical analysis of dependent competing risk data; see for example: Feizjavadian and Hashemi (2015), Kundu and Basu (2000), Wu et al. (2017) and Shi and Wu (2016).

2.4 Cure rate model

Cure rate models or long term survival models are the models for lifetime data which include a proportion, as a fraction of the total population, who are permanently cured. In the literature, this proportion is referred to as a cure fraction. In survival analysis, we usually consider that all patients are susceptible to the event of interest but in these models, there may is a cure fraction. Hence, these models are referred to as cure rate models. The population can then be separated into two categories: cured and non-cured or susceptible.

There have been many works on cure rate models. Boag (1949) first considered the proportion of cured in all the patients treated for cancer. The remainder of the patients were considered to be susceptible of dying due to disease, if not dead due to some other causes. So the cure rate model was introduced here in a form of a mixture model. The author assumed that the survival time follows a lognormal distribution and the maximum likelihood estimators for the model parameters were found. Also, the precision of the estimators of proportion cured, i.e., the cure rate, was examined.

Berkson and Gage (1952) developed a model with two parameters, one associated with the cured proportion and the other was associated with the susceptible group. As such, the

proportion of the population cured and remaining proportion who were not gave rise to a mixture distribution. The authors argued that the five year survival rate cannot be identified as the proportion of cured patients, since all patients are not guaranteed to die if untreated within five years and also a noncancerous patient can die within this period.

As previously mentioned, the Cox proportional hazards (Cox PH) model assumes that all the individuals in the study population are susceptible to the event of interest. A proportional hazards model was suggested for the susceptible group by Sy and Taylor (2000), which was called proportional hazards cure rate model. They used maximum likelihood techniques to estimate the incidence and latency regression parameters together by using nonparametric likelihood structure. The inverse of the observed information matrix was used to estimate standard errors of the estimates. A comparison between PH mixture model and Weibull mixture model was graphically shown.

A class of cure rate models was proposed by Yin and Ibrahim (2005), which included the mixture cure model and the promotion time cure model as special cases. The promotion time cure model is a cure model where the number of competing causes follows a Poisson distribution, whereas the mixture cure model has the number of competing causes to be a Bernoulli random variable. The proposed model was based on a Box-Cox transformation of the population survival function. In particular they defined a link parameter, where one value yields the mixture cure model and the other value yields the promotion time cure model. Moreover, a general covariate structure was suggested in order to accommodate for different covariate structures.

Rodrigues et al. (2009) proposed a Conway-Maxwell (COM) Poisson cure rate model. If a random variable M follows a COM-Poisson distribution, its probability mass function (p.m.f) is:

$$P[M = m; \eta, \phi] = \frac{1}{Z(\eta, \phi)} \frac{\eta^m}{(m!)^\phi}, \ m = 0, 1, 2, \ldots, \text{ where } Z(\eta, \phi) = \sum_{j=0}^{\infty} \frac{\eta^j}{(j!)^\phi}.$$

This distribution has Bernoulli, Poisson and geometric distribution as special cases. If $\phi = 1$, it reduces to Poisson distribution with parameter η. If $\phi = 0$, it reduces to geometric distribution with parameter $(1 - \eta)$ and $\eta < 1$. Moreover, if $\phi \to \infty$, the limiting case is a Bernoulli distribution with parameter $\frac{\eta}{1+\eta}$. In the case of discrete data, over dispersion and under dispersion are regularly encountered and that can be accounted for in this model. In this cure rate model, it is assumed that the number of competing risks follows a COM-Poisson distribution, call this M.

As mentioned earlier, in a competing risk scenario, it is commonly assumed that both M and the lifetimes due to these risks are unobservable and only the minimum lifetime is observed; call this Y. The survival function of this Y, which they refer to as long term survival function, is expressed as:

$$\begin{aligned} S_p(y) &= P[Y \geq y] \\ &= \sum_{m=0}^{\infty} P(M = m)[S(y)]^m \\ &= \frac{Z(\eta S(y), \phi)}{Z(\eta, \phi)}. \end{aligned} \tag{2.1}$$

The authors developed inference for their model and demonstrated it through an illustrative example.

Kannan et al. (2010) proposed a cure rate model which assumed the generalized exponential distribution for the lifetimes and a cure fraction. The authors also included covariates in their model through the cure proportion. They developed an estimation procedure in the form of the EM algorithm and illustrated their proposed model through a simulation study and a real life dataset involving drug abuse.

Balakrishnan and Pal (2015) considered a flexible cure rate model with generalized gamma lifetimes subject to right censoring. The authors developed their model in a competing

risk scenario where the number of competing risks follows a COM-Poisson distribution. Missingness was identified in the data within the censored observations, which includes both cured and non-cured observations; that is, for a censored observation, it is not known if the subject is susceptible or not. They considered Bernoulli, Poisson, Geometric as special cases of the COM-Poisson distribution. Using the Expectation-Maximization (EM), to be discussed soon, the parameters were estimated. Moreover, a simulation study was performed to test the fit of the models. The results of the simulation concluded that the EM algorithm converges almost accurately to the true parameter. Model discrimination within the generalized gamma family was also carried out using AIC and BIC criterion. Additionally, the model was tested on cutaneous melanoma data with two way model discrimination.

In Balakrishnan and Pal (2016) the authors considered a cure rate model incorporating Weibull distribution to model lifetimes and a COM-Poisson distribution to model the number of competing risks. They assumed right censoring in the data and used EM algorithm to estimate these right censored data. They also observed special cases of the model. The performance of the model was evaluated by Monte-Carlo simulation study.

In Pal and Balakrishnan (2016), they considered a Gamma distribution for the lifetimes of the susceptible group of patients and used the exponentially weighted Poisson distribution to model the number of competing causes while constructing a cure rate model. They considered the case where the number of competing risks were gradually diminishing. It was therefore called a destructive exponential cure rate model. In many real life problems, including naturally occurring mechanisms, it may be more appropriate to make this assumption. They adjusted the existing cure rate model by introducing a new random variable which represents the total number of competing causes, among the initial competing causes, which have not been destroyed. The authors used the EM algorithm to estimate the model parameters. A simulation study was conducted to observe the model fitting. AIC, BIC and likelihood ratio test were performed as ways to discriminate between the models. The authors observed

significant differences between the lognormal distribution and the Weibull, gamma and exponential distribution.

2.5 Generalized exponential distribution

The three parameter generalized exponential distribution has been proposed instead of Weibull and gamma families for lifetime data analysis by Kundu and Gupta (1999). It can be viewed as an exponentiated Weibull distribution when the location parameter is not present. If the location parameter were present, complexity in direct numerical calculations would be higher. When the shape parameter is not an integer, it is difficult to obtain the distribution function or the survival function of the gamma distribution. The c.d.f. of the generalized exponential family for lifetime data y is given as:

$$F(y) = (1 - e^{-\lambda y})^{\alpha} \tag{2.2}$$

for $y > 0, \alpha, \lambda > 0$, and with this, the survival function is defined as:

$$S(y) = 1 - F(y) = 1 - (1 - e^{-\lambda y})^{\alpha}. \tag{2.3}$$

By differentiating the c.d.f., we obtain the p.d.f. as

$$f(y) = \alpha\lambda e^{-\lambda y}(1 - e^{-\lambda y})^{\alpha-1}. \tag{2.4}$$

Therefore, the hazard function is:

$$h(y) = \frac{\alpha\lambda e^{-\lambda y}(1 - e^{-\lambda y})^{\alpha-1}}{1 - (1 - e^{-\lambda y})^{\alpha}}, \text{where} \quad \alpha, \lambda, y > 0. \tag{2.5}$$

Clearly when $\alpha = 1$, this model reduces to the simple exponential distribution. A graphical representation of the c.d.f and the p.d.f. of generalized exponential distribution are shown in Figures 2.3 and 2.4, respectively, for various parameter settings.

The parameters of the generalized exponential (GE) distribution were estimated using classical estimation procedures like maximum likelihood estimation (MLE), method of mo-

Figure 2.3: CDF of the generalized exponential distribution

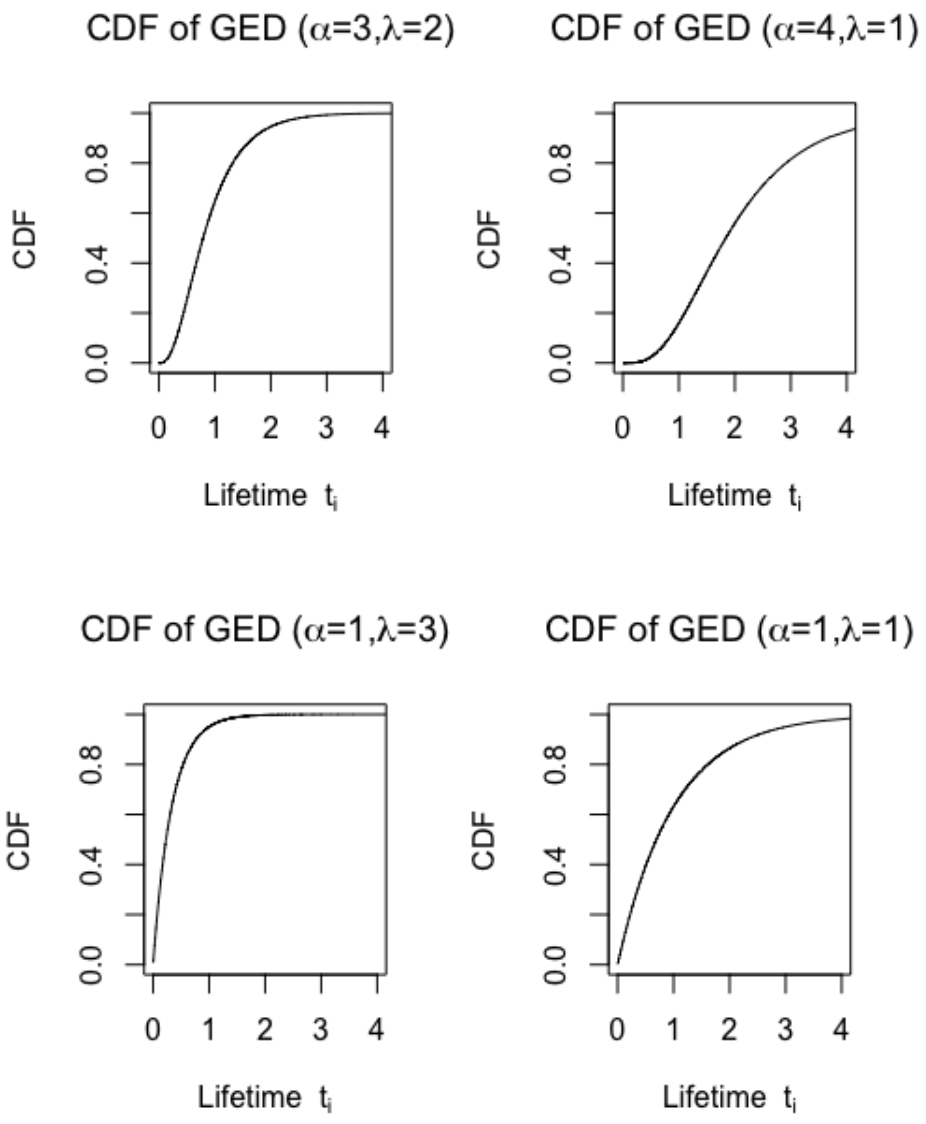
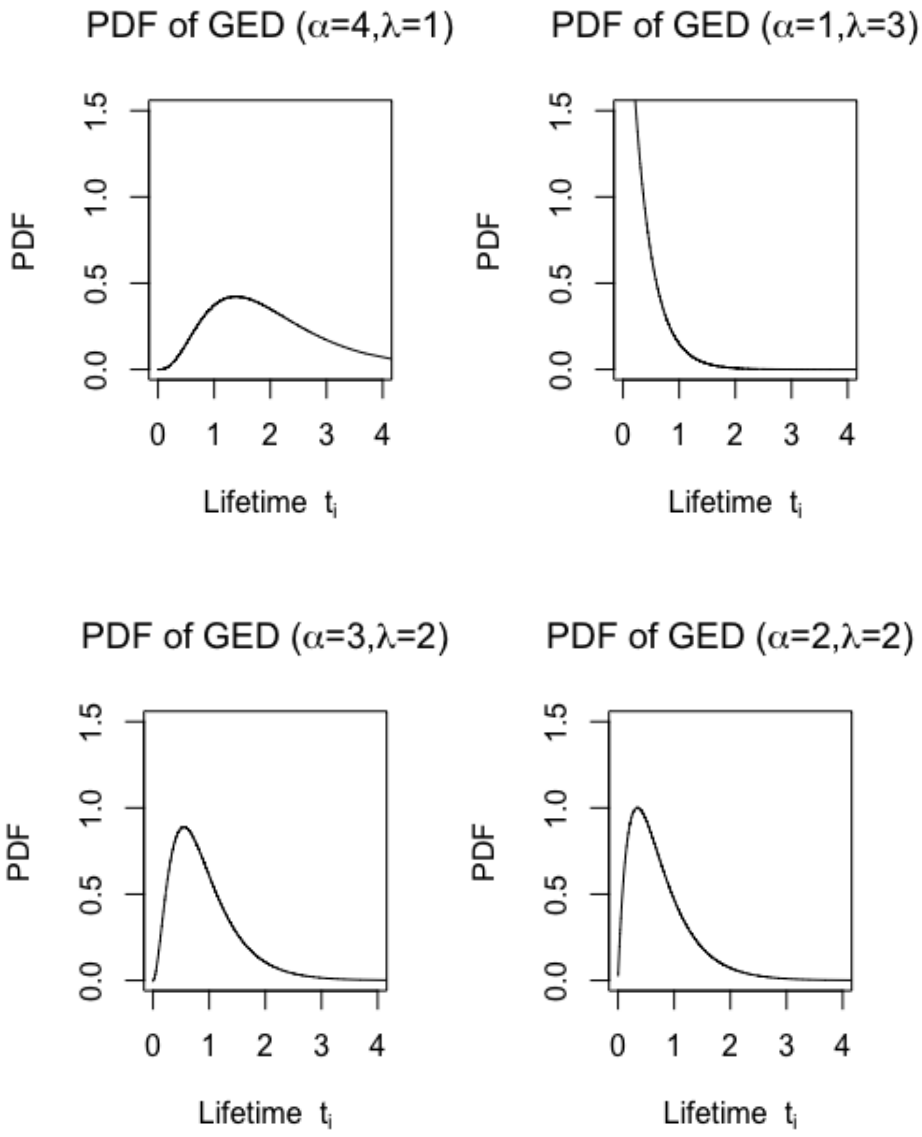


Figure 2.4: PDF of the generalized exponential distribution



ment estimation (MOM), Least square estimation (LSE), percentile estimation, L-moment estimation etc. by Kundu and Gupta (2008). The authors pointed out that the generalized exponential distribution model is quite flexible and can effectively analyze positive lifetime data instead of gamma, Weibull or log-normal model. However, it has properties similar to gamma distribution. It is more suitable for data purposes since gamma has an intractable distribution function. If the data is censored, this model can also be quite useful.

Gupta and Kundu (2007) compared approximate Bayes estimator and exact Bayes estimator with classical maximum likelihood estimations under the assumption of non-informative prior. Under the assumption of gamma prior for scale and shape parameters of GE and assuming the squared error loss function, they calculated the Bayes estimators. A data was used for illustration purposes.

In Gupta and Kundu (2000), the authors briefly discussed the different methods of estimation used on parameters of a generalized exponential distribution. Maximum likelihood estimation, least squares estimation, weighted least squares estimation, percentile estimation, method of moments and L-moment estimation were used for the estimation. The authors considered three cases: (1) α, the shape parameter, is known but λ, the reciprocal of a scale parameter, is unknown; (2) λ is known but α is unknown and; (3) when both α and λ are unknown.

Raqab and Madi (2006) in their paper have described some usefulness of generalized exponential distribution as it can be used in skewed distributions. The authors reflect on the simplicity of its distribution function and accordingly mentioned that the distribution can be extensively used to model lifetime data even with censoring or grouped data. They estimated the model parameters using a Bayesian approach. To use posterior and predictive inferences, they applied stochastic simulation based approaches. The Gibbs and Metropolis samplers were also applied in studying the future failure times and predictive density functions. For illustrative examples, two data sets were used. The authors Markiewicz et al. (2015) studied

the relative use of generalized exponential distribution and inverse Gaussian distribution in fitting the flood extremes data in Polish rivers. For model discrimination, they used various non-parametric methods. According to their study, the GED fits well to the extreme flood data of Polish rivers.

2.6 Inference and model discrimination

2.6.1 Maximum likelihood estimators

According to Casella and Berger (2002), the likelihood function can be used to summarize data. It is the most common and widely used technique to estimate model parameters. If $X_1, ..., X_n$ are an independently identically distributed (i.i.d.) sample from a population with p.d.f. or p.m.f $f(x|\theta_1, ..., \theta_k)$, the likelihood function is defined by

$$L(\theta|\boldsymbol{x}) = L(\theta_1, ..., \theta_k|x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} f(x_i|\theta_1, ..., \theta_k).$$

If $\boldsymbol{x}$ is an observed sample, let $\widehat{\boldsymbol{\theta}(\boldsymbol{x})}$ be a parameter value at which maximum value of $L(\theta|\boldsymbol{x})$ is attained as a function of θ, with $\boldsymbol{x}$ held fixed. A maximum likelihood estimator (MLE) of the parameter θ based on a sample $\boldsymbol{X}$ is then $\widehat{\boldsymbol{\theta}(\boldsymbol{x})}$.

We can also mention some properties of MLEs and the usefulness of such properties. MLEs satisfy the properties stated below:

- Consistency,
- Asymptotic normality.

Suppose n is the sample size, $\hat{\theta}$ is a estimate of θ and θ_0 is the true parameter of the sample distribution. Then as $n \to \infty$, if $\hat{\theta} \to \theta_0$, then this $\hat{\theta}$ is a consistent estimate. Moreover if the $\sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{d} N(0, {\sigma_{\theta_0}}^2)$, meaning the difference of estimator and parameter value converges

in distribution to a Normal density function, then $\hat{\theta}$ is said to be asymptotically Normal under regularity conditions. Here, ${\sigma_{\theta_0}}^2$ is called the asymptotic variance of the estimate $\hat{\theta}$. According to asymptotic normality, the estimator converges at a rate of $\frac{1}{\sqrt{n}}$.

Associated with the likelihood function, we can also look at the hessian matrix and the Fisher Information. The hessian matrix of some objective function is the matrix of its second partial derivative, and hence can be written as:

$$H(\theta) = \frac{\delta^2 L(\theta)}{\delta\theta\delta\theta'},$$

where L is the objective function. The hessian matrix above is directly related to the Fisher Information, which, for a random variable X, can be written as:

$$I(\theta) = -E(\frac{\delta^2 L(\theta)}{\delta\theta\delta\theta'}).$$

We call the sample based Fisher information the observed Fisher information. As pointed out in Efron and Hinkley (1978), the observed information can be used to determine the accuracy of the maximum likelihood estimates. In particular, since $\sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{d} N(0, I^{-1})$, where I^{-1} is the inverse of the observed information matrix, the variance (standard errors) can be approximated. The maximum likelihood estimation technique can be applied to all forms of data, including, for instance, censored data. In other words, the likelihood function can be appropriately modified and then maximized.

2.6.2 Expectation maximization algorithm

In order to estimate the MLEs in a missing data scenario, a useful technique is the expectation maximization (EM) algorithm. The EM algorithm converges to the MLEs almost surely (Casella and Berger (2002)). The main idea of the algorithm is to replace one difficult likelihood maximization with a sequence of easier maximizations and the original problem's answer is the limit of such maximizations. In this case, we are interested in solving the "incomplete data" problem but we actually address and solve the "complete-data problem". To incorporate the observations in an incomplete data setup, the EM algorithm was introduced by Dempster et al. (1977) as an iterative approach for computing MLEs. The algorithm estimates the parameters in such a way that involves two steps: first is the expectation step and the next step is the maximization step. Therefore, the process is referred to as Expectation-Maximization or EM algorithm.

Using the notation of Chauveau (1995), let v be complete data and x denote incomplete data, where the incomplete data is assumed to be what is observed. They denote the complete data density by $g(v|\varphi)$ and the incomplete data density by $f(x|\varphi)$, where $\varphi \in \Phi$ and Φ was the parameter space. The objective is to estimate φ in such a way that maximizes the incomplete log likelihood function, i.e.,

$$\underset{\varphi \in \Phi}{argmax}\ log\{f(x|\varphi)\}.$$

However, such maximization is very difficult due to the missingness in data. Using the current parameter value and the observed data value x, a conditional expectation of the unknown complete data log likelihood $log\{g(v|\varphi)\}$ can be done iteratively. In each iteration, such expectation is maximized.

Suppose $h(v|x,\varphi)$ is the conditional density of v given x, where

$$h(v|x,\varphi) = \frac{g(v|\varphi)}{f(x|\varphi)}$$

Then the objective function and the observed likelihood are:

$$Q(\varphi|\varphi') = E[log\{g(v|\varphi)\}|x,\varphi']$$

and

$$L(\varphi) = log\{f(x|\varphi)\},$$

respectively. The EM steps are:

E-step: Compute the density $h(v|x,\varphi^{(n)})$, where $\varphi^{(n)}$is the value of φ at n^{th} iteration.

M-step: Choose $\varphi^{(n+1)} \in \underset{\varphi\in\Phi}{argmax}\ Q(\varphi|\varphi^{(n)})$.

The EM steps continue until convergence of the sequence of $\varphi^{(n)}$ has been attained.

As pointed out in Diebolt and Ip (1994), many inference problems can be solved by the EM algorithm when they are formulated as missing value problems. The authors also point out, however, that the EM algorithm has shortcomings. For instance, it may converge to a local maxima or minima, and it is often seen to be sensitive to initial values. In Park (2005), the author described a technique to handle incomplete data in a competing risk scenario. Considering incompleteness to have arisen from either censoring and/or masking, the author applied the EM algorithm.

An alternative to the EM algorithm, as introduced in Celeux and Diebolt (1985), is the stochastic EM algorithm. It has been noted in the literature that this algorithm is particularly useful when the EM algorithm fails or is intractable. The main principle of this alternative technique is to impute missing data from a specified conditional distribution. That is, upon completion of the E-step where this distribution is determined, missing values are randomly generated from this distribution. This is called the S-step. From here, the dataset is referred

to as a pseudo-complete sample, and the MLEs can be computed. Hence, this procedure is often denoted by SEM. Over the years, the SEM has been used in a variety of inference problems. Dejardin and Lesaffre (2013) proposed the use of SEM for doubly interval-censored data and similarly, Zhang et al. (2013) demonstrated the use and efficiency of SEM in the case of progressively censored data. More recently, the algorithm was applied to reliability problems in Yang et al. (2016), where the authors considered system lifetime data.

2.6.3 Model discrimination

To discriminate between models, one can use Akaike's information criterion (AIC), Bayesian information criterion (BIC) and also the likelihood ratio test (LRT) to choose the best fit between two or more models.

AIC is measured by *Kullback-Leibler divergence* between the restricted density and the unrestricted density. So, there needs to be at least two likelihoods, one corresponding to a generalized model and another is obtained by putting constraints on the parameters of that generalized model. The latter is called the restricted likelihood and the former the unrestricted likelihood. Here, the AIC is calculated using the formula:

$$AIC = -2l + 2k,$$

where l is the maximized log-likelihood value of the model and k is the number of estimated parameters in the model. So, one can compute AIC for each model and compare values. The model with lower AIC has less missing information and hence is said to provide a better fit.

Similarly, the BIC can be calculated as:

$$BIC = -2l + klog(n),$$

where l is the maximized log-likelihood value of the model, k is the number of estimated parameters in the model and n is the sample size. Among the values of BIC, the one with lowest value is the best fitted model.

The likelihood ratio test considers the two models under study by taking their log-

likelihood values into account. In the null hypothesis, we assume that the lifetime distribution can be described by the restricted model and in the alternative hypothesis we say that the lifetime distribution can be described by the unrestricted model other than the one specified in the null hypothesis. If $\hat{l_0}$ is the restricted maximized log-likelihood value and $\hat{l}$ is the unrestricted maximized log-likelihood value, the formula to calculate the LRT test statistic is:

$$\Lambda = -2(\hat{l_0} - \hat{l}).$$

This test statistic is asymptotically distributed as $\chi^2(k)$, where k is the difference in the number of parameters between the two models. Therefore, one makes a conclusion whether to reject the null hypothesis or not by comparing the computed test statistic to percentiles of the appropriate χ^2 distribution.

Chapter 3

Proposed model

3.1 Model construction

In this chapter, we propose a cure rate model where the lifetimes have generalized exponential distribution. Moreover, we consider a competing risk scenario where the number of the competing risks follows a Poisson distribution and each lifetime is subject to right censoring. The whole population of interest is considered to be composed of two groups: cured and susceptibles (non-cured) and our model is developed in a competing risk scenario where we incorporate the existence of these separate groups.

Let M be the number of competing causes. We denote the p.m.f for this M as p_m, with $m = 0, 1, 2, \ldots$ We denote the lifetime due to j^{th} competing risk as W_j, where $j = 1, 2, \ldots,$. There is a common distribution function for all the W_j lifetimes, defined as $F(y) = 1 - S(y)$, where $S(y)$ is the survival function. We assume that both the competing causes M and the lifetime W_j associated to a particular cause are unobservable in the scenario of a competing causes. We assume to observe Y as:

$$Y = min\{W_0, W_1, W_2, ..., W_M\},$$

which can be interpreted as the minimum lifetime amongst the causes and includes W_0, where W_0 is such that $P[W_0 = \infty] = 1$. Inclusion of W_0 gives rise to a proportion of the population

who are not susceptible to the event of interest and hence can be considered cured. From hereon in, we refer to this proportion as the cure rate.

Under our proposed model, the probability mass function $p_m = P(M = m)$ for $m = 0, 1, 2, \ldots, m$ since $M \sim$ Poisson, can be expressed as:

$$p_m = P(M = m) = \frac{e^{-\eta}\eta^m}{m!}, \tag{3.1}$$

where η is the model parameter denoting the average number of competing causes. If there is no competing risk, then we obtain the cure rate, denoted as p_0, from the (3.1) by considering $M = 0$, i.e.,

$$p_0 = P(M = 0) = e^{-\eta}. \tag{3.2}$$

As noted in Chapter 2, in (2.1),

$$S_p(y) = \sum_{m=0}^{\infty} P(M = m)[S(y)]^m. \tag{3.3}$$

In our case, since the number of competing causes has a Poisson distribution with mean η, we then have:

$$\begin{aligned} S_p(y) &= \sum_{m=0}^{\infty} \frac{e^{-\eta}\eta^m[S(y)]^m}{m!} \\ &= e^{-\eta} \sum_{m=0}^{\infty} \frac{[\eta S(y)]^m}{m!} \\ &= e^{\eta S(y)} e^{-\eta} \\ &= e^{-\eta(1-S(y))}. \end{aligned} \tag{3.4}$$

Note, in this case, $S_p(y)$ is not a proper survival function since $\lim_{y\to\infty} S_p(y) = e^{-\eta} = p_0$ (i.e., not 0).

Here, as mentioned, we consider a situation where the lifetimes may not be completely observed and are subject to right censoring. Suppose, for the i^{th} lifetime, the censoring time is a random, call it C_i. Now, with cure rate p_0 and overall censoring proportion p, each lifetime will have its own censoring time. We assume that each C_i follows an exponential distribution with parameter γ, i.e., $C_i \sim exp(\gamma)$. Now, suppose there are n individuals. For each individual, we observe $T_i = min\{Y_i, C_i\}$. In order to generate data from our model, γ needs to be determined numerically. To find γ, and to ensure a cure proportion p_0, we take the ratio of susceptible proportion amongst the censored group with the total proportion of susceptible and equate it to the conditional probability of a subject living longer than the censored time, when at least 1 competing cause is present. That is:

$$\frac{p - p_0}{1 - p_0} = Pr[min(W_1, \ldots, W_M) > C | M > 0]. \tag{3.5}$$

To find γ, this implies we need to solve (3.5). In order to solve this under our model, we substitute in the appropriate quantities and (3.5) becomes:

$$\frac{p - p_0}{1 - p_0} = \frac{Pr[min(W_1, \ldots, W_M > C, M > 0)]}{Pr[M > 0]}$$

$$\frac{p - p_0}{1 - p_0} = \frac{\sum_{m=1}^{\infty} Pr[min(W_1, \ldots, W_M) > C] Pr[M = m]}{1 - Pr[M = 0]}$$

$$\frac{p - p_0}{1 - p_0} = \frac{1}{1 - p_0} \sum_{m=1}^{\infty} \frac{e^{-\eta} \eta^m}{m!} \int_0^{\infty} S(x)^m \gamma e^{-\gamma x} dx. \tag{3.6}$$

Let $t = \gamma x$ and so, $x = \frac{t}{\gamma}$. Then, the right hand side of (3.6) becomes:

$$\frac{1}{1-p_0}\sum_{m=1}^{\infty}\frac{e^{-\eta}\eta^m}{m!}\int_0^{\infty}(S(\frac{t}{\gamma}))^m e^{-t}dt = \frac{1}{1-p_0}\sum_{m=1}^{\infty}\frac{e^{-\eta}\eta^m}{m!}E(S(\frac{t}{\gamma}))^m$$

$$\approx \frac{1}{1-p_0}\sum_{m=1}^{\infty}\frac{e^{-\eta}\eta^m}{m!}\left[\frac{1}{N}\sum_{i=1}^{N}(S(\frac{t_i}{\gamma}))^m\right].$$

As can be seen, we approximate the expectation with a Monte Carlo approximation based on N simulations. For our purposes, we chose N=1000. Since $p_0 = e^{-\eta}$, then $\eta = -log(p_0)$. Rearranging the previous expression, the R.H.S of (3.6) becomes:

$$\frac{1}{1-p_0}\sum_{m=1}^{\infty}\frac{e^{-\eta}\eta^m}{m!}\int_0^{\infty}(S(\frac{t}{\gamma}))^m e^{-t}dt = \frac{1}{1-p_0}e^{-\eta}\frac{1}{N}\sum_{i=1}^{N}\sum_{m=1}^{\infty}\frac{\eta^m}{m!}[(S(\frac{t_i}{\gamma}))^m]$$

$$= \frac{1}{1-p_0}\frac{e^{-\eta}}{N}\left[\sum_{i=1}^{N}\left[\sum_{m=0}^{\infty}\frac{(\eta S(\frac{t_i}{\gamma}))^m}{m!}\right]-1\right]$$

$$= \frac{1}{1-p_0}\frac{e^{-\eta}}{N}\sum_{i=1}^{N}(e^{\eta S(\frac{t_i}{\gamma})}-1)$$

$$= \frac{p_0}{1-p_0}\frac{1}{N}\sum_{i=1}^{N}(e^{-log(p_0)S(\frac{t_i}{\gamma})}-1)$$

$$= \frac{p_0}{1-p_0}[\frac{1}{N}\sum_{i=1}^{N}p_0^{-S(\frac{t_i}{\gamma})}-1].$$

Now by equating the L.H.S and R.H.S of (3.6), we get:

$$\frac{p-p_0}{1-p_0} = \frac{p_0}{1-p_0}[\frac{1}{N}\sum_{i=1}^{N} {p_0}^{-S(\frac{t_i}{\gamma})} - 1]$$

$$\Rightarrow \frac{p-p_0}{p_0} = \frac{1}{N}\sum_{i=1}^{N} {p_0}^{-S(\frac{t_i}{\gamma})} - 1$$

$$\Rightarrow \frac{p}{p_0} - 1 = \frac{1}{N}\sum_{i=1}^{N} {p_0}^{-S(\frac{t_i}{\gamma})} - 1$$

$$\Rightarrow \frac{Np}{p_0} = \sum_{i=1}^{N} {p_0}^{-(1-(1-e^{-\lambda\frac{t_i}{\gamma}})^{\alpha})}$$

$$\Rightarrow \frac{Np}{p_0} = \sum_{i=1}^{N} {p_0}^{-1+(1-e^{-\lambda\frac{t_i}{\gamma}})^{\alpha}}$$

$$\Rightarrow \frac{Np}{p_0} = \sum_{i=1}^{N} {p_0}^{-1}{p_0}^{(1-e^{-\lambda\frac{t_i}{\gamma}})^{\alpha}}$$

$$\Rightarrow \frac{Np}{p_0} = \frac{1}{p_0}\sum_{i=1}^{N} {p_0}^{(1-e^{-\lambda\frac{t_i}{\gamma}})^{\alpha}}$$

$$\Rightarrow \sum_{i=1}^{N} p = \sum_{i=1}^{N} {p_0}^{(1-e^{-\lambda\frac{t_i}{\gamma}})^{\alpha}}$$

$$\Rightarrow \sum_{i=1}^{N} [{p_0}^{(1-e^{-\lambda\frac{t_i}{\gamma}})^{\alpha}} - p] = 0. \tag{3.7}$$

Using (3.7), by applying Newton-Raphson's iterative root finding technique, we can solve for γ. Therefore to generate data, one only needs to fix α, λ, p_0 and p, where of course p_0 and p need to be chosen appropriately, since, for a given group, p_0 must always less than p.

We introduce the covariate effect in the proposed model by relating cure fraction p_0 to covariates $\boldsymbol{x}$. We do this by setting $p_{0i} = \frac{1}{exp(exp(\boldsymbol{x'\beta}))}$,where $\boldsymbol{\beta = (\beta_0, \ldots, \beta_k)'}$ and k is the

number of covariates. Here, the log-linear link $exp(\boldsymbol{x'\beta})$ ranges from$-\infty$ to ∞ and after taking exponent of the negative of this term, it ranges from 0 to ∞. Therefore, this link is chosen to capture the parameter space of the Poisson distribution.

3.2 Likelihood construction

In this section, we use the quantities found in Chapters 2 and Section 3.1 to construct the likelihood function under our proposed model. To consider our population as a mixture population, we introduce an indicator variable I. If the subject is cured, I is to take the value 0 and it will take the value 1 if the subject is susceptible. Also, we denote the probability of being cured as p_0, so, $P[I = 0] = p_0$ and $P[I = 1] = 1 - p_0$. Suppose, the cumulative density of overall population is $F_p(y)$ and the cumulative density for susceptible is $F_1(y)$. Moreover, S_p and S_1 are the corresponding survival functions for overall population and the susceptible group. Then:

$$F_p(y) = (1 - p_0)F_1(y)$$

and

$$S_p(y) = p_0 + (1 - p_0)S_1(y), \tag{3.8}$$

$\lim_{y\to\infty} S_1(y) = 0$ and it is clear we can solve (3.8) for S_1 by using (3.4). Then, using the expression for $S_p(y)$ in (3.4), we have:

$$\begin{aligned} f_p(y) &= -S_p'(y) \\ &= \eta\alpha\lambda e^{-\lambda y}(1 - e^{-\lambda y})^{\alpha-1}e^{-\eta(1-e^{-\lambda y})^{\alpha}}. \end{aligned} \tag{3.9}$$

To clarify which observations are censored, an indicator variable δ_i is introduced, where $\delta_i = I(Y_i \leq C_i)$. Clearly, δ_i takes the value 1 if Y_i is a lifetime, and takes 0 if it is right censored, for $i = 1, \ldots, n$ where n is the sample size. Our data then consists of pairs of

times and censoring indicators, which we denote by $(t_1, \delta_1), (t_2, \delta_2), \ldots, (t_n, \delta_n)$. Therefore, the observed data likelihood function under our proposed model is:

$$\begin{aligned} L(\theta; t, x, \delta) &\propto \prod_{i=1}^{n} \{f_p(t_i, \boldsymbol{x_i}; \boldsymbol{\theta})^{\delta_i}\}\{S_p(t_i, \boldsymbol{x_i}; \boldsymbol{\theta})\}^{(1-\delta_i)} \\ &= \prod_{I_1} \eta\alpha\lambda e^{-\lambda t_i}(1 - e^{-\lambda t_i})^{\alpha-1} e^{-\eta(1-e^{-\lambda t_i})^{\alpha}} \prod_{I_0} e^{-\eta(1-e^{-\lambda t_i})^{\alpha}} \\ &= \prod_{I_1} e^{x'\beta}\alpha\lambda e^{-\lambda t_i}(1 - e^{-\lambda t_i})^{\alpha-1} e^{-e^{x'\beta}(1-e^{-\lambda t_i})^{\alpha}} \prod_{I_0} e^{-e^{x'\beta}(1-e^{-\lambda t_i})^{\alpha}}. \end{aligned} \tag{3.10}$$

Here, $\theta = (\beta', \alpha, \lambda)', t = (t_1, \ldots, t_n)', \delta = (\delta_1, \ldots, \delta_n)', I_1 = \{i : \delta_i = 1\}, I_0 = \{i : \delta_i = 0\}$, and the vector consisting of x_i values is denoted as $\boldsymbol{x}$. Now the log-likelihood becomes:

$$\begin{aligned} Log(L) &= \sum_{I_1} log[e^{x'\beta}\alpha\lambda e^{-\lambda t_i}(1 - e^{-\lambda t_i})^{\alpha-1} e^{-e^{x'\beta}(1-e^{-\lambda t_i})^{\alpha}}] \sum_{I_0} log[e^{-e^{x'\beta}(1-e^{-\lambda t_i})^{\alpha}}] \\ &= \sum_{I_1} [x'\beta + log(\alpha) + log(\lambda) - \lambda t_i + (\alpha - 1)log(1 - e^{-\lambda t_i}) - e^{x'\beta}(1 - e^{-\lambda t_i})^{\alpha}] \\ &- \sum_{I_0} e^{x'\beta}(1 - e^{-\lambda t_i})^{\alpha}. \end{aligned} \tag{3.11}$$

Having an expression for the likelihood (and log-likelihood) function, a natural next step is to develop maximum likelihood estimators(MLEs). In some cases, closed form expressions for MLEs can be obtained. In other cases, such as ours, one must turn to numerical methods and two such methods are discussed in the next chapter.

Chapter 4

Likelihood inference

In this chapter, we describe two numerical methods of estimation to find MLEs of the proposed model's parameters. The first is the direct optimization technique which takes the observed likelihood and maximizes it directly. The second approach is the EM algorithm, where we treat the data as an incomplete data problem. The details are discussed in the next two subsections.

4.1 Direct optimization

Our goal is to obtain values of the parameter which maximize (3.10) and after taking logarithm the log-likelihood function in (3.11). In order to maximize (3.11) numerically in R, initial values of the parameters are required. To obtain initial values, we carry out a grid search to find which values from the grid that maximize (3.11). Once initial values are found, we choose to use the function optim, which is an inbuilt function in R. Since two of our parameters are constrained (to the positive real line), we specify the method "L-BFGS-B" which allows box constraints as developed by Byrd et al. (1995). For a given dataset and associated log likelihood function, with appropriate initial values, optim, among other things, provides estimated parameters, the value of the log-likelihood functions at there parameter values and a convergence code. These estimated parameter values are hence the so calculated MLEs of

interest.

4.2 EM algorithm

As described in Chapter 3, data arising from our proposed model consist in part of censored observations. Among these censored observations, it is not known whether the subject is cured or susceptible. If we let I be the indicator variable taking the value 1 if the lifetime is observed, clearly this indicates I is unknown amongst censored observations. This indicates this is a missing data problem and hence EM the algorithm can be used. With this, the complete log-likelihood function is:

$$l_c(\boldsymbol{\theta; t, x, \delta, I}) = \sum_{I_1} I_i log f_p(t_i, \boldsymbol{x_i; \theta}) + \sum_{I_0} (1 - I_i) log p_0(\boldsymbol{\beta, x_i}) + \sum_{I_0} I_i log\{S_p(t_i, \boldsymbol{x_i; \theta}) - p_0(\boldsymbol{\beta, x_i})\}. \tag{4.1}$$

As described in Section (2.6.2), we need an E-step and M-step. For our proposed model, according to Balakrishnan and Pal (2016) and Pal and Balakrishnan (2016), these steps are as follows.

E-step: We compute the expected value of the complete data log-likelihood function with respect to the distribution of $I_i's$ given a current set of parameter values and the observed data. It is clear that $I_i's$ are Bernoulli random variables and we simply need to calculate ${\pi_i}^{(k)} = E(I_i|\theta^{(k)}, \boldsymbol{O}), i = 1, 2, \ldots, n,$ where $\boldsymbol{O} = \{$ observed $I_i's, (t_i, \boldsymbol{x_i, \delta_i})\}, \boldsymbol{\theta} = (\beta', \alpha, \lambda)$ and $\theta^{(k)}$ denotes the present value of parameters at the k^{th} iteration.

For $i \in I_0$, then:

$$\begin{aligned}
\pi_i{}^{(k)} &= P[I_i = 1 | T_i > t_i; \boldsymbol{\theta}^{(k)}] \\
&= \frac{P[T_i > t_i | I_i = 1] P[I_i = 1]}{p[T_i > t_i]} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(k)}} \\
&= \frac{(1 - p_0(\boldsymbol{\beta}, \boldsymbol{x_i})) S_1(t_i, \boldsymbol{x_i}; \boldsymbol{\theta})}{S_p(t_i, \boldsymbol{x_i}; \boldsymbol{\theta})} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(k)}} \\
&= \frac{S_p(t_i, \boldsymbol{x_i}; \boldsymbol{\theta}) - P_0(\boldsymbol{\beta}, \boldsymbol{x_i})}{S_p(t_i, \boldsymbol{x_i}; \boldsymbol{\theta})} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(k)}} \\
&= \frac{e^{-\eta(1-s(t_i))} - e^{-\eta}}{e^{-\eta(1-S(t_i))}} \\
&= w_i{}^{(k)}. \qquad (4.2)
\end{aligned}$$

For $i \in I_1$, we have $\pi_i{}^{(k)} = I_i = 1$. In the E-step, the $I_i's$ in (4.1) are substituted by $w_i{}^{(k)}$ for $i \in I_0$ and by 1 for $i \in I_1$. We denote the conditional expectation of the complete data log-likelihood function by $Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)})$, where $\boldsymbol{\pi}^{(k)}$ is the vector of $\boldsymbol{\pi_i}^{(k)}$ values. Under our model, the $Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)})$ function can be expressed as:

$$\begin{aligned}
Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)}) =& \sum_{I_1} [log\{exp(\boldsymbol{x'\beta})\} + log(\alpha\lambda) - \lambda t_i + (\alpha - 1) log(1 - exp(-\lambda t_i)) \\
& - exp(\boldsymbol{x'\beta}) + A] + \sum_{I_0} \{-exp(\boldsymbol{x'\beta})\} + \sum_{I_0} [-exp(\boldsymbol{x'\beta}) + log(exp(A) - 1)] \\
=& n_1 log(\alpha\lambda) + \sum_{I_1} \boldsymbol{x'\beta} - \lambda \sum_{I_1} t_i + \sum_{I_1} A + (\alpha - 1) \sum_{I_1} log(1 - exp(-\lambda t_i) \\
& + \sum_{I_0} \pi_i{}^{(k)} log(exp(A) - 1) - \sum_{I^*} exp(\boldsymbol{x'\beta}), \qquad (4.3)
\end{aligned}$$

where $A = \eta S(t_i) = exp(\boldsymbol{x'\beta}) S(t_i)$ and $I^* = I_0 \cup I_1$.

M-step: In this step, we maximize the $Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)})$ function with respect to θ over the parameter space $\boldsymbol{\Theta}$, given $\boldsymbol{\pi}^{(k)}$. This leads to an improved estimate of $\boldsymbol{\theta}$ given as:

$$\boldsymbol{\theta}^{(k+1)} = \underset{\boldsymbol{\theta} \in \boldsymbol{\Theta}}{argmax}\ Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)}).$$

The E-step and M-step are continued until convergence has been established according to a pre-specified criterion. Once convergence has occurred, the current value of θ are the MLEs.

Chapter 5

Numerical results

5.1 Simulation study

In order to determine the performance of our proposed model, we rigorously carried out a Monte Carlo Simulation study. We considered three different sample sizes: n= 100, 200 and 400. For each sample, we introduced the covariate effect by splitting the entire sample into four groups. For each group, for simplicity and as done in Pal and Balakrishnan (2016), the patients of j^{th} group are assigned a covariate value of j, for $j = 1, 2, 3, 4$. With only one covariate, our model has two regression parameters, β_0 and β_1. Using specific values for cure fraction p_0 and censoring proportion p, we can obtain the estimated values of these regression parameters. In order to calculate the values of the regression parameters, we exploit their relationship to p_0. Since, each p_{0i} is a monotone decreasing function of covariate x, we can see the inverse relationship from the link function. $p'_{0i}s$ decrease from group 1 to group 4 as covariates takes value from 1 to 4. We considered two choices for the cure rates for groups 1 and 4. For group 1 we used 0.65 and 0.40, and similarly for group 4, we used 0.25 and 0.15. In short, p_0 are monotonically deceasing in order to reflect the smaller chance of cure as the category goes from 1 to 4. For each group we also fixed the censoring proportion, p. We imposed one of three censoring settings: "High"=(0.85, 0.65, 0.50, 0.35), "Moderate"=(0.65, 0.55, 0.45, 0.35) and "Low"=(0.5, 0.4, 0.3, 0.2). Note, within each group, for a fixed value of

p, this allows for the calculation of the censoring parameter γ.

After fixing p_0 and p, to generate a lifetime, we first generate a censoring time C and a value of our Poisson random variable M, where $\eta = -log(p_0)$. In the special case that $M = 0$, this implies there are no competing risks and so the lifetime is unobservable and so we take the lifetime to be C. In all other cases we generate $W_1, \ldots, W_M$ from the GED as given in (3.1). The observed lifetime is then $T = min\{Y, C\}$, where Y is the minimum of lifetimes $W_1, \ldots, W_M$.

For the lifetime distribution, we considered two parameter settings: $(\alpha, \lambda) = (2, 3)$ and $(\alpha, \lambda) = (4, 1)$. In the case of the EM algorithm, for a simulated dataset, to initiate the iterative procedure, a grid search was carried out to find a set of parameter values within the parameter space that maximize Q as given in (4.3). For the M-step, like in the direct optimization method, we used R's inbuilt function optim. Moreover, for both methods, with the obtained estimates, the hessian matrix was calculated for each dataset. This allowed us to observe the Fisher Information and hence the standard errors of the estimates. This further allowed us to compute confidence intervals for the estimates. In addition to these calculations, we also calculated the empirical bias and root mean squared errors (RMSE) of the estimates, as well as coverage probabilities of the computed confidence intervals. For all settings considered, 1000 Monte Carlo simulations were run.

In Tables 5.1-5.6, we present the parameter estimation results for the direct optimization method. From these six tables, we can see that the parameters are efficiently estimated and the estimation gets better as the sample size increases. Moreover, the estimates of α display larger bias for the settings $\alpha = 4$ and $\lambda = 1$ when compared to $\alpha = 2$ and $\lambda = 3$. One possible explanation for this is the shape of the distribution since, in this case, compared to the other setting $\alpha = 2$ and $\lambda = 3$, there exists a larger portion of longer lifetimes.

These are followed by the estimates of the cure proportions using direct optimization in Tables 5.7 to 5.12. It is evident from these tables that the estimates of the cured proportion are mostly negatively biased. Also, the estimation improves as the sample size increases.

Another important observation is the estimates are almost same for both parameter settings.

Subsequently, in Tables 5.13 to 5.24, we have the analogous results using the EM algorithm. It is evident from these tables that the EM algorithm converges almost perfectly to the true parameter in larger sample sizes. Though the estimates have larger bias for small sample sizes, the estimates improve as n gets larger when we begin to observe smaller standard errors and see a gradual decrease in RMSE. The method seems to work best in the settings with low censoring time and cure proportion $p_{01} = 0.40$ and $p_{04} = 0.15$. Interestingly, the estimates of α give higher bias for $\alpha = 4$ and $\lambda = 1$ than lambda, whereas, λ gives higher bias than α in the $\alpha = 2$ and $\lambda = 3$ setup as censoring time decreases.

Unlike the direct method, estimates of the cured proportion using EM algorithm have very small biases irrespective of sample size. Noticeably, setups with high censoring proportion and cure proportion $p_{01} = 0.40$ and $p_{04} = 0.15$ for 1st and 4th group tend to have larger bias in all six tables. Analogous to the direct method, the estimates are almost same for both parameter settings.

We also observed the number of iterations per 1000 simulations for the EM algorithm. The observed average number of iterations ranged from 3 iterations to 16 iterations. We suspect that due to the large value of elpsilon (0.001) we have a small number of iterations. If the accuracy is increased, that is, epsilon is significantly small, we believe more iterations would be required.

Table 5.1: Estimates, bias and root mean square error(RMSE) and coverage probabilities(CP) for cure rate model using direct optimization method under different simulation settings with $\alpha = 2, \lambda = 3, n = 100$

n	p	Parameter	Estimate(S.E.)	Bias	RMSE	95% CP	90% CP
100(30,40,20,10)	High	α= 2.0000	2.1991(0.5309)	0.1991	0.6294	0.965	0.929
		λ= 3.0000	3.1826(0.8064)	0.1826	0.8726	0.955	0.905
		β_0=-1.2317	-1.2579(0.5074)	-0.0262	0.5113	0.955	0.908
		β_1= 0.3896	0.3985(0.1834)	0.0089	0.1819	0.943	0.909
100(30,40,20,10)	High	α= 2.0000	2.1688(0.5152)	0.1688	0.6184	0.958	0.909
		λ= 3.0000	3.1837(1.1444)	0.1837	1.2800	0.943	0.892
		β_0=-0.3300	-0.2845(0.6983)	0.0455	0.7190	0.949	0.900
		β_1= 0.2426	0.2559(0.1862)	0.0133	0.1925	0.946	0.897
100(30,40,20,10)	Moderate	α= 2.0000	2.1320(0.4393)	0.1320	0.4592	0.963	0.917
		λ= 3.0000	3.1236(0.8444)	0.1236	0.8527	0.953	0.912
		β_0=-0.3300	-0.3161(0.4105)	0.0139	0.4208	0.941	0.888
		β_1= 0.2426	0.2479(0.1547)	0.0053	0.1636	0.945	0.892
100(30,40,20,10)	Low	α= 2.0000	2.1021(0.3680)	0.1021	0.3793	0.971	0.931
		λ= 3.0000	3.1051(0.5757)	0.1051	0.5872	0.956	0.905
		β_0=-0.3300	-0.3287(0.3302)	0.0013	0.3360	0.948	0.901
		β_1= 0.2426	0.2475(0.1337)	0.0049	0.1384	0.943	0.896

Table 5.2: Estimates, bias and root mean square error(RMSE) and coverage probabilities(CP) using direct optimization method for cure rate model under different simulation settings with $\alpha = 4, \lambda = 1, n = 100$

n	p	Parameter	Estimate(S.E.)	Bias	RMSE	95% CP	90% CP
100(30,40,20,10)	High	α= 4.0000	4.5731(1.3333)	0.5731	1.6567	0.966	0.929
		λ= 1.0000	1.0516(0.2149)	0.0516	0.2312	0.948	0.892
		β_0=-1.2317	-1.2382(0.5012)	-0.0064	0.5022	0.951	0.900
		β_1= 0.3896	0.3928(0.1839)	0.0032	0.1883	0.946	0.899
100(30,40,20,10)	High	α= 4.0000	4.4647(1.2872)	0.4647	1.5669	0.954	0.924
		λ= 1.0000	1.0493(0.2682)	0.0493	0.2876	0.948	0.892
		β_0=-0.3300	-0.3108(0.5429)	0.0192	0.5320	0.958	0.914
		β_1= 0.2426	0.2457(0.1849)	0.0032	0.1831	0.951	0.896
100(30,40,20,10)	Moderate	α= 4.0000	4.2757(1.0532)	0.2757	1.1291	0.957	0.921
		λ= 1.0000	1.0218(0.2089)	0.0218	0.2095	0.958	0.908
		β_0=-0.3300	-0.3313(0.3963)	-0.0013	0.4131	0.949	0.893
		β_1= 0.2426	0.2540(0.1541)	0.0114	0.1606	0.947	0.889
100(30,40,20,10)	Low	α= 4.0000	4.2605(0.8964)	0.2605	1.0049	0.961	0.901
		λ= 1.0000	1.0190(0.1566)	0.0190	0.1633	0.948	0.898
		β_0=-0.3300	-0.3276(0.3287)	0.0024	0.3450	0.945	0.890
		β_1= 0.2426	0.2461(0.1336)	0.0035	0.1392	0.945	0.887

Table 5.3: Estimates, bias and root mean square error(RMSE) and coverage probabilities(CP) for cure rate model using direct optimization method under different simulation settings with $\alpha = 2, \lambda = 3, n = 200$

n	p	Parameter	Estimate(S.E.)	Bias	RMSE	95% CP	90% CP
200(55,60,45,40)	High	α= 2.0000	2.0887(0.3275)	0.087	0.3481	0.958	0.908
		λ= 3.0000	3.0864(0.5252)	0.0864	0.5615	0.944	0.886
		β_0=-1.2317	-1.2269(0.3515)	0.0048	0.3665	0.940	0.889
		β_1= 0.3896	0.3885(0.1138)	-0.0011	0.1166	0.947	0.890
200(55,60,45,40)	High	α= 2.0000	2.0743(0.3184)	0.0743	0.3578	0.948	0.885
		λ= 3.0000	3.0839(0.7075)	0.0839	0.7653	0.941	0.889
		β_0=-0.3300	-0.3348(0.3893)	-0.0048	0.3927	0.948	0.903
		β_1= 0.2426	0.2499(0.1149)	0.0073	0.1144	0.946	0.899
200(55,60,45,40)	Moderate	α= 2.0000	2.0595(0.2865)	0.0595	0.3042	0.952	0.901
		λ= 3.0000	3.0752(0.5712)	0.0752	0.5979	0.944	0.884
		β_0=-0.3300	-0.3275(0.2802)	0.0025	0.2896	0.941	0.895
		β_1= 0.2426	0.2444(0.0943)	0.0018	0.0974	0.947	0.896
200(55,60,45,40)	Low	α= 2.0000	2.0498(0.2457)	0.0498	0.2595	0.958	0.904
		λ= 3.0000	3.0499(0.3953)	0.0499	0.4073	0.938	0.885
		β_0=-0.3300	-0.3399(0.2292)	-0.0099	0.2319	0.951	0.899
		β_1= 0.2426	0.2463(0.0822)	0.0037	0.0840	0.950	0.893

Table 5.4: Estimates, bias and root mean square error(RMSE) and coverage probabilities(CP) for cure rate model using direct optimization method under different simulation settings with $\alpha = 4, \lambda = 1, n = 200$

n	p	Parameter	Estimate(S.E.)	Bias	RMSE	95% CP	90% CP
200(55,60,45,40)	High	α= 4.0000	4.2610(0.8021)	0.2610	0.9194	0.955	0.913
		λ= 1.0000	1.0265(0.1412)	0.0265	0.1453	0.959	0.907
		β_0=-1.2317	-1.2323(0.3483)	-0.0006	0.3431	0.953	0.907
		β_1= 0.3896	0.3911(0.1138)	0.0015	0.1130	0.962	0.907
200(55,60,45,40)	High	α= 4.0000	4.1726(0.7733)	0.1726	0.8562	0.949	0.901
		λ= 1.0000	1.0156(0.1714)	0.0156	0.1771	0.953	0.904
		β_0=-0.3300	-0.3429(0.3702)	-0.0129	0.3829	0.942	0.894
		β_1= 0.2426	0.2520(0.1146)	0.0094	0.1170	0.945	0.896
200(55,60,45,40)	Moderate	α= 4.0000	4.1784(0.6978)	0.1784	0.7607	0.958	0.907
		λ= 1.0000	1.0218(0.1431)	0.0218	0.1430	0.958	0.902
		β_0=-0.3300	-0.3338(0.2735)	-0.0038	0.2744	0.951	0.902
		β_1= 0.2426	0.2429(0.0946)	0.0003	0.0960	0.946	0.893
200(55,60,45,40)	Low	α= 4.0000	4.1198(0.5906)	0.1198	0.6510	0.951	0.893
		λ= 1.0000	1.0116(0.1087)	0.0116	0.1147	0.940	0.892
		β_0=-0.3300	-0.3353(0.2278)	-0.0053	0.2308	0.946	0.893
		β_1= 0.2426	0.2474(0.0820)	0.0048	0.0806	0.952	0.911

Table 5.5: Estimates, bias and root mean square error(RMSE) and coverage probabilities(CP) for cure rate model using direct optimization method under different simulation settings with $\alpha = 2, \lambda = 3, n = 400$

n	p	Parameter	Estimate(S.E.)	Bias	RMSE	95% CP	90% CP
400(110,120,90,80)	High	α= 2.0000	2.0453(0.2247)	0.0453	0.2304	0.959	0.910
		λ= 3.0000	3.0346(0.3661)	0.0346	0.3802	0.938	0.894
		β_0=-1.2317	-1.2327(0.2477)	-0.0009	0.2558	0.946	0.895
		β_1= 0.3896	0.3908(0.0803)	0.0012	0.0810	0.950	0.901
400(110,120,90,80)	High	α= 2.0000	2.0303(0.2172)	0.0303	0.2199	0.953	0.900
		λ= 3.0000	3.0284(0.4914)	0.0284	0.4788	0.954	0.905
		β_0=-0.3300	-0.3221(0.2712)	0.0079	0.2730	0.952	0.902
		β_1= 0.2426	0.2438(0.0807)	0.0012	0.0824	0.947	0.901
400(110,120,90,80)	Moderate	α= 2.0000	2.0337(0.1988)	0.0337	0.2064	0.946	0.904
		λ= 3.0000	3.0233(0.3978)	0.0233	0.4045	0.942	0.907
		β_0=-0.3300	-0.3303(0.1975)	-0.0003	0.2008	0.943	0.895
		β_1= 0.2426	0.2449(0.0667)	0.0024	0.0677	0.943	0.893
400(110,120,90,80)	Low	α= 2.0000	2.0217(0.1701)	0.0217	0.1705	0.961	0.914
		λ= 3.0000	3.0208(0.2770)	0.0208	0.2778	0.943	0.907
		β_0=-0.3300	-0.3355(0.1612)	-0.0055	0.1590	0.949	0.907
		β_1= 0.2426	0.2466(0.0578)	0.0041	0.0582	0.950	0.900

Table 5.6: Estimates, bias and root mean square error(RMSE) and coverage probabilities(CP) for cure rate model using direct optimization method under different simulation settings with $\alpha = 4, \lambda = 1, n = 400$

n	p	Parameter	Estimate(S.E.)	Bias	RMSE	95% CP	90% CP
400(110,120,90,80)	High	α= 4.0000	4.0797(0.5341)	0.0797	0.5480	0.952	0.911
		λ= 1.0000	1.0040(0.0983)	0.0040	0.0983	0.952	0.898
		β_0=-1.2317	-1.2376(0.2455)	-0.0058	0.2520	0.954	0.899
		β_1= 0.3896	0.3929(0.0802)	0.0034	0.0840	0.947	0.883
400(110,120,90,80)	High	α= 4.0000	4.0932(0.5292)	0.0932	0.5280	0.963	0.919
		λ= 1.0000	1.0112(0.1197)	0.0112	0.1143	0.962	0.917
		β_0=-0.3300	-0.3261(0.2586)	0.0039	0.2697	0.940	0.891
		β_1= 0.2426	0.2411(0.0805)	-0.0015	0.0813	0.952	0.902
400(110,120,90,80)	Moderate	α= 4.0000	4.0613(0.4735)	0.0613	0.4773	0.949	0.907
		λ= 1.0000	1.0067(0.1000)	0.0067	0.0986	0.947	0.902
		β_0=-0.3300	-0.3363(0.1927)	-0.0063	0.1946	0.952	0.907
		β_1= 0.2426	0.2444(0.0666)	0.0018	0.0672	0.949	0.898
400(110,120,90,80)	Low	α= 4.0000	4.0584(0.4085)	0.0584	0.4295	0.951	0.904
		λ= 1.0000	1.0064(0.0764)	0.0064	0.0788	0.951	0.892
		β_0=-0.3300	-0.3278(0.1608)	0.0022	0.1581	0.956	0.898
		β_1= 0.2426	0.2422(0.0579)	-0.0004	0.0575	0.947	0.897

Table 5.7: Estimates, bias and root mean square error(RMSE)for cure fraction using direct optimization method under different simulation settings with $\alpha = 2, \lambda = 3, n = 100$

n	p	p_0	Estimate	Bias	RMSE
100(30,40,20,10)	High	p_{01}=0.6500	0.6454	-0.0046	0.0951
		p_{02}=0.5294	0.5291	-0.0003	0.0747
		p_{03}=0.3910	0.3905	-0.0005	0.0768
		p_{04}=0.2500	0.2539	0.0039	0.1064
100(30,40,20,10)	High	p_{01}=0.4000	0.3886	-0.0114	0.1562
		p_{02}=0.3110	0.3022	-0.0088	0.1186
		p_{03}=0.2257	0.2194	-0.0063	0.0984
		p_{04}=0.1500	0.1537	0.0037	0.1013
100(30,40,20,10)	Moderate	p_{01}=0.4000	0.3925	-0.0075	0.1028
		p_{02}=0.3110	0.3040	-0.0070	0.0746
		p_{03}=0.2257	0.2211	-0.0046	0.0771
		p_{04}=0.1500	0.1558	0.0058	0.0936
100(30,40,20,10)	Low	p_{01}=0.4000	0.3971	-0.0029	0.0790
		p_{02}=0.3110	0.3077	-0.0033	0.0538
		p_{03}=0.2257	0.2233	-0.0025	0.0606
		p_{04}=0.1500	0.1546	0.0046	0.0784

Table 5.8: Estimates, bias and root mean square error(RMSE) for cure fraction using direct optimization method under different simulation settings with $\alpha = 4, \lambda = 1, n = 100$

n	p	p_0	Estimate	Bias	RMSE
100(30,40,20,10)	High	p_{01}=0.6500	0.6423	-0.0077	0.0925
		p_{02}=0.5294	0.5267	-0.0027	0.0703
		p_{03}=0.3910	0.3894	-0.0016	0.0784
		p_{04}=0.2500	0.2560	0.0060	0.1128
100(30,40,20,10)	High	p_{01}=0.4000	0.3916	-0.0084	0.1304
		p_{02}=0.3110	0.3059	-0.0052	0.0956
		p_{03}=0.2257	0.2243	-0.0014	0.0854
		p_{04}=0.1500	0.1596	0.0096	0.0984
100(30,40,20,10)	Moderate	p_{01}=0.4000	0.3953	-0.0047	0.1000
		p_{02}=0.3110	0.3047	-0.0064	0.0713
		p_{03}=0.2257	0.2194	-0.0063	0.0727
		p_{04}=0.1500	0.1520	0.0020	0.0878
100(30,40,20,10)	Low	p_{01}=0.4000	0.3972	-0.0028	0.0810
		p_{02}=0.3110	0.3083	-0.0027	0.0527
		p_{03}=0.2257	0.2240	-0.0017	0.0574
		p_{04}=0.1500	0.1551	0.0051	0.0765

Table 5.9: Estimates, bias and root mean square error(RMSE) for cure fraction using direct optimization method under different simulation settings with $\alpha = 2, \lambda = 3, n = 200$

n	p	p_0	Estimate	Bias	RMSE
200(55,60,45,40)	High	p_{01}=0.6500	0.6436	-0.0064	0.0720
		p_{02}=0.5294	0.5267	-0.0027	0.0579
		p_{03}=0.3910	0.3903	-0.0007	0.0502
		p_{04}=0.2500	0.2523	0.0023	0.0633
200(55,60,45,40)	High	p_{01}=0.4000	0.3984	-0.0016	0.1039
		p_{02}=0.3110	0.3094	-0.0016	0.0770
		p_{03}=0.2257	0.2233	-0.0025	0.0600
		p_{04}=0.1500	0.1494	-0.0006	0.0587
200(55,60,45,40)	Moderate	p_{01}=0.4000	0.3978	-0.0022	0.0764
		p_{02}=0.3110	0.3097	-0.0014	0.0570
		p_{03}=0.2257	0.2252	-0.0006	0.0513
		p_{04}=0.1500	0.1527	0.0027	0.0569
200(55,60,45,40)	Low	p_{01}=0.4000	0.4019	0.0019	0.0581
		p_{02}=0.3110	0.3123	0.0013	0.0397
		p_{03}=0.2257	0.2264	0.0006	0.0370
		p_{04}=0.1500	0.1520	0.0020	0.0454

Table 5.10: Estimates, bias and root mean square error(RMSE) and for cure fraction using direct optimization method under different simulation settings with $\alpha = 4, \lambda = 1, n = 200$

n	p	p_0	Estimate	Bias	RMSE
200(55,60,45,40)	High	p_{01}=0.6500	0.6423	-0.0077	0.0925
		p_{02}=0.5294	0.5267	-0.0027	0.0703
		p_{03}=0.3910	0.3894	-0.0016	0.0784
		p_{04}=0.2500	0.2560	0.0060	0.1128
200(55,60,45,40)	High	p_{01}=0.4000	0.3916	-0.0084	0.1304
		p_{02}=0.3110	0.3059	-0.0052	0.0956
		p_{03}=0.2257	0.2243	-0.0014	0.0854
		p_{04}=0.1500	0.1596	0.0096	0.0984
200(55,60,45,40)	Moderate	p_{01}=0.4000	0.3953	-0.0047	0.1000
		p_{02}=0.3110	0.3047	-0.0064	0.0713
		p_{03}=0.2257	0.2194	-0.0063	0.0727
		p_{04}=0.1500	0.1520	0.0020	0.0878
200(55,60,45,40)	Low	p_{01}=0.4000	0.3972	-0.0028	0.0810
		p_{02}=0.3110	0.3083	-0.0027	0.0527
		p_{03}=0.2257	0.2240	-0.0017	0.0574
		p_{04}=0.1500	0.1551	0.0051	0.0765

Table 5.11: Estimates, bias and root mean square error(RMSE) for cure fraction using direct optimization method under different simulation settings with $\alpha = 2, \lambda = 3, n = 400$

n	p	p_0	Estimate	Bias	RMSE
400(110,120,90,80)	High	p_{01}=0.6500	0.6473	-0.0027	0.0508
		p_{02}=0.5294	0.5281	-0.0014	0.0400
		p_{03}=0.3910	0.3900	-0.0010	0.0335
		p_{04}=0.2500	0.2497	-0.0003	0.0428
400(110,120,90,80)	High	p_{01}=0.4000	0.3962	-0.0038	0.0730
		p_{02}=0.3110	0.3080	-0.0030	0.0526
		p_{03}=0.2257	0.2232	-0.0025	0.0411
		p_{04}=0.1500	0.1494	-0.0006	0.0423
400(110,120,90,80)	Moderate	p_{01}=0.4000	0.3990	-0.0010	0.0531
		p_{02}=0.3110	0.3098	-0.0012	0.0390
		p_{03}=0.2257	0.2244	-0.0013	0.0351
		p_{04}=0.1500	0.1499	-0.0001	0.0394
400(110,120,90,80)	Low	p_{01}=0.4000	0.4003	0.0003	0.0404
		p_{02}=0.3110	0.3103	-0.0008	0.0281
		p_{03}=0.2257	0.2240	-0.0017	0.0270
		p_{04}=0.1500	0.1487	-0.0013	0.0329

Table 5.12: Estimates, bias and root mean square error(RMSE) and for cure fraction using direct optimization method under different simulation settings with $\alpha = 4, \lambda = 1, n = 400$

n	p	p_0	Estimate	Bias	RMSE
400(110,120,90,80)	High	p_{01}=0.6500	0.6483	-0.0017	0.0487
		p_{02}=0.5294	0.5284	-0.0010	0.0370
		p_{03}=0.3910	0.3894	-0.0016	0.0322
		p_{04}=0.2500	0.2485	-0.0015	0.0452
400(110,120,90,80)	High	p_{01}=0.4000	0.3985	-0.0015	0.0715
		p_{02}=0.3110	0.3113	0.0002	0.0500
		p_{03}=0.2257	0.2269	0.0012	0.0372
		p_{04}=0.1500	0.1529	0.0029	0.0389
400(110,120,90,80)	Moderate	p_{01}=0.4000	0.4013	0.0013	0.0502
		p_{02}=0.3110	0.3123	0.0013	0.0351
		p_{03}=0.2257	0.2267	0.0010	0.0312
		p_{04}=0.1500	0.1519	0.0019	0.0367
400(110,120,90,80)	Low	p_{01}=0.4000	0.3992	-0.0008	0.0398
		p_{02}=0.3110	0.3107	-0.0004	0.0267
		p_{03}=0.2257	0.2258	0.0001	0.0251
		p_{04}=0.1500	0.1514	0.0014	0.0316

Table 5.13: Estimates, bias and root mean square error(RMSE) and coverage probabilities(CP) for cure rate model using EM under different simulation settings with $\alpha = 2, \lambda = 3, n = 100$

n	p	Parameter	Estimate(S.E.)	Bias	RMSE	95% CP	90% CP
100(30,40,20,10)	High	α= 2.0000	2.1877(0.5305)	0.1877	0.6321	0.960	0.922
		λ= 3.0000	3.1525(0.8089)	0.1525	0.9170	0.934	0.883
		β_0=-1.2317	-1.2414(0.5105)	-0.0096	0.5430	0.939	0.888
		β_1= 0.3896	0.3965(0.1834)	0.0069	0.1846	0.937	0.902
100(30,40,20,10)	High	α= 2.0000	2.0979(0.5108)	0.0979	0.5922	0.956	0.901
		λ= 3.0000	2.8589(1.1689)	-0.1411	1.1867	0.954	0.902
		β_0=-0.3300	-0.1376(0.7084)	0.1924	0.6540	0.964	0.922
		β_1= 0.2426	0.2331(0.1851)	-0.0095	0.1872	0.947	0.899
100(30,40,20,10)	Moderate	α= 2.0000	2.0416(0.4328)	0.0416	0.4969	0.937	0.876
		λ= 3.0000	2.8256(0.8528)	-0.1744	1.0404	0.912	0.808
		β_0=-0.3300	-0.2318(0.4256)	0.0982	0.4436	0.948	0.888
		β_1= 0.2426	0.2460(0.1533)	0.0034	0.1600	0.938	0.886
100(30,40,20,10)	Low	α= 2.0000	2.1143(0.3711)	0.1143	0.4069	0.960	0.909
		λ= 3.0000	3.1193(0.5760)	0.1193	0.5867	0.950	0.902
		β_0=-0.3300	-0.3315(0.3301)	-0.0015	0.3389	0.948	0.907
		β_1= 0.2426	0.2445(0.1336)	0.0019	0.1385	0.946	0.891

Table 5.14: Estimates, bias and root mean square error(RMSE) and coverage probabilities(CP) using EM algorithm for cure rate model under different simulation settings with $\alpha = 4, \lambda = 1, n = 100$

n	p	Parameter	Estimate(S.E.)	Bias	RMSE	95% CP	90% CP
100(30,40,20,10)	High	α= 4.0000	4.5821(1.3358)	0.5821	1.6639	0.966	0.927
		λ= 1.0000	1.0540(0.2150)	0.0540	0.2314	0.948	0.893
		β_0=-1.2317	-1.2465(0.5018)	-0.0147	0.4998	0.949	0.902
		β_1= 0.3896	0.3949(0.1840)	0.0053	0.1878	0.945	0.901
100(30,40,20,10)	High	α= 4.0000	4.3164(1.2606)	0.3164	1.4532	0.946	0.919
		λ= 1.0000	0.9972(0.2708)	-0.0028	0.2721	0.948	0.902
		β_0=-0.3300	-0.1996(0.5435)	0.1305	0.5117	0.957	0.918
		β_1= 0.2426	0.2248(0.1836)	-0.0177	0.1810	0.955	0.896
100(30,40,20,10)	Moderate	α= 4.0000	4.2166(1.0430)	0.2166	1.1039	0.951	0.914
		λ= 1.0000	1.0027(0.2089)	0.0027	0.2153	0.952	0.902
		β_0=-0.3300	-0.3037(0.3959)	0.0263	0.4279	0.943	0.877
		β_1= 0.2426	0.2515(0.1537)	0.0090	0.1617	0.948	0.888
100(30,40,20,10)	Low	α= 4.0000	4.2615(0.8965)	0.2615	1.0032	0.961	0.901
		λ= 1.0000	1.0194(0.1566)	0.0194	0.1625	0.948	0.898
		β_0=-0.3300	-0.3292(0.3289)	0.0009	0.3460	0.945	0.887
		β_1= 0.2426	0.2464(0.1336)	0.0038	0.1394	0.945	0.887

Table 5.15: Estimates, bias and root mean square error(RMSE) and coverage probabilities(CP) using EM algorithm for cure rate model under different simulation settings with $\alpha = 2, \lambda = 3, n = 200$

n	p	Parameter	Estimate(S.E.)	Bias	RMSE	95% CP	90% CP
200(55,60,45,40)	High	α= 2.0000	2.0780(0.3266)	0.0780	0.3436	0.957	0.912
		λ= 3.0000	3.0898(0.5269)	0.0898	0.5293	0.958	0.908
		β_0=-1.2317	-1.2594(0.3535)	-0.0276	0.3645	0.939	0.892
		β_1= 0.3896	0.3950(0.1144)	0.0054	0.1190	0.942	0.882
200(55,60,45,40)	High	α= 2.0000	2.0279(0.3133)	0.0279	0.3289	0.950	0.895
		λ= 3.0000	2.8741(0.7069)	-0.1259	0.7086	0.940	0.902
		β_0=-0.3300	-0.2208(0.3932)	0.1092	0.3567	0.960	0.920
		β_1= 0.2426	0.2321(0.1140)	-0.0105	0.1075	0.969	0.927
200(55,60,45,40)	Moderate	α= 2.0000	1.9733(0.2790)	-0.0267	0.3229	0.906	0.846
		λ= 3.0000	2.7615(0.5665)	-0.2385	0.7526	0.863	0.759
		β_0=-0.3300	-0.2094(0.2836)	0.1206	0.3173	0.922	0.855
		β_1= 0.2426	0.2330(0.0934)	-0.0096	0.0954	0.945	0.897
200(55,60,45,40)	Low	α= 2.0000	2.0437(0.2443)	0.0437	0.2530	0.960	0.912
		λ= 3.0000	3.0481(0.3955)	0.0481	0.4048	0.954	0.907
		β_0=-0.3300	-0.3408(0.2290)	-0.0108	0.2398	0.936	0.898
		β_1= 0.2426	0.2480(0.0821)	0.0054	0.0877	0.935	0.877

Table 5.16: Estimates, bias and root mean square error(RMSE) and coverage probabilities(CP) using EM algorithm for cure rate model under different simulation settings with $\alpha = 4, \lambda = 1, n = 200$

n	p	Parameter	Estimate(S.E.)	Bias	RMSE	95% CP	90% CP
200(55,60,45,40)	High	α= 4.0000	4.2679(0.8032)	0.2679	0.9232	0.955	0.913
		λ= 1.0000	1.0286(0.1413)	0.0286	0.145608	0.959	0.909
		β_0=-1.2317	-1.2408(0.3489)	-0.0090	0.3410	0.959	0.911
		β_1= 0.3896	0.3931(0.1140)	0.0036	0.1126	0.964	0.908
200(55,60,45,40)	High	α= 4.0000	4.0799(0.7615)	0.0799	0.8181	0.942	0.898
		λ= 1.0000	0.9783(0.1717)	-0.0217	0.1716	0.952	0.904
		β_0=-0.3300	-0.2415(0.3669)	0.0886	0.3459	0.949	0.911
		β_1= 0.2426	0.2323(0.1134)	-0.0104	0.1107	0.948	0.909
200(55,60,45,40)	Moderate	α= 4.0000	4.1351(0.6925)	0.1351	0.7432	0.954	0.904
		λ= 1.0000	1.0068(0.1429)	0.0067	0.1460	0.948	0.888
		β_0=-0.3300	-0.3128(0.2730)	0.0172	0.3007	0.931	0.870
		β_1= 0.2426	0.2412(0.0943)	-0.0014	0.09948	0.940	0.879
200(55,60,45,40)	Low	α= 4.0000	4.1201(0.5906)	0.1201	0.6492	0.951	0.894
		λ= 1.0000	1.0118(0.1087)	0.0118	0.1139	0.943	0.897
		β_0=-0.3300	-0.3363(0.2278)	-0.0063	0.2331	0.946	0.887
		β_1= 0.2426	0.2476(0.0820)	0.0050	0.0812	0.949	0.909

Table 5.17: Estimates, bias and root mean square error(RMSE) and coverage probabilities(CP) using EM algorithm for cure rate model under different simulation settings with $\alpha = 2, \lambda = 3, n = 400$

n	p	Parameter	Estimate(S.E.)	Bias	RMSE	95% CP	90% CP
400(110,120,90,80)	High	α= 2.0000	2.0463(0.2247)	0.0463	0.2307	0.960	0.910
		λ= 3.0000	3.0395(0.3663)	0.0395	0.3801	0.940	0.895
		β_0=-1.2317	-1.2371(0.2478)	-0.0053	0.2550	0.947	0.895
		β_1= 0.3896	0.3918(0.0803)	0.0022	0.0808	0.951	0.899
400(110,120,90,80)	High	α= 2.0000	2.0001(0.2156)	0.0001	0.2114	0.950	0.897
		λ= 3.0000	2.8669(0.4908)	-0.1331	0.4485	0.961	0.915
		β_0=-0.3300	-0.2313(0.2718)	0.0987	0.2337	0.968	0.937
		β_1= 0.2426	0.2290(0.0802)	-0.0136	0.0756	0.967	0.927
400(110,120,90,80)	Moderate	α= 2.0000	1.9405(0.1930)	-0.0595	0.2337	0.886	0.814
		λ= 3.0000	2.6897(0.3945)	-0.3103	0.6370	0.730	0.619
		β_0=-0.3300	-0.2039(0.1990)	0.1261	0.2520	0.888	0.795
		β_1= 0.2426	0.2345(0.0657)	-0.0081	0.0670	0.948	0.891
400(110,120,90,80)	Low	α= 2.0000	2.0183(0.1698)	0.0183	0.1730	0.948	0.896
		λ= 3.0000	3.0196(0.2774)	0.0196	0.2831	0.953	0.895
		β_0=-0.3300	-0.3285(0.1612)	0.0015	0.1658	0.943	0.892
		β_1= 0.2426	0.2435(0.0578)	0.0009	0.0599	0.945	0.895

Table 5.18: Estimates, bias and root mean square error(RMSE) and coverage probabilities(CP) using EM algorithm for cure rate model under different simulation settings with $\alpha = 4, \lambda = 1, n = 400$

n	p	Parameter	Estimate(S.E.)	Bias	RMSE	95% CP	90% CP
400(110,120,90,80)	High	α= 4.0000	4.0858(0.5349)	0.0858	0.5498	0.952	0.910
		λ= 1.0000	1.0059(0.0984)	0.0059	0.0983	0.952	0.900
		β_0=-1.2317	-1.2456(0.2459)	-0.0138	0.2510	0.956	0.902
		β_1= 0.3896	0.3949(0.0803)	0.0053	0.0838	0.948	0.883
400(110,120,90,80)	High	α= 4.0000	4.0204(0.5227)	0.0204	0.5350	0.946	0.899
		λ= 1.0000	0.9768(0.1196)	-0.0232	0.1189	0.948	0.907
		β_0=-0.3300	-0.2267(0.2574)	0.1033	0.2322	0.962	0.923
		β_1= 0.2426	0.2217(0.0799)	-0.0209	0.0780	0.949	0.890
400(110,120,90,80)	Moderate	α= 4.0000	4.0208(0.4704)	0.0208	0.5195	0.926	0.861
		λ= 1.0000	0.9912(0.1001)	-0.0088	0.1145	0.910	0.848
		β_0=-0.3300	-0.3026(0.1922)	0.0274	0.2331	0.915	0.830
		β_1= 0.2426	0.2403(0.0663)	-0.0022	0.0712	0.941	0.877
400(110,120,90,80)	Low	α= 4.0000	4.0761(0.4112)	0.0761	0.4367	0.949	0.889
		λ= 1.0000	1.0067(0.0764)	0.0067	0.0804	0.950	0.887
		β_0=-0.3300	-0.3406(0.1611)	-0.0106	0.1656	0.951	0.901
		β_1= 0.2426	0.2467(0.0580)	0.0041	0.0603	0.950	0.900

Table 5.19: Estimates, bias and root mean square error(RMSE) using EM algorithm for cure fraction under different simulation settings with $\alpha = 2, \lambda = 3, n = 100$

n	p	p_0	Estimate	Bias	RMSE
100(30,40,20,10)	High	p_{01}=0.6500	0.6400	-0.0100	0.1088
		p_{02}=0.5294	0.5245	-0.0049	0.0875
		p_{03}=0.3910	0.3870	-0.0040	0.0836
		p_{04}=0.2500	0.2516	0.0016	0.1077
100(30,40,20,10)	High	p_{01}=0.4000	0.3450	-0.0550	0.1419
		p_{02}=0.3110	0.2650	-0.0460	0.1061
		p_{03}=0.2257	0.1930	-0.0327	0.0932
		p_{04}=0.1500	0.1394	-0.0106	0.0985
100(30,40,20,10)	Moderate	p_{01}=0.4000	0.3634	-0.0366	0.1140
		p_{02}=0.3110	0.2767	-0.0343	0.0906
		p_{03}=0.2257	0.1983	-0.0275	0.0877
		p_{04}=0.1500	0.1383	-0.0117	0.0937
100(30,40,20,10)	Low	p_{01}= 0.4000	0.3991	-0.0009	0.0797
		p_{02}=0.3110	0.3109	-0.0001	0.0544
		p_{03}=0.2257	0.2272	0.0015	0.0611
		p_{04}=0.1500	0.1585	0.0085	0.0803

Table 5.20: Estimates, bias and root mean square error(RMSE) using EM algorithm for cure fraction under different simulation settings with $\alpha = 4, \lambda = 1, n = 100$

n	p	p_0	Estimate	Bias	RMSE
100(30,40,20,10)	High	p_{01}=0.6500	0.6441	-0.0059	0.0914
		p_{02}=0.5294	0.5281	-0.0013	0.0696
		p_{03}=0.3910	0.3901	-0.0009	0.0782
		p_{04}=0.2500	0.2559	0.0059	0.1127
100(30,40,20,10)	High	p_{01}=0.4000	0.3597	-0.0403	0.1244
		p_{02}=0.3110	0.2811	-0.0300	0.0907
		p_{03}=0.2257	0.2087	-0.0170	0.0830
		p_{04}=0.1500	0.1529	0.0029	0.0964
100(30,40,20,10)	Moderate	p_{01}= 0.4000	0.3862	-0.0138	0.1060
		p_{02}=0.3110	0.2968	-0.0142	0.0776
		p_{03}=0.2257	0.2133	-0.0124	0.0761
		p_{04}=0.1500	0.1478	-0.0022	0.0881
100(30,40,20,10)	Low	p_{01}=0.4000	0.3976	-0.0024	0.0812
		p_{02}=0.3110	0.3087	-0.0024	0.0527
		p_{03}=0.2257	0.2242	-0.0015	0.0574
		p_{04}=0.1500	0.1552	0.0052	0.0766

Table 5.21: Estimates, bias and root mean square error(RMSE) using EM algorithm for cure fraction under different simulation settings with $\alpha = 2, \lambda = 3, n = 200$

n	p	p_0	Estimate	Bias	RMSE
200(55,60,45,40)	High	p_{01}= 0.6500	0.6511	0.0011	0.0695
		p_{02}= 0.5294	0.5334	0.0040	0.0546
		p_{03}= 0.3910	0.3950	0.0040	0.0483
		p_{04}= 0.2500	0.2544	0.0044	0.0652
200(55,60,45,40)	High	p_{01}= 0.4000	0.3641	-0.0359	0.0959
		p_{02}= 0.3110	0.2812	-0.0299	0.0710
		p_{03}= 0.2257	0.2032	-0.0225	0.0575
		p_{04}= 0.1500	0.1378	-0.0122	0.0565
200(55,60,45,40)	Moderate	p_{01}= 0.4000	0.3592	-0.0408	0.0898
		p_{02}= 0.3110	0.2760	-0.0350	0.0713
		p_{03}= 0.2257	0.1986	-0.0272	0.0614
		p_{04}= 0.1500	0.1339	-0.0161	0.0587
200(55,60,45,40)	Low	p_{01}= 0.4000	0.4016	0.0016	0.0595
		p_{02}= 0.3110	0.3114	0.0004	0.0398
		p_{03}= 0.2257	0.2250	-0.0008	0.0371
		p_{04}= 0.1500	0.1505	0.0005	0.0466

Table 5.22: Estimates, bias and root mean square error(RMSE) using EM algorithm for cure fraction under different simulation settings with $\alpha = 4, \lambda = 1, n = 200$

n	p	p_0	Estimate	Bias	RMSE
200(55,60,45,40)	High	p_{01}=0.6500	0.6471	-0.0029	0.0654
		p_{02}=0.5294	0.5286	-0.0008	0.0512
		p_{03}=0.3910	0.3903	-0.0008	0.0456
		p_{04}=0.2500	0.2504	0.0004	0.0614
200(55,60,45,40)	High	p_{01}=0.4000	0.3711	-0.0289	0.0898
		p_{02}=0.3110	0.2877	-0.0233	0.0637
		p_{03}=0.2257	0.2089	-0.0169	0.0518
		p_{04}=0.1500	0.1425	-0.0075	0.0551
200(55,60,45,40)	Moderate	p_{01}=0.4000	0.3937	-0.0063	0.0781
		p_{02}=0.3110	0.3066	-0.0044	0.0547
		p_{03}=0.2257	0.2231	-0.0027	0.0459
		p_{03}=0.1500	0.1512	0.0012	0.0522
200(55,60,45,40)	Low	p_{01}=0.4000	0.4001	0.0000	0.0593
		p_{02}=0.3110	0.3101	-0.0009	0.0405
		p_{03}=0.2257	0.2238	-0.0020	0.0356
		p_{04}=0.1500	0.1492	-0.0009	0.0426

Table 5.23: Estimates, bias and root mean square error(RMSE) using EM algorithm for cure fraction under different simulation settings with $\alpha = 2, \lambda = 3, n = 400$

n	p	p_0	Estimate	Bias	RMSE
400(110,120,90,80)	High	p_{01}=0.6500	0.6483	-0.0017	0.0505
		p_{02}=0.5294	0.5289	-0.0005	0.0398
		p_{03}=0.3910	0.3905	-0.0005	0.0334
		p_{04}=0.2500	0.2499	-0.0001	0.0428
400(110,120,90,80)	High	p_{01}=0.4000	0.3688	-0.0312	0.0628
		p_{02}=0.3110	0.2858	-0.0252	0.0450
		p_{03}=0.2257	0.2077	-0.0181	0.0381
		p_{04}=0.1500	0.1406	-0.0094	0.0409
400(110,120,90,80)	Moderate	p_{01}=0.4000	0.3567	-0.0433	0.0760
		p_{02}=0.3110	0.2726	-0.0385	0.0637
		p_{03}=0.2257	0.1944	-0.0313	0.0550
		p_{04}=0.1500	0.1284	-0.0216	0.0487
400(110,120,90,80)	Low	p_{01}=0.4000	0.3989	-0.0011	0.0422
		p_{02}=0.3110	0.3100	-0.0010	0.0292
		p_{03}=0.2257	0.2248	-0.0009	0.0274
		p_{04}=0.1500	0.1503	0.0003	0.0333

Table 5.24: Estimates, bias and root mean square error(RMSE) using EM algorithm for cure fraction under different simulation settings with $\alpha = 4, \lambda = 1, n = 400$

n	p	p_0	Estimate	Bias	RMSE
400(110,120,90,80)	High	p_{01}=0.6500	0.6500	0.0000	0.0481
		p_{02}=0.5294	0.5297	0.0003	0.0367
		p_{03}=0.3910	0.3902	-0.0009	0.0320
		p_{04}=0.2500	0.2486	-0.0014	0.0452
400(110,120,90,80)	High	p_{01}=0.4000	0.3698	-0.0302	0.0605
		p_{02}=0.3110	0.2893	-0.0218	0.0409
		p_{03}=0.2257	0.2131	-0.0126	0.0344
		p_{04}=0.1500	0.1472	-0.0028	0.0401
400(110,120,90,80)	Moderate	p_{01}=0.4000	0.3905	-0.0095	0.0628
		p_{02}=0.3110	0.3033	-0.0078	0.0444
		p_{03}=0.2257	0.2197	-0.0060	0.0345
		p_{04}=0.1500	0.1469	-0.0031	0.0359
400(110,120,90,80)	Low	p_{01}=0.4000	0.4022	0.0022	0.0414
		p_{02}=0.3110	0.3121	0.0010	0.0276
		p_{03}=0.2257	0.2256	-0.0001	0.0256
		p_{04}=0.1500	0.1500	0.0000	0.0323

5.2 Illustrative examples

In this section, we consider two real life datasets to illustrate our inference procedures. For each dataset, we use direct optimization and the EM algorithm to estimate the MLEs under our proposed model. Furthermore, we discriminate our model from exponential and for this purpose, we used AIC, BIC and LRT.

5.2.1 Melanoma data I

In this section, we consider a cutaneous melanoma data. This data, from Kirkwood et al. (2000) and also analyzed by Yin and Ibrahim (2005), were originally a part of a test on cutaneous melanoma, a malignant cancer type, for the evaluation of postoperative treatment performance with a high dose of a certain drug(interferon alpha-2b) for prevention of recurrence of the disease. Patients were included from 1991 to 1995 and followed up until 1998. We used survival time as our response variable. There are four nodule categories which are considered as the covariates and the data has been right censored. Relevant information of these data can be found in Appendix. The average survival time is 3.17 and the standard deviation is 1.69.

The results of the two maximum likelihood procedures are shown in Table 5.25. It is clear that the results from the two estimation methods give similar results, more noticeably in the case of GED. Cure fraction estimates and their standard errors are obsereved in Table 5.26 for the GED and in Table 5.27 for exponential distribution. Delta method is used to obtain these standard errors. The model discrimination results are shown in Table 5.28. From this table we can see that, according to both AIC and BIC, the model with GED lifetimes give smaller values and hence better fits the data. Moreover, as mentioned, we performed the likelihood ratio test. Our hypotheses of consideration are: $H_0 : \alpha = 1$ vs. $H_a : \alpha \neq 1$. For this dataset, $\Lambda = 45.5808$ for direct optimization and $\Lambda = 46.50491$ for the EM algorithm. Now, since the difference in the number of parameters between the two models is 1, we compare

our computed test statistic to a $\chi^2(1)$ percentile. We chose 5% level of significance and so the critical value is 3.84. We concluded that we reject the null hypothesis; in other words, GED gives a better fit to the dataset. We have also constructed QQ plots of normalized residuals as developed by Dunn and Smyth (1996). These are shown in Figure 5.1. It is clear from all four plots that both GED and exponential give good fits to the dataset.

Figure 5.1: QQ plot for the melanoma data I

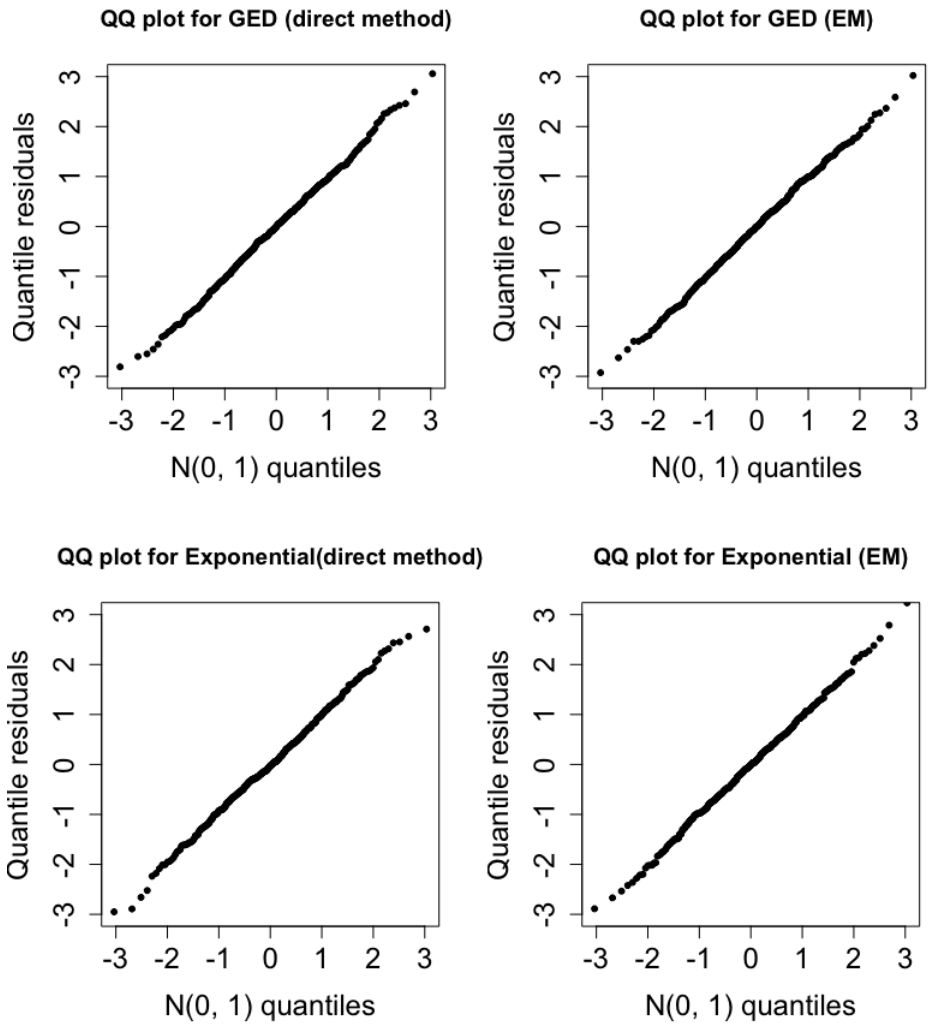


Table 5.25: Parameter estimates and SE of melanoma data I using the GED and exponential distribution

	GED estimates (SE)				Exponential estimates (SE)		
Estimation method	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\lambda}$	$\hat{\beta}_0$	$\hat{\beta}_1$
Direct optimization	2.5087 (0.3153)	0.6845 (0.0921)	-1.2326 (0.2002)	0.3648 (0.0686)	0.1182 (0.0536)	-0.4399 (0.4000)	0.3453 (0.0685)
EM	2.5088 (0.3154)	0.6845 (0.0921)	-1.2326 (0.2002)	0.3648 (0.0686)	0.1189 (0.0534)	-0.4459 (0.3967)	0.3457 (0.0685)

Table 5.26: Cure fraction estimates, SE and 95% confidence interval of melanoma data I using the GED

	GED estimates (SE)			
Estimation method	p_{01}	p_{02}	p_{03}	p_{04}
Direct optimization	0.6571 (0.0392)	0.5463 (0.0327)	0.4186 (0.0342)	0.2853 (0.0469)
95% CI with direct optimization	(0.5802, 0.7340)	(0.4822, 0.6103)	(0.3516, 0.4856)	(0.1933, 0.3773)
EM	0.6572 (0.0392)	0.5463 (0.0327)	0.4186 (0.0342)	0.2853 (0.0470)
95% CI with EM	(0.5803, 0.7340)	(0.4822, 0.6104)	(0.3516, 0.4856)	(0.1933, 0.3774)

Table 5.27: Cure fraction estimates, SE and 95% confidence interval of melanoma data I using exponential distribution

	Exponential estimates (SE)			
Estimation method	p_{01}	p_{02}	p_{03}	p_{04}
Direct optimization	0.4026 (0.5017)	0.2767 (0.4873)	0.1629 (0.4058)	0.0771 (0.2717)
95% CI with direct optimization	(-0.5806, 1.3859)	(-0.6785, 1.2319)	(-0.6325, 0.9583)	(-0.4555, 0.6096)
EM	0.4047 (0.1368)	0.2785 (0.1291)	0.1643 (0.1081)	0.0779 (0.0753)
95% CI with EM	(0.1365, 0.6729)	(0.0255, 0.5316)	(-0.0475, 0.3761)	(-0.0696, 0.2254)

Table 5.28: AIC and BIC of melanoma data I using the GED and exponential distribution

	Direct optimization		EM algorithm	
Distribution of lifetime	*AIC*	*BIC*	*AIC*	*BIC*
GED	1028.192	1044.315	1028.195	1044.318
Exponential	1071.773	1083.865	1072.7	1084.792

5.2.2 Melanoma data II

Our second illustrative example involves the dataset used by Rodrigues et al. (2009). This melanoma dataset, which can be found in the R package `timereg`, contains data relating to survival of patients after operation for malignant melanoma and relevant part of it is included in the Appendix. It contains several covariates including gender, tumour thickness and ulceration status, among others. For our analysis, we considered the single covariate tumour thickness, which we note to be continuous. Also, based on the dataset's status indicator variable, we constructed our censoring indicator. As in previous example, we fit our model to the survival times. The average survival time is 5.898 and the standard deviation is 3.07.

Table 5.29: Parameter estimates and SE of melanoma data II using the GED and exponential distribution

	GED estimates (SE)				Exponential estimates (SE)		
Estimation method	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\lambda}$	$\hat{\beta}_0$	$\hat{\beta}_1$
Direct optimization	2.1620 (0.4753)	0.2790 (0.0845)	-1.2311 (0.2383)	0.1608 (0.0312)	0.0321 (0.0494)	-0.0081 (1.3690)	0.1545 (0.0315)
EM algorithm	2.1641 (0.4758)	0.2794 (0.0845)	-1.2320 (0.2382)	0.1608 (0.0312)	0.0366 (0.0468)	-0.1262 (1.1217)	0.1545 (0.0314)

Table 5.30: Cure fraction estimates, SE and 95% confidence interval of melanoma data II using the GED

	GED estimates (SE)			
Estimation method	p_{01}	p_{02}	p_{03}	p_{04}
Direct optimization	0.7097 (0.0544)	0.6685 (0.0572)	0.6232 (0.0606)	0.5738 (0.0647)
95% CI with direct optimization	(0.6031, 0.8164)	(0.5564, 0.7806)	(0.5045, 0.7418)	(0.4469, 0.7007)
EM	0.7099 (0.0543)	0.6687 (0.0571)	0.6234 (0.0605)	0.5741 (0.0646)
95% CI with EM	(0.6034, 0.8164)	(0.5568, 0.7807)	(0.5049, 0.7419)	(0.4474, 0.7007)

Table 5.31: Cure fraction estimates, SE and 95% confidence interval of melanoma data II using exponential distribution

	Exponential estimates (SE)			
Estimation method	p_{01}	p_{02}	p_{03}	p_{04}
Direct optimization	0.3142 (0.4982)	0.2590 (0.4796)	0.2067 (0.4473)	0.1588 (0.4020)
95% CI with direct optimization	(-0.6622, 1.2907)	(-0.6811, 1.1991)	(-0.6701, 1.0834)	(-0.6290, 0.9466)
EM	0.3575 (0.4124)	0.3010 (0.4055)	0.2463 (0.3878)	0.1949 (0.3589)
95% CI with EM	(-0.4508, 1.1658)	(-0.4938, 1.0959)	(-0.5138, 1.006)	(-0.5086, 0.8984)

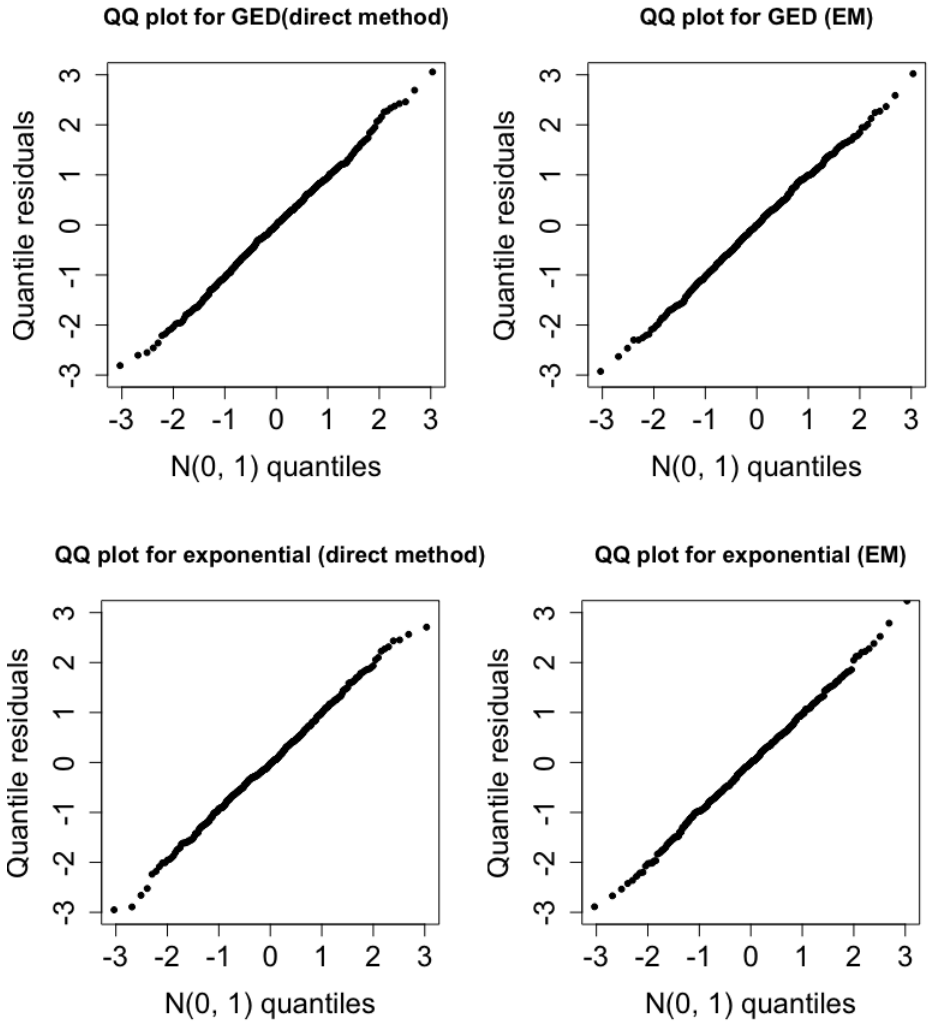
Table 5.32: AIC and BIC of melanoma data II using the GED and exponential distribution

	Direct optimization		EM algorithm	
Distribution of lifetime	*AIC*	*BIC*	*AIC*	*BIC*
GED	440.9625	454.2545	440.9625	454.2546
Exponential	449.5995	459.5685	449.6127	459.5817

The results of the two maximum likelihood procedures are shown in Table 5.29. It is evident that the results from two estimation methods gives similar results, once again more noticeably in GED. Moreover, we also observe the cure fraction separately for the GED in Table 5.30 and also for exponential in Table 5.31. To find the standard errors of cure fractions delta method is used. Also for the model discrimination, results are shown in Table 5.32. From this table, we can see that according to both AIC and BIC, the GED gave smaller values of the criteria and hence better fits the data.

We again performed the likelihood ratio test. Under the same hypotheses as the previous example, for this dataset, $\Lambda = 10.637$ for direct optimization and $\Lambda = 10.650$ for the EM algorithm. Again, at 5% level of significance, the critical value is 3.84. As before, we reject the null hypothesis; in other words, the GED gives a better fit to the dataset. We have also constructed QQ plots for the second dataset; these are shown in Figure 5.2. It is once again clear from all four plots both GED and exponential provides good fits to the dataset.

Figure 5.2: QQ plot for the melanoma data II



Chapter 6

Conclusion

In this thesis, a cure rate model based on generalized exponential lifetimes distribution in a competing risks scenario has been proposed. Our model also allows for a cure fraction in the population. Furthermore, to be more realistic, right censoring was also incorporated. The likelihood function for the proposed model was then constructed, and in relation to that, two methods of maximum likelihood estimation were developed. First we consider a direct optimization technique of the likelihood function through numerical methods. Secondly, an EM algorithm was developed. Here, the E-step was to find the conditional expectation of the complete log-likelihood of our proposed model and the M-step involved maximization of the conditional log-likelihood incorporating the expectation of cure fraction found in the E-step.

We have demonstrated these two estimation methods through a simulation study and two illustrative examples. We were then able to observe their performance and make comments. In the simulation study, we found each method gave reasonable estimates of the parameter values in different settings. It was evident that model fitting works better for large sample sizes compared to small sample sizes, as expected. Two cuteneous melanoma datasets were chosen illustrate the model fitting. For each dataset, we fitted both the GED and exponential distributions. After the analyses, it was clear that both estimation methods gave very similar results overall. Three model discrimination methods were then carried out and the cure rate model with GED lifetimes fitted both datasets better than one assuming exponential

lifetimes.

Overall, the proposed estimation methods seem to work well. However, in our numerical studies, we did not allow the situation when there is no cure proportion. Here, the results are limited to the case of $p_0 > 0$, and further work is needed to handle the special case of $p_0 = 0$. Finally, we only considered one form of censoring, that is of right censoring. It would be interesting to extend our inference procedures to other forms of censoring, such as progressive censoring. Also, the work in this thesis could be extended to look at other distributions for the number of competing risks, such as the COM-Poisson.

APPENDIX

A

Melanoma data I

Table 1: Melanoma data I

Surv.Time	Censoring	Nodule Category
0.7228	1	3
6.2998	0	1
6.6229	0	3
3.4716	1	4
7.0116	0	1
6.6475	0	1
6.9760	0	2
0.6708	1	4
6.8802	0	1
6.6092	0	1
1.0678	1	4
6.3299	0	2
6.3217	0	1
1.5031	1	3
0.7721	1	3
1.7823	1	4
6.6585	0	2
3.5373	1	1
5.7495	0	2
3.8138	1	1
5.9767	0	2
1.6892	1	1
5.7385	1	3
5.8754	0	1
1.5743	1	1
5.5688	0	3
1.5715	1	3
2.4367	1	1
5.7878	0	3
1.5989	1	2

Continued on next page

Table 1 – *Melanoma data I*

Surv.Time	**Censoring**	**Nodule Category**
0.6078	1	2
5.6756	0	1
0.2409	1	4
5.5469	0	3
6.0452	0	4
1.9001	1	2
2.1328	1	4
0.3559	1	2
2.4339	1	4
1.4538	1	3
5.2485	0	2
1.5305	1	4
5.3279	0	3
0.9336	1	4
2.2231	1	3
4.9446	0	1
5.0760	0	4
5.1061	0	1
5.0513	1	4
5.7303	0	2
1.2457	1	4
1.4127	1	3
0.9418	1	2
5.4648	0	2
1.6290	1	3
5.4155	0	2
5.6400	0	2
3.1485	1	1
5.6345	0	3
1.6920	1	4
5.1335	0	3
3.2416	1	3
5.2430	0	2
5.3114	0	4
2.2752	1	1
2.5681	1	2
5.0431	0	3
4.9500	0	1
4.9802	0	3
1.5441	1	3
4.7200	0	4
5.1581	0	1
5.0678	0	1
1.7139	1	3
5.1526	0	3
1.9220	1	4
4.6872	0	3
4.9582	0	1

Continued on next page

Table 1 – *Melanoma data I*

Surv.Time	**Censoring**	**Nodule Category**
4.1478	0	2
1.4456	1	2
4.7392	0	2
5.1088	0	1
4.3696	0	2
3.3347	1	4
3.3046	1	4
3.6112	0	2
0.5859	1	3
1.4866	1	4
4.6379	0	1
4.3203	0	2
4.4627	0	3
4.9035	0	1
2.1958	1	1
1.1855	1	4
4.4271	0	2
1.5359	1	2
4.3778	0	2
4.3368	0	1
4.5010	0	3
4.4654	0	1
2.7981	1	1
4.7146	0	3
3.0308	1	3
3.3374	1	2
2.7789	1	1
4.1396	0	1
4.4873	0	1
0.7666	1	1
4.1944	0	2
4.5777	0	1
1.5332	1	2
1.4812	1	4
4.4955	0	3
4.4627	0	2
4.4901	0	1
4.0739	0	3
4.0767	0	2
3.3949	0	3
3.9617	0	1
3.9398	0	3
4.1862	0	1
2.0096	1	4
4.2190	0	2
4.0219	0	2
1.9685	1	3
2.0835	1	4

Continued on next page

Table 1 – *Melanoma data I*

Surv.Time	**Censoring**	**Nodule Category**
3.5072	0	1
3.7673	0	1
1.1526	1	3
0.8131	1	2
4.0685	0	2
2.9897	0	2
1.3416	1	2
3.8494	0	2
2.2204	1	1
3.7454	0	2
3.6030	0	4
0.6845	1	4
1.0623	1	2
3.5592	0	1
3.3867	0	2
3.4661	0	1
3.5373	0	2
3.6824	0	2
3.5756	0	4
0.9610	1	2
1.4237	0	2
2.1547	0	1
3.6167	0	2
3.2361	0	1
0.9254	1	1
0.9555	1	4
3.2115	0	2
3.2033	0	2
3.4442	0	2
2.6420	0	3
1.9411	1	3
0.9802	1	2
0.9555	1	4
3.1020	0	2
3.2005	0	1
2.9651	0	2
3.1239	0	2
2.3792	0	2
2.9268	0	4
3.0664	0	3
3.0472	0	3
2.5845	0	1
2.0589	1	4
0.7912	0	1
2.8967	0	1
3.1266	0	3
0.3477	1	4
5.9986	0	4

Continued on next page

Table 1 – *Melanoma data I*

Surv.Time	**Censoring**	**Nodule Category**
5.9685	1	4
6.2943	0	1
6.5818	0	3
6.6749	0	2
0.6133	1	2
0.8761	1	2
5.8453	0	2
6.1054	0	3
1.7029	1	3
6.0151	0	1
1.6646	1	2
2.1109	1	4
1.5962	1	2
0.7721	1	3
1.4073	1	2
1.6810	1	2
1.6044	1	4
1.8179	1	3
3.4114	1	2
5.1937	0	4
5.9548	0	4
5.7112	0	2
3.3922	0	1
1.8864	1	1
3.9617	0	1
1.0075	1	4
2.0862	1	1
1.6290	1	2
6.0397	0	3
1.2046	1	3
3.3073	1	1
5.5031	0	4
1.7358	1	1
3.7235	1	1
2.0753	1	3
0.8214	1	3
5.8070	0	1
2.8720	1	2
1.5797	1	2
4.9829	0	2
5.7714	0	4
0.9637	1	1
1.3251	0	2
5.2266	0	1
1.5140	1	1
0.6762	1	4
1.0157	1	3
5.0431	0	4

Continued on next page

Table 1 – *Melanoma data I*

Surv.Time	**Censoring**	**Nodule Category**
5.2238	0	2
4.8843	0	2
0.4600	1	3
5.0130	0	1
5.0924	0	1
5.4045	0	2
2.8638	1	2
4.9829	0	2
5.0897	0	4
0.8022	1	4
3.6578	1	2
4.8049	0	2
2.4038	1	1
4.1725	0	2
1.7385	1	2
0.6023	1	3
5.1636	0	2
1.2950	1	2
4.8077	0	4
5.1499	0	4
4.5229	0	1
4.7557	0	2
4.2710	0	1
0.5257	1	3
5.0157	0	1
2.5161	1	1
4.8022	0	1
0.8597	1	2
4.7365	0	2
4.9227	0	3
4.4846	0	3
0.4600	1	2
4.8515	0	1
3.3073	1	1
4.4928	0	1
0.4791	1	3
3.8823	0	1
4.6817	0	2
4.5996	0	3
3.5044	0	4
4.3450	0	3
1.6701	1	3
2.4011	1	4
4.3450	0	3
4.6133	1	2
3.4415	0	1
3.7372	0	2
3.2279	0	4

Continued on next page

Table 1 – *Melanoma data I*

Surv.Time	**Censoring**	**Nodule Category**
4.1287	0	4
0.5476	1	4
1.2129	1	2
4.1478	0	3
1.6071	1	4
4.5421	0	4
2.0205	1	1
0.1698	1	4
4.2656	0	2
4.2081	0	1
1.3388	1	4
3.5181	0	3
0.5968	1	4
3.2964	0	1
4.1205	0	1
1.7632	1	1
3.0390	1	3
2.2587	0	3
0.5558	1	2
1.2430	1	2
2.8172	0	1
4.0575	0	2
3.7454	0	1
2.8665	0	3
3.6140	0	1
4.0328	0	2
2.8008	0	3
1.4730	1	2
0.7420	1	4
4.0630	0	2
3.9891	0	2
3.5702	1	4
3.5838	0	1
3.5975	0	4
1.8590	1	2
0.8597	1	4
1.6372	1	1
1.6975	1	1
1.5086	1	3
3.8686	0	3
3.6578	0	1
1.9959	1	3
0.3313	1	3
3.6386	0	3
3.6030	0	2
2.5462	1	1
2.1738	1	1
3.5209	0	2

Continued on next page

Table 1 – *Melanoma data I*

Surv.Time	**Censoring**	**Nodule Category**
2.8227	1	2
2.6366	1	2
2.6256	0	1
2.8036	0	2
3.5455	0	2
0.9938	1	4
2.7187	0	1
0.9199	1	4
1.8508	1	2
0.9117	1	4
5.1253	1	3
5.7988	0	2
3.2005	0	2
4.4326	0	2
1.1581	1	1
0.2656	1	3
5.4976	0	4
1.2430	1	4
5.5031	0	2
5.6509	0	2
3.4935	1	2
0.1478	1	3
5.9302	0	2
0.7830	1	2
1.9247	1	2
3.1102	1	2
4.9336	0	3
1.7029	1	4
1.7358	1	4
1.4374	1	4
2.2478	1	2
5.4100	0	4
4.2628	1	1
4.7392	0	3
4.4764	0	1
0.5366	1	3
0.9774	1	1
5.0760	0	4
5.0020	0	2
3.0582	1	2
4.4134	0	1
1.9192	1	3
0.8898	1	4
1.3169	1	4
4.0739	0	2
2.6804	1	1
3.8905	0	2
4.0465	0	3

Continued on next page

Table 1 – *Melanoma data I*

Surv.Time	**Censoring**	**Nodule Category**
4.4928	0	3
1.5441	1	3
2.0917	1	4
3.8029	1	2
3.1567	0	2
1.1389	0	3
2.3080	1	2
4.6489	0	2
0.2847	1	2
1.2676	1	4
3.7043	0	1
3.2580	1	3
4.1369	1	3
0.4244	1	4
2.9377	0	2
1.8371	1	2
4.0027	0	3
3.0856	0	1
3.6523	0	3
3.7892	0	1
1.5113	0	2
3.3758	0	1
2.9952	0	3
1.3196	1	1
3.4114	0	2
0.2930	1	2
2.9322	0	1
2.8912	0	2
3.0198	0	1
3.1732	0	2
3.2526	0	1
2.7378	0	2
0.8843	1	4
3.3621	0	1
0.6899	1	4
2.9870	0	1
3.2690	0	4
3.2224	0	4
1.7823	1	2
1.8070	1	2
2.8830	0	3
2.7159	0	2
3.1923	0	4
2.9432	0	1
0.7337	1	4
2.7707	0	1
1.3689	1	1
2.6776	0	2

Continued on next page

Table 1 – *Melanoma data I*

Surv.Time	Censoring	Nodule Category
2.7762	0	2
2.9788	1	1

Melanoma data II

Table 2: Melanoma data II

Surv.Time	Censoring	Tumour Thickness
0.0274	0	6.76
0.0822	0	0.65
0.0959	0	1.34
0.2712	0	2.90
0.5068	1	12.08
0.5589	1	4.84
0.5753	1	5.16
0.6356	1	12.88
0.6356	0	3.22
0.7644	1	7.41
0.8082	1	4.19
0.9726	0	0.16
1.0575	1	3.87
1.1671	1	4.84
1.2849	1	2.42
1.3507	0	12.56
1.4493	1	5.80
1.7014	1	7.06
1.7233	1	5.48
1.8055	1	7.73
1.8274	1	13.85
1.9671	1	2.34
2.0603	1	4.19
2.1342	1	4.04
2.1726	1	4.84
2.2384	1	0.32
2.2630	0	8.54
2.2822	1	2.58
2.3507	1	3.56
2.3808	1	3.54
2.3890	1	0.97
2.6493	1	4.83
2.6767	1	1.62
2.6904	1	6.44
2.8521	1	14.66
2.8904	1	2.58
2.9096	1	3.87

Continued on next page

Table 2 – *Melanoma data II*

Surv.Time	**Censoring**	**Tumour Thickness**
2.9452	1	3.54
3.1671	1	1.34
3.3644	1	2.24
3.4301	1	3.87
3.4822	1	3.54
3.5945	1	17.42
3.9096	0	1.29
3.9315	1	3.22
4.1068	0	1.29
4.1260	1	4.51
4.1315	0	8.38
4.1370	0	1.94
4.1425	0	0.16
4.1534	1	2.58
4.1781	0	1.29
4.2247	0	0.16
4.2411	1	1.62
4.2658	0	1.29
4.2740	1	2.10
4.2822	0	0.32
4.3397	1	0.81
4.3973	0	1.13
4.4411	1	5.16
4.4575	0	1.62
4.4767	0	1.37
4.4959	0	0.24
4.4959	0	0.81
4.5151	0	1.29
4.5260	0	1.29
4.5315	0	0.97
4.5315	0	1.13
4.5671	1	5.80
4.5973	0	1.29
4.6164	0	0.48
4.6301	1	1.62
4.6849	0	2.26
4.6849	0	0.58
4.7288	1	0.97
4.7808	0	2.58
4.8274	0	0.81
4.8740	0	3.54
4.8959	0	0.97
4.8959	0	1.78
4.9123	0	1.94
4.9425	0	1.29
4.9644	0	3.22
5.0301	0	1.53
5.0384	0	1.29

Continued on next page

Table 2 – *Melanoma data II*

Surv.Time	**Censoring**	**Tumour Thickness**
5.0384	0	1.62
5.0795	0	1.62
5.0849	0	0.32
5.0959	0	4.84
5.1068	0	1.29
5.2027	0	0.97
5.2438	0	3.06
5.2575	0	3.54
5.2603	0	1.62
5.2795	0	2.58
5.2959	1	1.94
5.3205	0	0.81
5.3562	0	7.73
5.3589	0	0.97
5.3644	0	12.88
5.3781	0	2.58
5.3973	0	4.09
5.4932	0	0.64
5.4986	0	0.97
5.5096	0	3.22
5.5452	0	1.62
5.5562	0	3.87
5.5836	0	0.32
5.6329	0	0.32
5.6411	0	3.22
5.6466	1	2.26
5.6493	1	3.06
5.6849	0	2.58
5.7123	0	0.65
5.7589	0	1.13
5.7616	1	0.81
5.7644	0	0.97
5.7753	1	1.76
5.7863	0	1.94
5.8904	0	0.65
5.9068	0	0.97
5.9315	0	5.64
6.0521	0	9.66
6.1014	0	0.10
6.1014	0	5.48
6.1808	1	2.26
6.2027	0	4.83
6.4082	0	0.97
6.4685	0	0.97
6.5397	0	5.16
6.5425	1	0.81
6.5836	0	2.90
6.6466	0	3.87

Continued on next page

Table 2 – *Melanoma data II*

Surv.Time	**Censoring**	**Tumour Thickness**
6.6466	0	1.94
6.6603	0	0.16
6.7397	0	0.64
6.7589	1	2.26
6.8274	0	1.45
6.8301	0	4.82
6.9068	0	1.29
6.9644	0	7.89
7.0110	0	0.81
7.0274	1	3.54
7.0411	0	1.29
7.2877	0	0.64
7.3041	0	3.22
7.3315	0	1.45
7.5014	0	0.48
7.6219	1	1.94
7.6356	0	0.16
8.1753	0	0.16
8.3068	0	1.29
8.3288	0	1.94
8.3342	1	3.54
8.4027	0	0.81
8.4356	0	0.65
8.4959	0	7.09
8.6137	0	0.16
8.6356	0	1.62
8.6411	0	1.62
8.7123	0	1.29
8.7178	0	6.12
8.7260	0	0.48
8.7644	0	0.64
8.8438	0	3.22
8.8466	0	1.94
8.9808	0	2.58
9.0329	0	2.58
9.1178	0	0.81
9.1233	0	0.81
9.1452	1	3.22
9.2685	0	0.32
9.2712	0	3.22
9.2740	0	2.74
9.2822	0	4.84
9.3205	0	1.62
9.4274	0	0.65
9.4740	0	1.45
9.4767	0	0.65
9.4767	0	1.29
9.5233	0	1.62

Continued on next page

Table 2 – *Melanoma data II*

Surv.Time	**Censoring**	**Tumour Thickness**
9.6521	0	3.54
10.0466	0	3.22
10.1233	0	0.65
10.1233	0	1.03
10.3452	0	7.09
10.3452	0	1.29
10.4932	0	0.65
10.5644	0	1.78
10.6082	0	12.24
10.7096	0	8.06
10.8712	0	0.81
10.9616	0	2.10
11.2411	0	3.87
11.2849	0	0.65
11.2986	0	1.94
11.5260	0	0.65
11.8082	0	2.10
12.0274	0	1.94
12.2712	0	1.13
12.3068	0	7.06
12.7890	0	6.12
12.8438	0	0.48
13.4959	0	2.26
15.2466	0	2.90

B

R codes for simulation study

```
#### we need to find theta in order to obtain a random censoring time from exponential distribution

####The function to calculate theta parameter of censoring time t is:


parameter<-function(theta,p,pnot,N,alpha,lambda)

{

a<-c() # a is a vector

x<-rexp(N) #N observations from exp(1) distribution

for (i in 1:N) #for loop to obtain 'a' vector

{

a[i]<- pnot^(1-exp(-lambda*x[i]/theta))^alpha- p

}

return(sum(a))

}
```

```
#function to obtain the observed time

data.fun<-function(p,pnot,N,alpha,lambda,x,s_size)
{
ind<-list()
cov<-list()
lifetime<-list()

for (j in 1:4)
{
theta<-uniroot(parameter,c(exp(-10),exp(10)),p=p[j],
pnot=pnot[j],N=N,alpha=alpha,lambda=lambda)$root
observed<-NULL ##vector of observed lifetime data
unobserved<-NULL  ##vector of unobserved lifetime data
for( i in 1:s_size[j])
{
m<-rpois(1,-log(pnot[j]))#one observation from poisson
if (m!=0)
{
ged<-rgen.exp(m, alpha, lambda) #m observation from GED
y<-min(ged) #minimum of GED observations
cens<-rexp(1, rate=theta) #an observation from exp(theta)
t <-min(y,cens) #minimum of y and c which is the observed time
}
else {
cens<-rexp(1, rate=theta) #an observation from exp(theta)
t <-cens
}
```

```
if (t!=cens){
observed <- c(observed, t)
}
else {
unobserved<-c(unobserved,t)
}

}

lifetime[[j]]<- c(observed,unobserved)
ind[[j]]<-c(rep(1,times=length(observed)),rep(0,times=length(unobserved)))
cov[[j]]<-rep(j,s_size[j])
}

result<-list(lifetime=lifetime,ob=observed,un=unobserved,
indicator=ind,covariate=cov)
output<-matrix(c(unlist(result$lifetime),unlist(result$indicator),
unlist(result$covariate)),ncol=3)
output
}
```

```
################################### the log likelihood function


neg.loglikelihood<-function(param,gdata)
{
alpha<-param[1]
lambda<-param[2]
beta0<-param[3]
beta1<-param[4]

m11<-c()
m22<-c()

beta<-matrix(c(beta0,beta1),ncol=1)

for (j in 1:4)
{
m1<-c()
m2<-c()
y1 <- gdata[which(gdata[,2]==1&gdata[,3]==j),1]
x <- c(1,j)
y2 <- gdata[which(gdata[,2]==0&gdata[,3]==j),1]

if (length(y1)==0){m1=0}

else {for (i in 1:length(y1))
{m1[i]<-log(alpha)+log(lambda)+(x%*%beta)+(-lambda*y1[i])+
(alpha-1)*log(1-exp(-lambda*y1[i]))-(exp(x%*%beta))*(1-exp(-lambda*y1[i]))^alpha
}}

if (length(y2)==0){m2=0}
```

```
else {for (k in 1:length(y2))
{m2[k]<- -(exp(x%*%beta))*(1-exp(-lambda*y2[k]))^alpha
}
}

m11[j]<-sum(m1)
m22[j]<-sum(m2)

}

a<- -(sum(m11,m22))
return(a)
}
```

```
##################function to get betas

betas<-function(p01,p04)

{

b1<-((log(-log(p04)))-log(-log(p01)))/3

b0<- log(-log(p01))-b1

b<-c(b0,b1)

p02<-exp(-exp(b[[1]]+b[[2]]*2))

p03<-exp(-exp(b[[1]]+b[[2]]*3))

p02

p03

a<-list(b0,b1,p01,p02,p03,p04)

return(a)

}


#####################grid search

grid<-function(gdata)

{

alpha<-seq(0.5,4.5,by=0.5)

lambda<-seq(0.5,4.5,by=0.5)

beta0<-seq(-5,5,by=0.5)

beta1<-seq(-5,5,by=0.5)

min.like<-Inf

for (i in 1:9){

for (j in 1:9){

for (k in 1:21){

for (l in 1:21){
```

```
like<-neg.loglikelihood(c(alpha[i],lambda[j],beta0[k],beta1[l]),gdata)
if (like < min.like) {min.like<-like
alpha.min<-alpha[i]
lambda.min<-lambda[j]
beta0.min<-beta0[k]
beta1.min<-beta1[l]
}
}
}
}
}
minima<-c(alpha.min,lambda.min,beta0.min,beta1.min)
return(minima)
}


log.likelihood <- function(mles,data)
{-neg.loglikelihood(param = mles,gdata=data)}
```

```
############EM section:
pi.em<-function(param,gdata)
{
p0<-c()
N<-c()
D<-c()
alpha<-param[1]
lambda<-param[2]
beta0<-param[3]
beta1<-param[4]

beta<-matrix(c(beta0,beta1),ncol=1)

y<-gdata[,1]
x<-matrix(c(rep(1,times=length(y)),gdata[,3]),ncol=2)
s_t<-1-(1-exp(-lambda*y))^alpha
eta<-exp(x%*%beta)
pi.e<-c()
for (i in (1 :length(y)))
{
p0[i]<-exp(-exp(x[i,]%*%beta))
if(s_t[i]==0){s_t[i]<-.Machine$double.eps}
if(p0[i]==0){p0[i]<-.Machine$double.eps}

N[i]<-(p0[i]^(-s_t[i]))*p0[i]-p0[i]
D[i]<-(p0[i]^(-s_t[i]))*p0[i]
if(round(N[i],15)==0){N[i] <- .Machine$double.eps}
if(round(D[i],15)==0){D[i] <- .Machine$double.eps}
pi.e[i]<-(N[i]/D[i])
```

```
}

return(pi.e)

}
```

```
#############conditional log likelihood

Q2<-function(param,gdata,pi.e) ##need to use only censored lifetime

{

alpha<-param[1]

lambda<-param[2]

beta0<-param[3]

beta1<-param[4]

m1<-c()

m2<-c()

beta<-matrix(c(beta0,beta1),ncol=1)

A<-c()

B<-c()

y1<-gdata[which(gdata[,2]==1),1]

x1<-matrix(c(rep(1,times=length(gdata[which(gdata[,2]==1),3])),

gdata[which(gdata[,2]==1),3]),ncol=2)

x2<-matrix(c(rep(1,times=length(gdata[which(gdata[,2]==0),3])),

gdata[which(gdata[,2]==0),3]),ncol=2)

y2<-gdata[which(gdata[,2]==0),1]

s_t1<-(1-(1-exp(-lambda*y1))^alpha)

s_t2<-(1-(1-exp(-lambda*y2))^alpha)

eta1<-exp(x1%*%beta)

eta2<-exp(x2%*%beta)
```

```
if (length(y1)==0){m1=0}

if(length(y1)!=0 ) {for (i in 1:length(y1))

{if(s_t1[i]==0){s_t1[i]<-.Machine$double.eps}

m1[i]<-log(alpha*lambda)+(x1[i,]%*%beta)-lambda*y1[i]

+(alpha-1)*log(1-exp(-lambda*y1[i]))-exp(x1[i,]%*%beta)+eta1[i]*s_t1[i]}}


if (length(y2)==0){m2=0}


if(length(y2)!=0) {for (k in 1:length(y2))

{if(s_t2[k]==0){s_t2[k]<-.Machine$double.eps}

A[k]<-exp(x2[k,]%*%beta)*s_t2[k]

if(A[k]<40){

B[k]<-exp(A[k])*exp(-exp(x2[k,]%*%beta))-exp(-exp(x2[k,]%*%beta))

if (B[k]==0){

B[k]<-.Machine$double.eps

m2[k]<- pi.e[k]*log(B[k])-(1-pi.e[k])*exp(x2[k,]%*%beta)}

else{m2[k]<- pi.e[k]*log(exp(A[k])-1)-exp(x2[k,]%*%beta)}}

if(A[k]>=40){

m2[k]<- pi.e[k]*(A[k])-exp(x2[k,]%*%beta)}

}

}

a<- -sum(sum(m1),sum(m2))

#####a<- list((m1),(m2),x2,s_t2,eta2,pi.e,A)

return(a)

}
```

```
####################grid search
grid.em<-function(gdata)
{
alpha<-seq(0.5,4.5,by=0.5)
lambda<-seq(0.5,4.5,by=0.5)
beta0<-seq(-5,5,by=0.5)
beta1<-seq(-5,5,by=0.5)
min.Qfun<-Inf
for (i in 1:9){
for (j in 1:9){
for (k in 1:21){
for (l in 1:21){

pi.grid<-pi.em(c(alpha[i],lambda[j],beta0[k],
beta1[l]),gdata[which(gdata[,2]==0),])

Qfun<-(Q2(c(alpha[i],lambda[j],beta0[k],
beta1[l]),gdata,pi.grid))

if (Qfun < min.Qfun) {min.Qfun<-Qfun
alpha.min<-alpha[i]
lambda.min<-lambda[j]
beta0.min<-beta0[k]
beta1.min<-beta1[l]
}
}
}
}
```

```
}
minima<-c(alpha.min,lambda.min,beta0.min,beta1.min)
return(minima)
}

################################## EM.ALGORITHM

EM<-function(gdata,epsilon)
{
abs.diff<-Inf
inits<-grid.em(gdata)
pi.e<-pi.em(inits,gdata[which(gdata[,2]==0),])
Q.old<-Q2(inits,gdata,pi.e)
k=0
while (abs.diff>epsilon) {
k<-k+1
pi.e<-pi.em(inits,gdata[which(gdata[,2]==0),])
out <- optim(inits, fn=Q2,method="L-BFGS-B",
lower= c(1e-3,1e-3,-Inf,-Inf),gdata=gdata,pi.e=pi.e)
inits<-out$par
Q.new<-out$value
abs.diff<-abs((Q.new-Q.old)/Q.old)
Q.old<-Q.new
output<-c(inits,k)
if(k>500){output<-NULL
epsilon<-Inf
}
}

return(output)
```

```
}
```

```
##############simulation with direct method

set.seed(123)

library(reliaR)

library(numDeriv)

######### case 1(b0=-1.231746,b1=0.3895951,alpha=2,lambda=3)

betas(.65,.25)

para1<-c(alpha=2,lambda=3,b0=-1.231746,b1=0.3895951)

p_all<-c(.85,.65,.5,.35)

pnot_all1<-c(0.65,0.5294055,0.3910264,0.25)

x_all<-c(1,2,3,4)

s_all<- c(30,40,20,10)

test11 <- matrix(nrow=1000 ,ncol=4)

ged.data1<-list()

H<-matrix(nrow=1000,ncol=4)

s<-matrix(nrow=1000,ncol=4)

count11<-0

count21<-0

count31<-0

count41<-0

count12<-0

count22<-0

count32<-0

count42<-0

po1<-c()

po2<-c()

po3<-c()
```

```
po4<-c()

d1<-c()
d2<-c()
d3<-c()
d4<-c()

e1<-c()
e2<-c()
e3<-c()
e4<-c()

for(i in 1:1000 ){
ged.data1[[i]]<-data.fun(p_all,pnot_all1,1000 ,2,3,x_all,s_all)
inits <- grid(ged.data1[[i]])
out <- optim(inits, fn=neg.loglikelihood,method="L-BFGS-B",
lower= c(1e-3,1e-3,-Inf,-Inf),gdata=ged.data1[[i]])
test11[i,] <- out$par               #######vector of estimates
H<- hessian(f=log.likelihood,x=test11[i,],data=ged.data1[[i]])
Hinv <- solve(H)

VAR1 <- -Hinv[1,1]
VAR2 <- -Hinv[2,2]
VAR3 <- -Hinv[3,3]
VAR4 <- -Hinv[4,4]
###
SE1 <- sqrt(VAR1)
SE2 <- sqrt(VAR2)
SE3 <- sqrt(VAR3)
SE4 <- sqrt(VAR4)
```

```
###95%CI

CI11<-c(test11[i,1]-1.96*SE1,test11[i,1]+1.96*SE1)

CI21<-c(test11[i,2]-1.96*SE2,test11[i,2]+1.96*SE2)

CI31<-c(test11[i,3]-1.96*SE3,test11[i,3]+1.96*SE3)

CI41<-c(test11[i,4]-1.96*SE4,test11[i,4]+1.96*SE4)

###counts to make coverage probabilities

if(para1[1]>=CI11[1] & para1[1]<=CI11[2] ){count11<-count11+1} else {count11<-count11+0}

if(para1[2]>=CI21[1] & para1[2]<=CI21[2] ){count21<-count21+1} else {count21<-count21+0}

if(para1[3]>=CI31[1] & para1[3]<=CI31[2] ){count31<-count31+1} else {count31<-count31+0}

if(para1[4]>=CI41[1] & para1[4]<=CI41[2] ){count41<-count41+1} else {count41<-count41+0}

######90%CI

CI12<-c(test11[i,1]-1.645*SE1,test11[i,1]+1.645*SE1)

CI22<-c(test11[i,2]-1.645*SE2,test11[i,2]+1.645*SE2)

CI32<-c(test11[i,3]-1.645*SE3,test11[i,3]+1.645*SE3)

CI42<-c(test11[i,4]-1.645*SE4,test11[i,4]+1.645*SE4)

###counts to make coverage probabilities

if(para1[1]>=CI12[1] & para1[1]<=CI12[2] ){count12<-count12+1} else {count12<-count12+0}

if(para1[2]>=CI22[1] & para1[2]<=CI22[2] ){count22<-count22+1} else {count22<-count22+0}

if(para1[3]>=CI32[1] & para1[3]<=CI32[2] ){count32<-count32+1} else {count32<-count32+0}

if(para1[4]>=CI42[1] & para1[4]<=CI42[2] ){count42<-count42+1} else {count42<-count42+0}

#########

s[i,]<-c(SE1,SE2,SE3,SE4)

#####squared difference of true and estimated parameter

d1[i]<-{test11[i,1]-para1[1]}^2
```

```
d2[i]<-{test11[i,2]-para1[2]}^2

d3[i]<-{test11[i,3]-para1[3]}^2

d4[i]<-{test11[i,4]-para1[4]}^2


###cured fraction for four groups

po1[i]<-exp(-exp(test11[i,3]+test11[i,4]))

po2[i]<-exp(-exp(test11[i,3]+2*test11[i,4]))

po3[i]<-exp(-exp(test11[i,3]+3*test11[i,4]))

po4[i]<-exp(-exp(test11[i,3]+4*test11[i,4]))

########squared difference of true and estimated cured fraction

e1[i]<-{po1[i]-pnot_all1[1]}^2

e2[i]<-{po2[i]-pnot_all1[2]}^2

e3[i]<-{po3[i]-pnot_all1[3]}^2

e4[i]<-{po4[i]-pnot_all1[4]}^2

}



est11 <- c(est.a=mean(test11[,1]), est.l=mean(test11[,2])

,est.b0=mean(test11[,3]),est.b1=mean(test11[,4]))

SE11 <- c(est.se.a=mean(s[,1]), est.se.l=mean(s[,2]),

est.se.b0=mean(s[,3]),est.se.b1=mean(s[,4]))

bias11<- est11-para1

CP95_11<-c(count11/1000,count21/1000,count31/1000,count41/1000)
```

```
CP90_11<-c(count12/1000,count22/1000,count32/1000,count42/1000)

RMSE11<- c(rmse.a=sqrt(sum(d1)/999),rmse.l= sqrt(sum(d2)/999),
rmse.b0= sqrt(sum(d3)/999),rmse.b1= sqrt(sum(d4)/999))

RMSE_p11<-c(rmse.po1= sqrt(sum(e1)/999),
rmse.po2= sqrt(sum(e2)/999),
rmse.po3= sqrt(sum(e3)/999),
rmse.po4= sqrt(sum(e4)/999))

###mean estimates of cured fraction for four groups
est.p11<-c(mean(po1),
mean(po2),
mean(po3),
mean(po4))

biasp11<- est.p11-pnot_all1
```

```
##EM simulation study


set.seed(123)

library(reliaR)

library(numDeriv)

######### case 1(b0=-1.231746,b1=0.3895951,alpha=2,lambda=3)

betas(.65,.25)

para1<-c(alpha=2,lambda=3,b0=-1.231746,b1=0.3895951)

p_all<-c(.85,.65,.5,.35)

pnot_all1<-c(0.65,0.5294055,0.3910264,0.25)

x_all<-c(1,2,3,4)

s_all<- c(30,40,20,10)

test11 <- matrix(nrow=1000 ,ncol=4)

ged.data1<-list()

H<-matrix(nrow=1000,ncol=4)

s<-matrix(nrow=1000,ncol=4)

count11<-0

count21<-0

count31<-0

count41<-0

count12<-0

count22<-0

count32<-0

count42<-0

po1<-c()

po2<-c()
```

```
po3<-c()
po4<-c()

d1<-c()
d2<-c()
d3<-c()
d4<-c()

e1<-c()
e2<-c()
e3<-c()
e4<-c()

iter<-c()
i=1
while(i<= 1000 ){
ged.data1[[i]]<-data.fun(p_all,pnot_all1,1000,2,3,x_all,s_all)
temp<- EM(ged.data1[[i]],epsilon=.001)
if(is.null(temp)==FALSE){test11[i,]<-temp[1:4]
iter[i]<-temp[5]

H<- hessian(f=log.likelihood,x=test11[i,],data=ged.data1[[i]])
Hinv <- solve(H)

VAR1 <- -Hinv[1,1]
VAR2 <- -Hinv[2,2]
VAR3 <- -Hinv[3,3]
VAR4 <- -Hinv[4,4]
###
if (min(c(VAR1,VAR2,VAR3,VAR4))>0){
SE1 <- sqrt(VAR1)
```

```
SE2 <- sqrt(VAR2)

SE3 <- sqrt(VAR3)

SE4 <- sqrt(VAR4)

###95%CI

CI11<-c(test11[i,1]-1.96*SE1,test11[i,1]+1.96*SE1)

CI21<-c(test11[i,2]-1.96*SE2,test11[i,2]+1.96*SE2)

CI31<-c(test11[i,3]-1.96*SE3,test11[i,3]+1.96*SE3)

CI41<-c(test11[i,4]-1.96*SE4,test11[i,4]+1.96*SE4)

###

if(para1[1]>=CI11[1] & para1[1]<=CI11[2] ){count11<-count11+1} else {count11<-count11+0}

if(para1[2]>=CI21[1] & para1[2]<=CI21[2] ){count21<-count21+1} else {count21<-count21+0}

if(para1[3]>=CI31[1] & para1[3]<=CI31[2] ){count31<-count31+1} else {count31<-count31+0}

if(para1[4]>=CI41[1] & para1[4]<=CI41[2] ){count41<-count41+1} else {count41<-count41+0}

######90%CI

CI12<-c(test11[i,1]-1.645*SE1,test11[i,1]+1.645*SE1)

CI22<-c(test11[i,2]-1.645*SE2,test11[i,2]+1.645*SE2)

CI32<-c(test11[i,3]-1.645*SE3,test11[i,3]+1.645*SE3)

CI42<-c(test11[i,4]-1.645*SE4,test11[i,4]+1.645*SE4)

###

if(para1[1]>=CI12[1] & para1[1]<=CI12[2] ){count12<-count12+1} else {count12<-count12+0}

if(para1[2]>=CI22[1] & para1[2]<=CI22[2] ){count22<-count22+1} else {count22<-count22+0}

if(para1[3]>=CI32[1] & para1[3]<=CI32[2] ){count32<-count32+1} else {count32<-count32+0}

if(para1[4]>=CI42[1] & para1[4]<=CI42[2] ){count42<-count42+1} else {count42<-count42+0}

#########
```

```
s[i,]<-c(SE1,SE2,SE3,SE4)

d1[i]<-{test11[i,1]-para1[1]}^2

d2[i]<-{test11[i,2]-para1[2]}^2

d3[i]<-{test11[i,3]-para1[3]}^2

d4[i]<-{test11[i,4]-para1[4]}^2

###

po1[i]<-exp(-exp(test11[i,3]+test11[i,4]))

po2[i]<-exp(-exp(test11[i,3]+2*test11[i,4]))

po3[i]<-exp(-exp(test11[i,3]+3*test11[i,4]))

po4[i]<-exp(-exp(test11[i,3]+4*test11[i,4]))

########

e1[i]<-{po1[i]-pnot_all1[1]}^2

e2[i]<-{po2[i]-pnot_all1[2]}^2

e3[i]<-{po3[i]-pnot_all1[3]}^2

e4[i]<-{po4[i]-pnot_all1[4]}^2

i=i+1

}

}

}

#test11

est11 <- c(est.a=mean(test11[,1]), est.l=mean(test11[,2]),

est.b0=mean(test11[,3]),est.b1=mean(test11[,4]))

SE11 <- c(est.se.a=mean(s[,1]), est.se.l=mean(s[,2]),

est.se.b0=mean(s[,3]),est.se.b1=mean(s[,4]))

bias11<- est11-para1
```

```
CP95_11<-c(count11/1000,count21/1000,count31/1000,count41/1000)

CP90_11<-c(count12/1000,count22/1000,count32/1000,count42/1000)

RMSE11<- c(rmse.a=sqrt(sum(d1)/999),rmse.1= sqrt(sum(d2)/999),
rmse.b0= sqrt(sum(d3)/999),rmse.b1= sqrt(sum(d4)/999))

RMSE_p11<-c(rmse.po1= sqrt(sum(e1)/999),
rmse.po2= sqrt(sum(e2)/999),
rmse.po3= sqrt(sum(e3)/999),
rmse.po4= sqrt(sum(e4)/999))

est.p11<-c(mean(po1),
mean(po2),
mean(po3),
mean(po4))

biasp11<- est.p11-pnot_all1
```

Model discrimination codes

```
############## the log likelihood function
# R code for automatically loading and installing required packages.
libPath = Sys.getenv("R_LIBS_USER")
if (!file.exists(libPath))
{
dir.create(libPath, recursive=TRUE)
}

# Set repository options:
local({r <- getOption("repos")
r["CRAN"] <- "http://cran.stat.sfu.ca/"
options(repos=r)})

# Set a vector of strings: package names to use (and install, if necessary)
pkg_list = c(‘survival’, ‘timereg’)

for (pkg in pkg_list)
{
# Try loading the library.
if ( ! library(pkg, logical.return=TRUE, character.only=TRUE, lib=libPath) )
{
# If the library cannot be loaded, install it; then load.
install.packages(pkg, lib=libPath)
library(pkg, character.only=TRUE, lib=libPath)
}
}

source("ged_exp_aic bic_melanoma.R")
source("exp_aic bic_melanoma.R")
```

```
data("melanoma")
mel.data<-data.matrix(melanoma)
t<-mel.data[,3]/365 ###columns for day ###survival time
I<-mel.data[,2]###columns for ulc ###censoring indicator
u<-mel.data[,4]
x<-mel.data[,5]/100###columns for tumour thickness ### category variable
a<-replace(I, I==2 | I==3, 0)

gdata<-matrix(cbind(t,a,x),ncol=3)

inits.exp1 <- grid.exp(gdata)
inits.exp1
out.exp1 <- optim(inits.exp1, fn=neg.loglikelihood.exp,
method="L-BFGS-B", lower= c(1e-3,-Inf,-Inf),gdata=gdata)
test.dat.exp1<- out.exp1$par
test.dat.exp1
library(numDeriv)
H11<- hessian(f=log.likelihood.exp,x=test.dat.exp1,data=gdata)
Hinv11 <- solve(H11)

em.dat.exp1<-EM.exp(gdata,epsilon=.001)
H1.em1<- hessian(f=log.likelihood.exp,x=em.dat.exp1[1:3],data=gdata)
Hinv1.em1 <- solve(H1.em1)


inits1 <- grid(gdata)
out1 <- optim(inits1, fn=neg.loglikelihood,method="L-BFGS-B",
lower= c(1e-3,1e-3,-Inf,-Inf),gdata=gdata)
test.dat1<- out1$par
library(numDeriv)
```

```
H1.g<- hessian(f=log.likelihood,x=test.dat1[2],data=gdata)

Hinv1.g <- solve(H1.g)

em.dat1g<-EM(gdata,epsilon=.001)

H1.g.em<- hessian(f=log.likelihood,x=em.dat1g[1:4],data=gdata)

Hinv1.g.em <- solve(H1.g.em)


aic.g<-(2*4-2*log.likelihood(test.dat1,gdata))

aic.exp<-(2*3-2*log.likelihood.exp(test.dat.exp1,gdata))

aic.g.em<-(2*4-2*log.likelihood(em.dat1g,gdata))

aic.exp.em<-(2*3-2*log.likelihood.exp(em.dat.exp1,gdata))

bic.g<-log(205)*4-2*log.likelihood(test.dat1,gdata)

bic.exp<-log(205)*3-2*log.likelihood.exp(test.dat.exp1,gdata)

bic.g.em<-log(205)*4-2*log.likelihood(em.dat1g,gdata)

bic.exp.em<-log(205)*3-2*log.likelihood.exp(em.dat.exp1,gdata)

LRT<-2*(log.likelihood(test.dat1,gdata)-log.likelihood.exp(test.dat.exp1,gdata))

LRT.em<-2*(log.likelihood(em.dat1g,gdata)-log.likelihood.exp(em.dat.exp1,gdata))
```

Code for QQ plot

```
datafile<-read.table("dataset1.txt",header=T)

t<-datafile$SURVTIME ### this is lifetimes

x<-datafile$NODES1 ### this is X

I<-datafile$SURVCENS  ### this is censoring indicator

a<-replace(I, I==1, 0)

d<-replace(a, a==2, 1)


dataset1<-matrix(c(t,d,x),ncol=3)


set.seed(123)


par(mfrow=c(2,2))


#########

b0=-1.2326388

b1=0.3647754

alpha=2.5086833

lambda=0.6845316


eta=exp(b0+b1*x)

F=(1-exp(-lambda*t))^alpha

s=1-F


Sp=exp(-eta*(1-(s)))


nrep = 5

mqresid = NULL

u = d * (1 - Sp) + (1 - d) * runif(length(t), 1 - Sp)

for (i in 1:nrep) {
```

```
qresid = sort(qnorm(runif(u)))
mqresid = cbind(mqresid, qresid)
}
qresid = apply(mqresid, 1, median)

ks.test(qresid, "pnorm")


qqnorm(qresid, pch = 20, main = "QQ plot for GED (direct method)", xlab = "N(0, 1) quantiles",
ylab = "Quantile residuals", cex.lab = 1.5,
cex.axis = 1.5, ylim = c(-3,3))

########################
##melanoma data II uses the similar code
```

Bibliography

Arnold, B. C., N. Balakrishnan, and H. N. Nagaraja (1992). *A first course in order statistics*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM, 3600 Market Street, Floor 6, Philadelphia, PA 19104). (Cited on page 5.)

Austin, P. C., D. S. Lee, and J. P. Fine (2016). Introduction to the analysis of survival data in the presence of competing risks. *Circulation 133*, 601–609. (Cited on page 9.)

Balakrishnan, N. and R. Aggarwala (2000). *Progressive censoring: theory, methods and applications*. Statistics for Industry and Technology. Birkhäuser Boston. (Cited on page 7.)

Balakrishnan, N. and E. Cramer (2014, 08). *The art of progressive censoring: applications to reliability and quality*. Birkhäuser Basel. (Cited on page 6.)

Balakrishnan, N. and S. Pal (2015). An EM algorithm for the estimation of parameters of a flexible cure rate model with generalized gamma lifetime and model discrimination using likelihood and information-based methods. *Computational Statistics 30*(1), 151–189. (Cited on page 12.)

Balakrishnan, N. and S. Pal (2016). Expectation maximization-based likelihood inference for flexible cure rate models with weibull lifetimes. *Statistical Methods in Medical Research 25*(4), 1535–1563. (Cited on pages 1, 13 and 34.)

Berkson, J. and R. P. Gage (1952, September). Survival curve for cancer patients following treatment. *Journal of the American Statistical Association 47*(259), 501–515. (Cited on page 10.)

Boag, J. W. (1949). Maximum likelihood estimates of the proportion of patients cured by cancer therapy . *Journal of the Royal Statistical Society.Series B (Methodological) 11*(1), 15–53. (Cited on page 10.)

Byrd, R. H., P. Lu, J. Nocedal, and C. Zhu (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing, 16*(5), 1190–1208. (Cited on page 33.)

Casella, G. and R. Berger (2002). *Statistical inference (Second Edition)*. The Wadsworth group. (Cited on pages 18 and 20.)

Celeux, G. and J. Diebolt (1985). The sem algorithm: a probabilistic teacher algorithm derived from the em algorithm for the mixture problem. *Computational Statistics Quarterly 2*, 73–82. (Cited on page 21.)

Chauveau, D. (1995). A stochastic em algorithm for mixtures with censored data. *Journal of Statistical Planning and Inference 46*, 1–25. (Cited on page 20.)

Dejardin, D. and E. Lesaffre (2013). Stochastic em algorithm for doubly interval-censored data. *Biostatistics 14*(4), 766–778. (Cited on page 22.)

Dempster, A. P., N. M. Lourd, and D. B. Rubin (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological), 39*(1), 1–38. (Cited on page 20.)

Diebolt, J. and E. Ip (1994, September). A stochastic em algorithm for approximating the maximum likelihood estimate. techreport 301, Department of Statistics, Stanford University. (Cited on page 21.)

Dunn, P. K. and G. K. Smyth (1996). Randomized quantile residuals. *Journal of Computational and Graphical Statistics, 5*(3), 236–244. (Cited on page 53.)

Efron, B. and D. V. Hinkley (1978, December). Assessing the accuracy of the maximum likelihood estimator: observed versus expected fisher information. *Biometrika 65*(3), 457–482. (Cited on page 19.)

Eryilmaz, S. (2018). Modeling systems with two dependent components under bivariate shock models. *Communications in Statistics - Simulation and Computation*, 1–15. (Cited on page 9.)

Feizjavadian, S. and R. Hashemi (2015). Analysis of dependent competing risks in the presence of progressive hybrid censoring using marshall–olkin bivariate weibull distribution. *Computational Statistics and Data Analysis 82*, 19–34. (Cited on page 10.)

Finkelstein, M. (2008). *Failure rate modelling for reliability and risk*. Springer-Verlag London Limited. (Cited on page 3.)

Gross, A. J. and V. A. Clark (1975). *Survival distributions: reliability applications in the biomedical sciences*. John Wiley & Sons,Inc. (Cited on page 8.)

Gupta, R. and D. Kundu (2000). Generalized exponential distribution: Different method of estimations. *Journal of Statistical Computation and Simulation*, 1–22. (Cited on page 17.)

Gupta, R. D. and D. Kundu (2007). Generalized exponential distribution : Existing results and some recent developments. *Journal of Statistical Planning and Inference 137*, 3537–3547. (Cited on page 17.)

Gåsemyr, J. and B. Natvig (1995). Some aspects of reliability analysis in shock models. *Scandinavian Journal of Statistics 22*(3), 385–393. (Cited on page 10.)

Hong, Y. and W. Q. Meeker (2014). Confidence interval procedures for system reliability and applications to competing risks models. *Lifetime data analysis 20*, 161–184. (Cited on page 9.)

Ishioka, T. and Y. Nonaka (1991, April). Maximum likelihood estimation of weibull parameters for two independent competing risks. *IEEE Transaction on reliability 40*(1). (Cited on page 9.)

Kannan, N., D. Kundu, P. Nair, and R. C. Tripathi (2010). The generalized exponential cure rate model with covariates. *Journal of Applied Statistics 37*(10), 1625–1636. (Cited on page 12.)

Kirkwood, J. M., J. G. Ibrahim, V. K. Sondak, J. Flaherty, L. Richards, M. Ernstoff, T. Smith, U. Rao, M. Steele, and R. Blum (2000). High- and low-dose interferon alfa-2b in high-risk melanoma: first analysis of intergroup trial e1690/s9111/c9190. *Journal of Clinical Oncology 18*, 2444–2458. (Cited on page 52.)

Kundu, D. and S. Basu (1991). Analysis of incomplete data in presence of dependent competing risks. *American Mathematical Society*. (Cited on page 9.)

Kundu, D. and S. Basu (2000). Analysis of incomplete data in presence of dependent competing risks. *Journal of Statistical Planning and Inference 87*, 221–239. (Cited on page 10.)

Kundu, D. and R. D. Gupta (1999). Generalized exponential distributions . *Australian and New Zealand Journal of Statistics 41*(2), 173–188. (Cited on page 14.)

Kundu, D. and R. D. Gupta (2008). Generalized exponential distribution : bayesian estimations . *Computational Statistics & data analysis 52*, 1873–1883. (Cited on page 17.)

Lawless, J. (2003). *Statistical models and methods for lifetime data (Second Edition)*. John Wiley & Sons, Inc., Hoboken, New Jersey. (Cited on pages 3, 4 and 8.)

Markiewicz, I., W. G. Strupczewski, E. Bogdanowicz, and K. Kochanek (2015). Generalized exponential distribution in flood frequency analysis for Polish rivers. *PLoS ONE 10(12): e0143965*. (Cited on page 17.)

Miyakawa, M. (1982). Statistical analysis of incomplete data in competing risks model. *Journal of the Japanese Society of Quality Control 12*, 49–52. (Cited on page 9.)

Moeschberger, M. L. (1974). Life tests under dependent competing causes of failure. *Technometrics 16*(1), 39–47. (Cited on page 9.)

Moeschberger, M. L. and J. P. Klein (1995). Statistical methods for dependent copmpeting risks. *Lifetime data analysis 1(2)*, 195–204. (Cited on page 10.)

Pal, S. and N. Balakrishnan (2016). An EM type estimation procedure for the destructive exponentially weighted Poisson regression cure model under generalized gamma lifetime. *Journal of Statistical Computation and Simulation*, 1–23. (Cited on pages 1, 2, 13, 34 and 37.)

Park, C. (2005). Parameter estimation of incomplete data in competing risks using the em algorithm. *IEEE transactions on reliability 54*(2). (Cited on page 21.)

Pascual, F., W. Meeker, and L. Escobar (2006). *Springer handbook of engineering statistics*, Chapter Accelerated life test models and data analysis, pp. 397–426. Springer. (Cited on page 5.)

Prentice, R. L., J. D. Kalbfleisch, P. A. V., N. Flourney, and V. T. Farewell (1978). The analysis of failure times in the presence of competing risks. *Biometrics 34*(4), 541–554. (Cited on page 8.)

Raqab, M. Z. and M. T. Madi (2006). Bayesian inference for the generalized exponential distribution. *Journal of Statistical Computation and Simulation 75*(10), 841–852. (Cited on page 17.)

Rodrigues, J., M. d. Castro, V. G. Cancho, and N. Balakrishnan (2009, October). COM-Poisson cure rate survival models and an application to cutaneous melanoma data. *Journal of Statistical Planning And Inference 139*(10), 3605–3611. (Cited on pages 11 and 55.)

Shi, Y. and M. Wu (2016). Statistical analysis of dependent competing risks model from gompertz distribution under progressively hybrid censoring. *SpringerPlus 5*(1745). (Cited on page 10.)

Sy, J. P. and J. M. G. Taylor (2000). Estimation in a Cox proportional hazards cure model. *Biometrics* (March), 227–236. (Cited on page 11.)

Wu, M., Y. Shi, and C. Zhang (2017). Statistical analysis of dependent competing risks model in accelerated life testing under progressively hybrid censoring using copula function. *Communications in Statistics - Simulation and Computation 46*(5), 4004–4017. (Cited on page 10.)

Yang, Y., H. K. T. Ng, and N. Balakrishnan (2016). A stochastic expectation-maximization algorithm for the analysis of system lifetime data with known signature. *Computational Statistics 31*, 609–641. (Cited on page 22.)

Yin, G. and J. G. Ibrahim (2005). Cure rate models: a unified approach. *The Canadian Journal of Statistics 33*(4), 559–570. (Cited on pages 11 and 52.)

Zhang, M., Z. Ye, and M. Xie (2013). A stochastic em algorithm for progressively censored data analysis. *Quality and Reliability Engineering International 30*(5), 711–722. (Cited on page 22.)